

МАТЕМАТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ СЛОЖНОГО ЭКСПЕРИМЕНТА

ПРОБЛЕМЫ
ПОСТРОЕНИЯ
МАТЕМАТИЧЕСКОГО
И ПРОГРАММНОГО
ОБЕСПЕЧЕНИЯ
ИЗМЕРИТЕЛЬНО-
ВЫЧИСЛИТЕЛЬНЫХ
КОМПЛЕКСОВ

5

МАТЕМАТИЧЕСКОЕ

ОБЕСПЕЧЕНИЕ

СЛОЖНОГО

ЭКСПЕРИМЕНТА

**ПРОБЛЕМЫ
ПОСТРОЕНИЯ
МАТЕМАТИЧЕСКОГО
И ПРОГРАММНОГО
ОБЕСПЕЧЕНИЯ
ИЗМЕРИТЕЛЬНО-
ВЫЧИСЛИТЕЛЬНЫХ
КОМПЛЕКСОВ**

5

МАТЕМАТИЧЕСКОЕ
ОБЕСПЕЧЕНИЕ
СЛОЖНОГО
ЭКСПЕРИМЕНТА

МАТЕМАТИЧЕСКОЕ

**ИЗДАНИЕ
В ПЯТИ
ТОМАХ**

КИЕВ НАУКОВА ДУМКА 1990

ОБЕСПЕЧЕНИЕ

СЛОЖНОГО

ЭКСПЕРИМЕНТА

**ПРОБЛЕМЫ
ПОСТРОЕНИЯ
МАТЕМАТИЧЕСКОГО
И ПРОГРАММНОГО
ОБЕСПЕЧЕНИЯ
ИЗМЕРИТЕЛЬНО-
ВЫЧИСЛИТЕЛЬНЫХ
КОМПЛЕКСОВ**

5

УДК 519.95 : 518.0 : 621.391 : 681.325

Математическое обеспечение сложного эксперимента : В 5 т. /
Под общ. ред. И. И. Ляшко. Т. 5. Проблемы построения математи-
ческого и программного обеспечения измерительно-вычислительных
комплексов / Белов Ю. А., Егоров Б. М., Козлов Н. Н., Ляш-
ко И. И., Макаров В. Л. — Киев : Наук. думка, 1990. — 368 с. —
ISBN 5-12-000387-7, ISBN-12-000388-5.

В монографии обсуждаются проблемы, возникающие в связи с развитием измерительно-вычислительных комплексов, которые находят применение в различных областях науки и техники. Рассматриваются вопросы разработки и обоснования методов анализа данных и алгоритмов, входящих в состав специального математического обеспечения измерительно-вычислительного комплекса, а также эргономический анализ деятельности оператора автоматизированных систем. Приводится комплекс программ статистической обработки результатов качественных измерений.

Для специалистов в области теоретической и прикладной кибернетики, прикладной математики, вычислительной техники и всех интересующихся вопросами использования математических методов при анализе и синтезе сложных систем.

Ил. 70. Табл. 18. Библиогр.: 361=364 (89 назв.).

Авторы:

Ю. А. БЕЛОВ, Б. М. ЕГОРОВ, Н. Н. КОЗЛОВ,
И. И. ЛЯШКО, В. Л. МАКАРОВ

*Утверждено к печати ученым советом Киевского
государственного университета им. Т. Г. Шевченко*

Редакция математики и механики

Редактор Т. С. Мельник

2404000000-468
М. _____ подписное
M221 (04)-90

ISBN 5-12-000387-7
ISBN 5-12-000388-5

© Ю. А. Белов, Б. М. Егоров, Н. Н. Козлов,
И. И. Ляшко, В. Л. Макаров, 1990

ПРЕДИСЛОВИЕ

*Памяти наших товарищей и коллег
В. П. Диденко и О. Е. Цитрицкого
посвящается.*

Настоящая книга продолжает цикл монографий, освещающих вопросы разработки математического и программного обеспечения современных измерительно-вычислительных комплексов (ИВК). Следует отметить, что среди задач математического и технического обеспечения ИВК сегодня можно выделить традиционные и нетрадиционные. И если первые хорошо известны специалистам в данной предметной области, то вторые задачи возникли совсем недавно, с одной стороны, в связи с бурным развитием таких направлений, как автоматизация проектирования и испытаний ИВК, с другой — в результате внедрения методологии и технологии математического моделирования на ЭВМ. Эти задачи связаны с формализацией процессов проектирования и испытаний как процессов принятия решений, например при формировании облика ИВК.

В настоящее время можно говорить о том, что созданы основы теории принятия проектных решений, которая существенно использует методы системного анализа и многокритериальной оптимизации, математического и имитационного моделирования. Много в этом направлении сделано коллективами советских ученых, возглавляемых П. С. Краснощековым, В. С. Михалевичем, Н. Н. Моисеевым, Г. С. Поспеловым, А. А. Самарским, В. В. Федоровым и др.

Данный том является пятым в серии «Математическое обеспечение сложного эксперимента», и в некотором смысле можно считать, что им подытоживается первый этап исследований. Напомним, что целью издания пятитомника было изложение основных математических проблем и методов их решения, возникающих в связи с развитием ИВК и находящих широкое применение в различных областях науки и техники. В смысле поставленной цели авторы считают свою задачу выполненной. Однако следует отметить, что к моменту завершения оформления рукописи монографии нами совместно с учениками получен ряд результатов, позволяющих по-новому взглянуть как на некоторые вопросы, изложенные ранее, так и на проблему в целом. Завершена первая версия разработки инструментальных средств проектирования и оценки эффективности ИВК.

Следует отметить, что наши исследования имеют ярко выраженный системный, междисциплинарный характер, что также наложило отпечаток на содержание и этого тома.

Авторы выражают благодарность А. Н. Нечаеву за проведенную работу по подготовке комплекса программ статистической обработки результатов качественных измерений для публикации, В. Н. Матвиенко за помощь при оформлении рукописи, а также товарищам, принимавшим участие в ее обсуждении

Авторы

МЕТОДЫ И АЛГОРИТМЫ ОБРАБОТКИ ИНФОРМАЦИИ В ИЗМЕРИТЕЛЬНО- ВЫЧИСЛИТЕЛЬНОМ КОМПЛЕКСЕ

§ 1. Специальное математическое обеспечение измерительно-вычислительного комплекса

Сокращение временных и экономических затрат на испытания сложных технических систем (СТС) возможно только на основе получения из каждого, как правило, дорогостоящего, порой уникального, эксперимента максимального количества информации, а также на основе полной автоматизации процесса обработки. Достижение этих целей можно обеспечить в настоящее время как совершенствованием технических средств ИВК, так и совершенствованием математического обеспечения (МО).

Рассмотрим состав, структуру и принципы построения системы обработки траекторной измерительной информации (ТИИ) на примере ИВК, решающего задачи определения параметров движения летящих объектов.

Под системой МО будем понимать всю совокупность алгоритмов, программ, программной документации, обеспечивающих автоматическое или автоматизированное функционирование системы обработки ИВК.

Математическое обеспечение подразделяется на общее (ОМО) и специальное (СМО). Программное обеспечение ОМО предназначено для организации вычислительного процесса и, как правило, ориентируется на конкретный тип ЭВМ. Оно обычно поставляется в комплекте с вычислительными средствами, куда входят программы, необходимые в процессе функционирования и развития системы МО, программы для автоматизации разработки программ, организации функционирования ЭВМ, а также другие служебные и стандартные программы (организующие, транслирующие и т. д.).

Программное обеспечение СМО предназначено для ввода в ЭВМ исходной измерительной информации, оперативного анализа и первичной ее обработки, совместной статистической обработки разнородной ТИИ с целью определения эталонных параметров движения летательных аппаратов (ЛА), формирования выходных результатов, информационного обеспечения системы обработки в целом, а также для управления процессом математической обработки на ЭВМ в различных режимах.

В связи с этим чрезвычайно большое значение приобретают вопросы развития МО и в первую очередь СМО, без которого невозможно использование современных ЭВМ для решения задач, стоящих перед ИВК.

Требования к СМО зависят от большого числа факторов. Основными из них являются следующие:

разнородность входного потока измерительной информации как по составу, так и по точности измерения параметров движения ЛА;

наличие различных типов ЛА;

необходимость максимального учета и использования всех аппаратных и программных возможностей, представляемых современными вычислительными средствами;

позышение требований к точности определения оценок эталонных параметров движения ЛА с одновременным сокращением сроков представления конечных результатов обработки;

возрастание требований к уровню автоматизации процесса сбора и обработки измерительной информации;

дальнейшее развитие математических методов обработки результатов натурных экспериментов и описания траекторий движения летательных аппаратов различных типов;

расширение круга задач по информационному обеспечению различных подсистем ИВК.

На основании изложенного можно заключить, что СМО является одним из наиболее важных элементов ИВК, а реализация указанных факторов требует тщательной разработки теории специального математического обеспечения.

Измерительная информация, полученная в ходе летного эксперимента, имеет ряд особенностей, которые необходимо учитывать при решении задач автоматизации процессов обработки. Основными из них являются:

большое разнообразие носителей информации (перфолисты, перфокарты, магнитные ленты, кинофотоплёнки, цифропечатать на бумажной ленте);

циклический характер изменения параметров, регистрируемых радиотехническими средствами;

структурная и временная избыточность измерений;

наличие случайных, медленно меняющихся и систематических составляющих погрешностей измерений;

косвенный характер и отсутствие наглядности представления результатов измерений;

неодновременное поступление измерительной информации в обработку.

Указанные особенности позволяют выделить несколько самостоятельных этапов математической обработки ТИИ, включающих в себя предварительную обработку и ввод информации в ЭВМ, первичную и вторичную (совместную) обработки и формирование требуемой выходной информации.

Разработку конкретных алгоритмов первичной и вторичной обработок необходимо осуществлять с учетом моделей ЛА, состава измерительных средств, частоты поступающей измерительной информации. Кроме того, эти алгоритмы должны обеспечивать возможность использования радиотелеизмерений, позволяющих повысить точность и достоверность результатов обработки. Однако в этом случае из-за боль-

шого различия в точности траекторных измерений и данных телеметрии корреляционные матрицы погрешностей оказываются близкими к вырожденным, что приводит к неустойчивости алгоритмов оценивания. Это требует использования известных методов регуляризации и устойчивых процедур совместного оценивания.

Модели движения, используемые в алгоритмах обработки, необходимо разрабатывать с учетом типов ЛА и требований к точности оценивания параметров движения. Они должны достаточно точно отражать реальное движение объектов и быть простыми для реализации на ЭВМ.

Система СМО должна решать задачи информационного обеспечения подсистем ИВК и процесса математической обработки информации на ЭВМ, обеспечивать длительное хранение, автоматизированный поиск и выдачу в требуемом виде исходной информации, результатов обработки и справочных данных. При разработке отдельных алгоритмов необходимо учитывать возможность организации взаимообмена информацией как непосредственно между отдельными модулями, так и централизованно через информационную подсистему. Для этого информационная подсистема должна быть открытой, удовлетворять требованиям экономии памяти, обеспечивать выдачу данных с наибольшей скоростью и гибкость процесса математической обработки информации.

Операционные системы современных ЭВМ позволяют осуществлять процесс обработки с помощью формализованного задания на обработку, а также проводить автоматически или в диалоговом режиме проверку готовности средств системы обработки. В свою очередь, и программно-реализованное СМО должно обеспечивать возможность реализации многоуровневых итерационных циклов с целью выбора оптимального процесса обработки, контроль и управление со стороны оператора процесса прохождения задания на обработку в системе; проводить одновременную обработку измерительной информации нескольких заданий; осуществлять протоколирование прохождения задания на обработку с выдачей необходимых сообщений.

Состав программ СМО обработки ТИИ должен обеспечивать решение задач в режимах экспресс-обработки, оперативной и полной обработки.

Ввиду многообразного характера использования элементов СМО разработку СМО целесообразно проводить путем создания автономных программных модулей, прикладных программ и программных комплексов, отвечающих системным требованиям.

Для реализации системного подхода программы, включаемые в состав СМО, должны иметь модульную структуру, обеспечивать оперативность применения. Необходимо, чтобы информационное сопряжение с другими программами СМО осуществлялось через общие информационные базы данных, а сходная информация справочного характера запрашивалась программой из базы справочных данных.

Из наборов модулей в соответствии с заданием на обработку формируются технологические цепочки обработки ТИИ. Они и реализуются на ЭВМ при решении конкретных задач.

Совершенствование и модернизация структуры и состава СМО осуществляются путем замены модулей без нарушения основной структуры самого СМО. Модульный принцип позволяет обеспечить эффективную и качественную разработку, отладку, эксплуатацию, корректировку и расширение программного обеспечения СМО при широком использовании программных средств ЭВМ.

Алгоритмы СМО должны обладать свойством адаптивности к различным факторам, возникающим при проведении испытаний, т. е. способностью выполнять заданные функции при возможном изменении исходных данных в определенных пределах и отвечать требованиям алгоритмической, информационной и программной надежности. Вместе с тем важно, чтобы и сама система СМО была адаптивной к изменению версий ОС, поля внешних устройств и т. д.

Математическое обеспечение содержит средства автоматизированного технологического комплекса разработки, отладки и документирования программ. Использование этих средств позволяет значительно сократить сроки создания СМО.

Немаловажным требованием является обеспечение рационального соотношения между функциональными возможностями, обозримостью и простотой системы. Необходимо также минимизировать время, затрачиваемое на изучение ее возможностей и подготовку конкретных специалистов.

И, наконец, СМО должно предоставлять широкие возможности по проведению паспортизации измерительных средств, отдельных трасс полета и измерительного комплекса в целом, содержать набор комплексных моделей, обеспечивающих решение задач моделирования и планирования работы ИВК, для определения путей развития, совершенствования и оптимизации структуры ИВК.

Режимы обработки траекторной измерительной информации. Наряду с высокими требованиями к точности и надежности получения результатов обработки ТИИ предъявляются всевозрастающие требования к пропускной способности и срокам обработки результатов летных экспериментов.

Возможности по точности нахождения параметров движения и срокам их получения определяются тактико-техническими характеристиками траекторных измерительных средств и комплекса в целом. К их числу относятся:

- точность определения параметров траектории ЛА отдельными измерительными средствами;

- количество ЛА, одновременно участвующих в летном эксперименте;

- пропускная способность радиотехнических и оптических измерительных средств;

- условия проведения летного эксперимента (время суток, класс ЛА и состав бортовой аппаратуры траекторных измерений, трассы полета и т. п.);

- состав и размещение измерительного комплекса;

- пропускная способность системы передачи данных (СПД) и других видов доставки информации в центр обработки;

возможности вычислительных средств центра обработки информации.

В зависимости от целей летного эксперимента и требований к срокам представления результатов в системе обработки реализуются режимы оперативной и полной обработки.

Оперативная обработка проводится для качественной оценки функционирования испытываемых средств, выявления и локализации неисправностей и принятия решения о проведении или переносе очередного этапа летного эксперимента.

В зависимости от состава измерительных средств, возможностей системы передачи данных и вычислительных средств оперативная обработка может осуществляться в темпе поступления информации от одного из измерительных средств или через небольшой промежуток времени по информации, переданной от ограниченного числа средств.

Основными источниками ТИИ в режиме оперативной обработки являются измерительные средства, имеющие в своем составе специализированные ЭВМ, позволяющие осуществлять обработку и передачу информации в реальном масштабе времени. Кроме того, в радиотехнических средствах автоматически осуществляется ввод различного рода поправок, что значительно сокращает сроки обработки информации этих средств в центре обработки и соответственно повышает точность вычисляемых параметров.

К комплексу алгоритмов и программ, обеспечивающих решение задач обработки в оперативном режиме, предъявляются высокие требования по достоверности получаемых результатов и их быстродействию на ЭВМ.

Полная обработка проводится с целью количественной оценки характеристик испытываемых систем, диагностики аномальных ситуаций в их работе. Она осуществляется в двух вариантах: в первом варианте совместно обрабатываются все измерения, поступившие от измерительных средств; во втором варианте, реализуемом при наличии больших ресурсов времени и высоких требований к точностным характеристикам результатов, совместно обрабатываются все полученные измерения.

Для режима полной обработки в состав СМО должны входить алгоритмы, позволяющие производить совместную статистическую обработку всей совокупности разнородных измерений.

Структура, состав и задачи, решаемые системой специального математического обеспечения. Эффективность функционирования ИВК, как уже отмечалось, зависит от характеристик измерительных средств, их территориального расположения относительно трасс полета ЛА, а также от характеристик алгоритмов системы СМО и их вычислительной структуры. Для успешного решения широкого круга задач математической обработки результатов измерений, отображения и формирования выходных данных, управления процессом обработки в ЭВМ, информационного обеспечения ИВК, решения обеспечивающих задач система СМО должна иметь сложную иерархическую структуру и включать в себя следующие подсистемы: ввода информации в ЭВМ; первичной обработки; вторичной обработки; информационную; управле-

ния процессом обработки; формирования выходных результатов; обеспечивающих программ; создания и испытаний СМО.

Рассмотрим назначение и задачи, которые необходимо решать каждой подсистемой, чтобы обеспечить выполнение требований, предъявляемых к СМО.

Подсистема ввода информации в ЭВМ непосредственно решает задачи ввода информации, передаваемой по каналам СПД или зарегистрированной на различных носителях (перфокартах, перфоленте, магнитной ленте и др.), а также специализированных программ экспресс-анализа информации, отбраковки недостоверной информации, распаковки и записи в информационную подсистему. Особенно жесткие требования предъявляются к быстродействию канальных программ, обеспечивающих ввод информации из каналов СПД в режиме оперативной обработки.

Подсистема первичной обработки предназначена для преобразования измеренных параметров в числовые массивы значений параметров траектории в системах координат измерительных средств с привязкой к шкале системы единого времени (СЕВ), а также определения реальных статистических характеристик результатов измерения.

Учитывая существенные различия физических принципов получения измерений различными измерительными средствами, рассмотрим задачи подсистемы первичной обработки по двум направлениям:

- 1) первичную обработку информации радиотехнических измерительных средств;
- 2) первичную обработку информации оптических измерительных средств.

В общем случае после ввода информации в ЭВМ основные этапы первичной обработки осуществляются в такой последовательности: логическая, математическая обработка информации; накопление результатов первичной обработки в базе данных информационной подсистемы с последующей, при необходимости, регистрацией на различных накопителях; отображение и документирование результатов первичной обработки в виде графиков, таблиц и т. д.

Радиотехнические средства в силу независимости от погодных условий, высокого уровня автоматизации позволяют оперативно получать и передавать в центр обработки измерительного комплекса высокоточную траекторную информацию. Кроме того, в режиме воспроизведения рассчитываются параметры траектории движения ЛА, которые также могут быть переданы в центр обработки.

С этой точки зрения в структуре организации процесса обработки можно выделить следующие три подхода:

- 1) последовательную обработку информации, передаваемой в центр обработки;
- 2) централизованную обработку;
- 3) децентрализованную обработку.

При последовательной обработке измеренные параметры в режиме реального масштаба времени (РМВ) передаются и обрабатываются в центре обработки.

Сущность централизованной обработки информации заключается в том, что на отдельных средствах измеренные параметры регистрируются на различных накопителях и затем в режиме воспроизведения передаются в центр обработки. Определение параметров траектории движения ЛА производится после сбора измерительной информации всех средств.

При децентрализованной схеме обработки на измерительных средствах, имеющих в своем составе ЭВМ, в режиме послеполетной обработки определяются параметры траектории движения и записи передаются в центр обработки. С точки зрения сокращения количества операций в центре обработки предпочтительным является подход, основанный на децентрализованной обработке, однако достичь требуемых точностей получаемых оценок в ряде случаев не удастся, так как в этом случае теряется ряд преимуществ, которые дают структурная избыточность измерений и известный метод уравнивания неравноточных измерений. Поэтому в случаях, когда речь идет о максимальной точности оценивания, необходимо в совместной обработке использовать сами измеренные параметры, т. е. использовать подход централизованной обработки. Кроме того, первичная обработка в центре обработки позволяет повысить достоверность и точность результатов за счет устранения ряда ограничений алгоритмов, реализованных в специализированных ЭВМ измерительных средств и работающих в РМВ, применения более современных методов статистической обработки, использования для решения задач отбраковки аномальных измерений геометрических связей между первичными траекторными параметрами и структурной избыточностью измерений.

Основными задачами первичной обработки измерений радиотехнических средств являются:

- обработка служебной, тестовой и калибровочной информации, определение нулевых уровней отсчета и калибровочных коэффициентов; спрямление циклических параметров;

- статистическая обработка спрямленной информации измерительных шкал (проверка на достоверность, отбраковка, сглаживание, восстановление);

- приведение отсчетов к единым моментам времени по шкале СЕВ;

- расчет и ввод поправок и определение значений первичных траекторных параметров;

- формирование выходных массивов первичной обработки и их запись в информационную подсистему.

В зависимости от режима обработки меняют объем и содержание операций первичной обработки, функциональные связи между ее программными модулями.

К первичной обработке траекторной информации в темпе их поступления предъявляются жесткие временные требования, так как все вычисления должны быть выполнены до поступления очередного информационного кадра.

Получение первичных измерений оптических средств — наиболее трудоемкий и длительный по времени процесс, поэтому эти измерения, как правило, используются в режиме полной обработки.

Подсистема алгоритмов вторичной обработки представляет собой совокупность наиболее сложных алгоритмов СМО. Ее основное назначение — определение эталонных траекторий ЛА всех типов.

Алгоритмы, входящие в подсистему, должны основываться на современных математических методах совместной обработки траекторных измерений, полученных от разнотипных измерительных средств ИВК и уже прошедших этап первичной обработки. В ряде случаев вместе с траекторными измерениями в совместную обработку необходимо включать соответствующую телеметрическую траекторную информацию (о величинах перегрузок, переключения рулей и т. д.), получаемую по запросам от системы сбора и обработки телеметрической информации. Наконец, в алгоритмах, входящих в подсистему, должна использоваться хранящаяся в информационной подсистеме априорная информация о характере движения наблюдаемого ЛА, точностных характеристиках погрешностей измерений, характеристиках внешней среды и т. д.

Результаты обработки — эталонные значения искомых параметров, полученные в подсистеме, передаются для хранения в информационную подсистему и подсистему формирования выходных результатов, где они преобразуются к виду, удобному для потребителя.

Подсистема вторичной обработки должна обеспечивать решение своих задач в режимах оперативной и полной обработки.

Алгоритмы, входящие в подсистему, могут использоваться также при решении задач послеполюсного статистического анализа точностных характеристик измерений, полученных от средств ИВК, при моделировании процессов измерений и т. д.

В соответствии с принципами построения СМО рассматриваемая подсистема должна быть открытой: при совершенствовании существующих и внедрении новых методов и алгоритмов обработки траекторных измерений состав алгоритмов в подсистеме может постоянно расширяться. В результате для решения каждой из задач, возлагаемых на СМО, должны привлекаться алгоритмы, которые больше других соответствуют конкретным реальным условиям проведения эксперимента.

Для обработки траекторных измерений из алгоритмов, входящих в подсистему, должны формироваться технологические цепи. Предусматривается два типа технологических цепей: цепи, реализующие неименные последовательности обработки (жесткие пакеты), многократно проведенные, предназначенные для обработки траекторных измерений в стандартных условиях, и переменные цепи, которые могут изменяться и уточняться в процессе обработки в новых, изменившихся условиях.

Цепи выбираются или строятся на основе полученного задания на обработку. Выбор алгоритмов для включения их в технологическую цепь производится в соответствии с инструкциями и рекомендациями по их применению, которые могут дополняться и совершенствоваться в процессе проведения экспериментов.

В подсистеме реализуются различные способы описания движения ЛА: дифференциальные уравнения движения, дискретная, полиномиальная и сплайновая аппроксимации.

В подсистеме наряду с алгоритмами оценивания неизвестных параметров, основанными на методе наименьших квадратов, должны быть реализованы алгоритмы оценивания, основанные на современных робастных методах математической статистики.

Подсистема должна также включать в себя группы алгоритмов, реализующих различные методы: коррекцию систематических погрешностей траекторных измерений; идентификацию оптических траекторных измерений; интегрирование уравнений движения; минимизацию функций многих переменных; решение плохо обусловленных систем линейных алгебраических уравнений; вычисление частных производных от измеряемых величин по определяемым параметрам; динамическую фильтрацию.

Одним из основных элементов автоматизированной системы обработки траекторной информации является информационная подсистема, которая предназначена для информационного обслуживания всех подсистем с целью формирования, регистрации, централизованного хранения и коллективного использования информации, отражающей состояние и функционирование как подсистем обработки, так и необходимых элементов ИВК.

Информационные потоки траекторной информации имеют достаточно сложную структуру, что обусловлено разнородностью измерительных средств и сложностью организации обработки измерительной информации, требующей привлечения больших объемов служебной, справочной и управляющей информации.

Информационная подсистема должна удовлетворять следующим основным требованиям:

- быть открытой (расширяемой); это требование связано с возможностью поэтапного развития и с появлением новых функций информационного обеспечения задач обработки;

- обеспечивать заявку на выдачу данных с наибольшей скоростью; производить процедуры обмена информацией в плановом (из прикладных программ) и автономном (в виде запросов) режимах;

- обеспечивать хранение баз данных, входящих в подсистему, на носителях, совместимых с ЭВМ, на которой реализована система управления данными;

- быть интегрированной, т. е. должен соблюдаться принцип единой информационной базы;

- удовлетворять требованию экономии памяти;

- обеспечивать независимость пользователей от физической организации данных;

- обеспечивать средствами защиты данные многократного использования от искажения.

Для разработки информационной подсистемы, удовлетворяющей указанным требованиям, необходимо решить следующие основные задачи:

- определить состав и структуру баз данных, учитывая число исходных данных, их организацию и связь с другими массивами;

- определить технические средства для хранения баз данных;

- выбрать единый метод идентификации и кодирования информации, так как информационная подсистема может выдавать лишь правиль-

но идентифицированную информацию, а способ кодирования должен располагать резервом кодов для идентификации новых элементов;

определить способ хранения данных по каждой базе и метод доступа, разработав стандартные документы по вводу (выводу) данных в соответствующую базу (из базы) и приведя полный их перечень;

разработать язык запросов для обслуживания потребителей-программистов в автономном режиме, выявив типовые запросы и зафиксировав их в специально разработанных блоках, предусмотрев возможность реализации запросов в ЭВМ в диалоговом режиме;

установить основные функции системы управления базами данных, которые на ЭВМ реализуются в виде программ.

Существует два пути создания информационной подсистемы в рамках системы обработки траекторной информации на ЭВМ:

1) автономное построение на основе использования возможностей, предоставляемых ОС ЭВМ с учетом особенностей обслуживаемых потоков информации;

2) использование разработанных на базе ОС автоматизированных банков данных (АБД), которые в большей части удовлетворяют требованиям, предъявляемым к информационной подсистеме.

Первый путь требует квалифицированной проработки большого комплекса вопросов, связанных с организацией структуры данных, определения принципов и разработки методов размещения, хранения, поиска и др. Второй путь — более эффективен и требует разработки рациональных структур данных, учитывающих особенности использования информации в системе обработки, однако предъявляет повышенные требования по комплектации вычислительного комплекса.

Подсистема управления процессом математической обработки предназначена для управления программными средствами системы СМО. На нее возлагается решение следующих задач:

обеспечение функционирования системы СМО в диалоговом режиме;

проведение оперативного анализа исходной информации и планирование работы системы обработки на этапах первичной и вторичной обработки и формирования выходных результатов;

ввод заданий на математическую обработку информации, оформленных на проблемно-ориентированном языке;

выполнение введенных заданий;

анализ результатов выполнения задания и удовлетворения их предъявляемым требованиям.

Программные средства подсистемы управления процессом математической обработки обеспечивают управление перестройкой технологических цепей обработки в соответствии с потребностями задач, решаемых системой обработки. Кроме того, на них возлагаются функции защиты вычислительных процессов от сбоев и отказов аппаратуры, документирования процесса управления математической обработкой, изменения режимов функционирования программных средств системы обработки.

При создании многомашинных вычислительных комплексов на системы управления возлагаются задачи организации совместного функ-

ционирования средств вычислительного комплекса, формирования единой очереди заданий и планирования их выполнения на различных ЭВМ.

Подсистема формирования выходных результатов предназначена для расчета и выдачи траекторных параметров и их точностных характеристик в виде, удобном для потребителя. В составе подсистемы целесообразно иметь специализированный пакет программ, позволяющих автоматизировать процесс составления отчетов по каждому эксперименту, что позволит существенно сократить количество ручных операций и общее время обработки результатов эксперимента.

Для обеспечения решения задач первичной и вторичной обработки траекторных измерений, а также планирования работы средств и систем ИВК, развития самого СМО в состав программ СМО автоматизированной системы обработки траекторных измерений должны быть включены программы, позволяющие решать «обеспечивающие» задачи.

Основными из таких задач являются следующие:

- вычисление рефракционных поправок к измерениям оптических и радиотехнических измерительных средств;

- оценка параметров атмосферы (для вычисления рефракционных поправок и учета влияния атмосферы на движение наблюдаемых ЛА);

- моделирование рельефа местности;

- статистическая оценка точностных характеристик измерительных средств ИВК;

- периодическая паспортизация измерительных средств ИВК, проводимая для уточнения их точностных характеристик;

- расчет целеуказаний для планирования работы средств и систем ИВК, решение задач устранения неоднозначности многошкальных фазометрических средств, корректировки результатов первичной обработки и т. д.;

- задачи, связанные с моделированием процессов измерений: моделирование работы измерительных средств, траекторий движения ЛА, погрешностей измерений, процессов обработки и т. д.

С целью сокращения сроков разработки и реализации системы СМО дополнительно с поставляемым в комплекте с ЭВМ общим математическим обеспечением в рамках подсистемы создания и испытаний СМО необходимо иметь:

- пакеты прикладных программ широкого назначения;

- технологический комплекс создания программ с реализацией диалоговых режимов отладки;

- автоматизированные средства документирования программ;

- комплекс тестов и математических моделей различных исходных данных;

- средства ведения библиотек программных модулей и программ СМО в виде исходных, объектных, загрузочных модулей и каталогизированных процедур.

Таким образом, система специального математического обеспечения является важным компонентом ИВК и представляет собой сложный комплекс алгоритмических и программных средств. Создание эффективной системы СМО требует решения широкого круга сложных

научно-технических задач. Научные результаты, заложенные в основу создания системы СМО, во многом определяют уровень решения задач, стоящих перед ИВК в целом.

§ 2. Адаптивный метод определения параметров движения ЛА по результатам траекторных измерений

При использовании методов фильтрации для обработки траекторных измерений (ТИ) возникают две основные практические трудности — выбор математической модели оцениваемого динамического процесса и выбор вероятностных характеристик ошибок измерений. Полное статистическое описание реальных процессов как правило невозможно. Кроме того, математическая модель оцениваемого процесса хотя и должна быть достаточно полной, тем не менее с вычислительной точки зрения желательна ее относительная простота.

Процедура выбора параметров, характеризующих состояние динамической системы, с целью включения их в математическую модель не может быть полностью формализована. При этом необходимо стремиться к максимально возможному уменьшению размерности вектора состояния, поскольку известно, что объем вычислительных операций, например, при использовании стандартного фильтра Калмана пропорционален третьей степени размерности системы [74].

Построение математической модели движения неразрывно связано с конкретным содержанием решаемой задачи. Исходя из этого, рассмотрим особенности процесса траекторных измерений.

Вектор измеряемых параметров $Z(t)$ связан с вектором состояния $X(t)$ уравнением измерений

$$Z(t) = f\{X(t)\} + V(t),$$

в котором $f\{X(t)\}$ — детерминированная функция; $V(t)$ — погрешность измерений.

На практике функция $f\{X(t)\}$ нелинейна. Поэтому линеаризуя ее, разлагая ее в окрестности точки X_0 в ряд Тейлора и учитывая относительно малость отклонения опорной точки от реальной траектории, ограничиваемся двумя первыми членами разложения и получаем

$$\begin{aligned} Z(t) &= f(X_0) + \frac{df(X_0)}{d(X)} [X(t) - X_0] + V(t) = f(X_0) + \\ &+ \frac{df(X_0)}{dt} X_0 + \frac{df(X_0)}{dt} X(t) + V(t). \end{aligned}$$

Преобразуем полученное выражение к виду

$$Z(t) = H_0 + HX(t) + V(t), \quad (2.1)$$

где H — матрица измерений.

Наличие в измерениях погрешностей $V(t)$ приводит к ошибкам оценивания. По своим свойствам и влиянию на последующую обработку ошибки измерений можно разделить на грубые, систематические и случайные.

Грубые ошибки возникают в результате внезапного изменения условий проведения конкретного эксперимента, например сбоя аппаратуры, ошибки операторов и т. п. Измерения, содержащие такие ошибки, обычно сильно отличаются от других результатов, поэтому они легко выявляются в процессе экспресс-анализа и исключаются или заменяются на достоверные на этапе первичной обработки.

Систематические ошибки характеризуются тем, что сохраняют свои значения от измерения к измерению или изменяются по определенному закону. В зависимости от причин возникновения такие ошибки можно разделить на следующие группы:

- инструментальные, обусловленные конструктивными особенностями ИС;

- ошибки установки, зависящие от особенностей размещения ИС вдоль трасс полета и их геодезической привязки;

- методические ошибки, являющиеся следствием особенностей методов измерений.

Систематические ошибки имеют достаточно сложную структуру и представляют собой медленно меняющиеся функции времени. Частично эти ошибки выявляются при калибровке и паспортизации ИС и учитываются в процессе статистической обработки. Нескомпенсированная часть рассматривается как случайная медленно меняющаяся ошибка, значения которой сильно коррелированы.

Случайные ошибки обусловлены действием случайных факторов в процессе измерений. В реальных ИС приходится сталкиваться со случайными возмущениями типа волновой структуры и шума. Для ошибок, имеющих волновую структуру, характерно волнообразное поведение (по крайней мере на коротких интервалах времени). Ошибки типа шума обладают хаотическим характером, имеют резкие изломы и скачки. Их поведение характеризуется отсутствием какой-либо сглаженности и регулярности. Такие ошибки невозможно предсказать, но их можно описать статистическими характеристиками.

В зависимости от характера возмущений, приводящих к появлению случайных ошибок, по длительности интервала корреляции они делятся на две составляющие:

- быстро меняющаяся, представляющая собой эргодическую стационарную слабо коррелированную функцию, причинами возникновения которой являются шумы следящих систем и приемных устройств ИС, флуктуации атмосферы и сигнала и т. п.;

- медленно меняющаяся составляющая, представляющая собой сильнокоррелированную функцию, определяемую в основном начальными возмущениями.

Используя результаты экспериментальных исследований влияния на точностные характеристики измерений различных факторов, таких, как дальность до ЛА, динамика сопровождения, угол места, характеристики бортовой аппаратуры, время суток, климатические условия и другие, можно сделать следующие выводы о характере погрешностей ТИ:

погрешность измерений в пределах требований технических условий испытаний слабо изменяется во времени и в зависимости от измеряемых параметров;

систематическая составляющая может существенно (в несколько раз) превышать случайную составляющую;

закон распределения ошибок близок к нормальному;

возможны грубые отклонения отдельных измерений и целых реализаций.

При условии отбраковки недостоверных (с грубыми ошибками) измерений модель измерений можно представить в следующем виде:

$$Z(t) = Z_{\text{н}}(t) + u + v_1(t) + v_2(t),$$

где $Z(t)$ — измеренное значение параметра; $Z_{\text{н}}(t)$ — истинное значение параметра; u — систематическая, постоянная для данной реализации составляющая; $v_1(t)$ — медленно меняющаяся составляющая погрешности; $v_2(t)$ — случайная составляющая.

Такая модель позволяет оценить характеристики каждой из составляющих погрешности. Дисперсия случайной составляющей $\sigma_{v_2}^2$ может определяться сравнением измерений со «сглаженными» значениями или по разностям измерений в соседние моменты времени, так как частота съема информации в существующих ИС выше частоты, необходимой для описания маневров ЛА. Оценка σ_u^2 и $\sigma_{v_1}^2$ может осуществляться сравнением измерений с эталонными, которые имеют обычную частоту, недостаточную для оценки $\sigma_{v_2}^2$, но достаточную для оценки $\sigma_{v_1}^2$.

Погрешности u и $v_1(t)$ могут быть представлены в виде конечного числа неслучайных линейно независимых функций

$$u + v_1(t) = \sum_m Q_m g_m(t),$$

где Q_m — неизвестные числовые параметры; g_m — детерминированные линейно независимые функции.

Величины Q_m постоянны в одной реализации, но случайны на множестве реализаций. Причинами, вызывающими их случайные изменения от работы к работе, могут быть погрешности: калибровки; определения времени ретрансляции приемоответчиков, частоты сигналов; определения поправок на распространение радиоволн и др.

Случайная составляющая, которую можно представить в виде медленно меняющейся и быстро меняющейся компонент

$$v_2(t) = v_{2\text{м}}(t) + v_{2\text{б}}(t),$$

подавляется при статистической обработке в три-четыре раза в высокочастотной части, а для неманеврирующих ЛА — практически полностью. Однако низкочастотную часть весьма трудно отделить от систематических погрешностей.

Таким образом, можно записать, что вектор $V(t)$ в выражении (2.1) представляет собой комбинацию медленно меняющихся компонент $v_1(t)$ и независимой быстро меняющейся составляющей $v_2(t)$:

$$V(t) = \sum_i v_{1i}(t) \delta_i + v_2(t), \quad i = 1, 2, \dots, m, \quad (2.2)$$

где m — число составляющих;

$$\delta_i = \begin{cases} 1, & \text{если } v_{li} \text{ присутствует в измерениях;} \\ 0, & \text{если } v_{li} \text{ отсутствует в измерениях.} \end{cases}$$

Исходя из предположения о нормальном законе распределения погрешности измерений, в качестве основного критерия оценки точности измерений выбирается среднеквадратическое отклонение погрешности, а составляющие ошибки можно представить в виде нормальных случайных процессов с корреляционными функциями вида

$$K(\tau) = \sigma^2 \exp \{-\alpha |\tau|\}.$$

При дискретных измерениях период корреляции для выбора шага дискретности определяется выражением

$$T = \int_0^\infty |R(\tau)| d\tau,$$

где $R(\tau)$ — ковариационная функция.

В этом случае при $R(\tau) = \exp \{-\alpha |\tau|\}$ период корреляции записывается как $T = 1/\alpha$.

Тогда можно выразить корреляционные функции составляющих ошибки измерений через их статистические характеристики — дисперсию σ^2 и период корреляции T :

$$K_i(\tau) = \sigma_i^2 \exp \left\{ -\frac{|\tau|}{T_i} \right\}, \quad i = 1, 2, \dots, l, \quad (2.3)$$

где l — число составляющих.

В работе [32] показано, что случайный процесс с корреляционной функцией (2.3) может быть описан в виде линейного стохастического дифференциального уравнения (СДУ)

$$v_i(t) = v_i(t_0) + \int_{t_0}^t -\alpha_i v_i(\tau) d\tau + \int_{t_0}^t \beta_i d\omega_\tau, \quad (2.4)$$

в котором $\alpha_i = \frac{1}{T_i}$; $\beta_i = \sqrt{\sigma_i^2 2\alpha_i}$.

Таким образом, каждую составляющую погрешности измерений (2.2) можно описать с помощью выражения (2.4), при этом они будут отличаться значениями α_i и β_i , т. е. статистическими характеристиками этих погрешностей T_i и σ_i^2 .

При моделировании движение ЛА считается известным, если для заданного интервала времени t удастся найти зависимость параметров движения X от времени t . Искомые параметры движения могут быть получены интегрированием соответствующей системы дифференциальных уравнений, описывающих движение ЛА,

$$\frac{dX}{dt} = f\{X, Q, t\},$$

где $Q = \{q_1, q_2, \dots, q_n\}$ — вектор параметров модели.

В реальной ситуации параметры модели известны не точно, а часть из них имеет случайный характер. Кроме того, функция $f\{X, Q, t\}$ нелинейна. В этом случае задача фильтрации оказывается нелинейной и ее решение основывается на минимизации дисперсии ошибки фильтрации (оптимальность в среднеквадратическом смысле) и сводится к нахождению функции условной плотности вероятности состояния при заданных измерениях и определению эволюции во времени этой условной плотности. Однако, хотя и получены соответствующие СДУ [32], трудности сугубо вычислительного характера при оценке условной плотности вероятности состояния оказываются непреодолимыми в большинстве практических задач. В связи с этим будем рассматривать процесс движения в момент времени t как сумму компоненты, являющейся функцией наблюдений $Z(t)$ до момента времени t , и случайной составляющей. Тогда уравнение для $X(t)$ имеет следующий вид:

$$X(t) = f\{Z(t), t\} + V(t).$$

Здесь $V(t)$ — случайный процесс, обусловленный погрешностями измерений, которые, как показано выше, может быть описан линейными СДУ вида (2.4).

Случайный процесс, движение ЛА по недетерминированным траекториям также может быть описан линейными неоднородными СДУ вида

$$X(t) = X(t_0) + \int_{t_0}^t \mathfrak{F}_0 d\tau + \int_{t_0}^t \mathfrak{F} X(\tau) d\tau + \int_{t_0}^t \mathcal{L} d\omega_\tau,$$

где $\mathfrak{F}_0, \mathfrak{F}, \mathcal{L}$ — матрицы перехода.

Для учета влияния медленно меняющихся составляющих ошибок измерений на выходе оценки расширим вектор состояния, введя в него эти компоненты, в результате чего получим

$$X(t) = \begin{bmatrix} X(t) \\ v^2(t) \end{bmatrix} = X(t_0) + \int_{t_0}^t A_0 d\tau + \int_{t_0}^t A X(\tau) d\tau + \int_{t_0}^t B d\omega_\tau, \quad (2.5)$$

где A_0 — детерминированная функция, определяемая полученными измерениями;

$$A = \begin{bmatrix} \mathfrak{F} & 0 \\ 0 & \alpha_{v_i} \end{bmatrix}; \quad B = \begin{bmatrix} \mathcal{L} & 0 \\ 0 & \beta_{v_i} \end{bmatrix}.$$

Тогда расширенный вектор состояния связан с вектором наблюдений посредством уравнения измерений, имеющего следующий вид:

$$Z(t) = H(t_0) + H(t) X(t) + v_2(t). \quad (2.6)$$

Выражение (2.5) полностью определяет состояние динамической системы по наблюдениям (2.6), в которых имеет место ошибка, описываемая системой уравнений (2.4).

Таким образом, в общем виде выражение (2.5) определяет стохастическую модель движения ЛА в виде линейных неоднородных стохастических дифференциальных уравнений. В ее основе лежит представление детерминированной составляющей траектории в виде функции

измерений, а в качестве стохастической составляющей используется быстро меняющаяся составляющая погрешности измерений. Сильно коррелированная медленно меняющаяся компонента погрешности измерений учитывается за счет введения в вектор оцениваемых параметров, тем самым предотвращая смещение получаемых на выходе алгоритма обработки оценок. Такая модель проста в реализации и наглядна с физической точки зрения. Наличие нулевых элементов в переходных матрицах позволяет значительно сократить вычислительные затраты.

Для учета дискретного характера измерений (2.6), описывающихся выражением

$$Z_K = H_0 + HX_K + v_{2K}, \quad (2.7)$$

необходимо иметь дискретные аналоги непрерывных описаний случайных процессов (2.5) и (2.4), которые, как показано в работе [40], в форме рекуррентных стохастических уравнений (PCY) имеют следующий вид:

$$X_K = F_0 + FX_{K-1} + GW_K, \quad (2.8)$$

$$v_{iK} = a_i v_{iK-1} + b_i e_K. \quad (2.9)$$

Здесь e_K — последовательность гауссовских случайных величин с нулевым математическим ожиданием и единичной дисперсией;

$$a_i = \exp \left\{ -\frac{\Delta t}{T_i} \right\}; \quad b_i = \sigma_i \sqrt{1 - a_i^2}, \quad i = 1, 2.$$

Связь между непрерывным и дискретным описаниями, а значит, и элементами матриц в уравнениях (2.5) и (2.8) определяется следующими соотношениями:

$$F(\Delta t) = \Phi(\Delta t), \quad (2.10)$$

$$G(\Delta t) G^T(\Delta t) = \Phi(\Delta t) \left[\int_0^{\Delta t} \Phi^{-1}(\tau) \mathcal{L} \mathcal{L}^T (\Phi^{-1}(\tau))^T d\tau \right] \Phi^T(\Delta t), \quad (2.11)$$

где Δt — дискретность измерений; $\Phi(\Delta t)$ — решение однородного матричного дифференциального уравнения с постоянными коэффициентами: $\frac{d\Phi(\Delta t)}{dt} = A\Phi(\Delta t)$; $\Phi(0) = I$; I — единичная матрица.

Подставив выражение (2.8) в формулу (2.7), получим уравнение наблюдаемого процесса

$$Z_K = H_0 + H(FX_{K-1} + GW_K) + v_{2K}.$$

Выразив шумовую компоненту через измерения из выражения (2.7) в $(K-1)$ -й момент, получим

$$v_{2K-1} = Z_{K-1} - (H_0 + HX_{K-1}).$$

Подставим это выражение в уравнение наблюдаемого процесса, используя выражение (2.9), и проведем ряд аналитических преобразований, в результате чего получим рекуррентное уравнение

$$Z_K = a_2 Z_{K-1} + H_0 - a_2 H_0 + (HF - a_2 H) X_{K-1} + HGW_K + b_2 e_K.$$

Введем обозначения

$$H_{0K}^* = H_0 - a_2 H_0 + a_2 Z_{K-1}, \quad (2.12)$$

$$H^* = HF - a_2 H, \quad (2.13)$$

$$V_{2K}^* = HGW_K + b_2 e_K \quad (2.14)$$

и запишем РСУ частично наблюдаемого процесса (X_t, Z_t) , $t = 0, 1, \dots, N$,

$$\begin{cases} X_K = F_0 + FX_{K-1} + GW_K, \\ Z_K = H_{0K}^* + H^*X_{K-1} + V_{2K}^*. \end{cases} \quad (2.15)$$

Построение оптимальных оценок ненаблюдаемых значений X по наблюдениям Z заключается в определении вероятностных характеристик распределения последовательности (X, Z) . Для этого используем теорему Липцера — Ширяева [40], в которой доказывается, что если задана частично наблюдаемая случайная последовательность

$$(\Theta, \xi) = (\Theta_t, \xi_t), \quad t = 0, 1, \dots, N,$$

$$\Theta_t = \{\Theta_1(t), \dots, \Theta_K(t)\},$$

$$\xi_t = \{\xi_1(t), \dots, \xi_l(t)\},$$

определяемая рекуррентными уравнениями

$$\Theta_{t+1} = a_0(t, \xi) + a_1(t, \xi) \Theta_t + b_1(t, \xi) e_1(t+1) + b_2(t, \xi) e_2(t+1),$$

$$\xi_{t+1} = A_0(t, \xi) + A_1(t, \xi) \Theta_t + B_1(t, \xi) e_1(t+1) + B_2(t, \xi) e_2(t+1),$$

где $e_1(t) = \{e_{11}(t), \dots, e_{1K}(t)\}$, $e_2(t) = \{e_{21}(t), \dots, e_{2l}(t)\}$ — независимые гауссовские векторы с независимыми компонентами, каждая из которых нормально распределена; $a_0(t, \xi)$, $A_0(t, \xi)$ — вектор-функции; $a_1, A_1, b_1, B_1, b_2, B_2$ — матричные функции, то условное математическое ожидание m_t и дисперсия γ_t находятся из рекуррентных уравнений (для наглядности опущены аргументы t, ξ)

$$m_{t+1} = (a_0 + a_1 m_t) + (b \circ B + a_1 \gamma_t A_1^T) (B \circ B + A_1 \gamma_t A_1^T)^{-1} (\xi_{t+1} - A_0 - A_1 m_t),$$

$$\gamma_{t+1} = (a_1 \gamma_t a_1^T + b \circ b) - (b \circ B) + a_1 \gamma_t A_1^T B \circ B + A_1 \gamma_t A_1^T)^{-1} (b \circ B + a_1 \gamma_t A_1^T)^T,$$

где $b \circ b = b_1 b_1^T + b_2 b_2^T$; $b \circ B = b_1 B_1^T + b_2 B_2^T$; $B \circ B = B_1 B_1^T + B_2 B_2^T$.

Согласно этой теореме вероятностные характеристики частично наблюдаемого процесса (2.15) определяются следующими выражениями:

$$M_K = (F_0 + FM_{K-1}) + [G(HG)^T + FD_{K-1}H^{*T}][HG(HG)^T + b_2 b_2^T + H^*D_{K-1}H^{*T}]^{-1}(Z_K - H_{0K}^* - H^*M_{K-1}),$$

$$D_K = (FD_{K-1}F^T + GG^T) - [G(HG)^T + FD_{K-1}H^{*T}][HG(HG)^T + b_2 b_2^T + H^*D_{K-1}H^{*T}]^{-1}[G(HG)^T + FD_{K-1}H^{*T}]^T.$$

Анализ полученных соотношений показывает, что выражение для условного математического ожидания M_K частично наблюдаемого процесса (2.15) можно записать в форме алгоритма калмановской фильтрации. Тогда оценка фильтрации вычисляется в виде

$$M_K = F_0 + FM_{K-1} + K_K P_K^{-1} \tilde{Z}_K, \quad (2.16)$$

где выражение

$$K_K = GG^T H^T + FD_{K-1} H^{*T} \quad (2.17)$$

определяет коэффициент усиления фильтра, формула

$$\tilde{Z}_K = Z_K - H_{0K}^* - H^* M_{K-1} \quad (2.18)$$

есть не что иное, как невязка измерений, а

$$P_K = HGG^T H^T + b_2 b_2^T + H^* D_{K-1} H^{*T} \quad (2.19)$$

ковариационная матрица ошибок прогноза.

Нахождение аналитического представления решения уравнения для матрицы дисперсий D_K требует громоздких выкладок. Преобразуем его, как показано в работе [33], к виду

$$D_K = C_1 D_K C_1^T + C_2 - C_1 D_{K-1} H^{*T} [(HG(HG)^T + b_2 b_2^T)^T + \\ + H^* D_{K-1} H^{*T}]^{-1} H^* D_{K-1} C_1^T,$$

где

$$C_1 = F - G(HG)^T C_3^{-1} H^{*T}; \quad C_2 = GG^T - C_3^{-1} [G(HG)^T]^T; \quad C_3 = \\ = HG(HG)^T + b_2 b_2^T.$$

Приняв допущение о невырожденности обращаемых матриц, применим к полученному уравнению лемму об обращении матриц [64] и будем иметь следующую форму записи дискретного уравнения Риккати:

$$D_K = C_1 (D_{K-1}^{-1} + C_4)^{-1} C_1^T + C_2,$$

в котором $C_4 = H^{*T} C_3^{-1} H^*$.

Преобразуем это уравнение. Имеем

$$D_K = C_1 (I + C_4 D_{K-1} D_{K-1}^{-1}) C_1^T + C_2 = C_1 D_{K-1} (I + C_4 D_{K-1}) C_1 + C_2 = \\ = C_1 D_{K-1} (C_1^{-T} + C_1^{-T} C_4 D_{K-1})^{-1} + C_2,$$

где индекс «—т» обозначает транспонирование обратной матрицы.

Умножим полученное уравнение на $(C_1^{-T} + C_1^{-T} C_4 D_{K-1})$ справа и приведем его к виду

$$D_K q_1 D_{K-1} + D_K q_2 - q_3 D_{K-1} - q_4 = 0,$$

где $q_1 = C_1^{-T} C_4$; $q_2 = C_1^{-T}$; $q_3 = C_1 + C_2 C_1^{-T} C_4$; $q_4 = C_2 C_1^{-T}$.

В результате получим рекуррентное уравнение Риккати, решением которого является значение ковариационной матрицы ошибок фильтрации. На практике начальные значения этой матрицы априорно неизвестны. Поэтому сделаем допущение о стационарности системы, т. е.

решение рекуррентного уравнения стремится к стационарному значению D . В этом случае D выбираем в качестве начального приближения $D_0 = D$ и рекуррентное уравнение вырождается в алгебраическое уравнение Риккати

$$Dq_1D + Dq_2 - q_3D - q_4 = 0. \quad (2.20)$$

Таким образом, выражения (2.16), (2.20) определяют вероятностные характеристики распределения последовательности ненаблюдаемых значений фильтруемого процесса $X(t)$ по известным наблюдениям $Z(t)$.

Перейдем к синтезу адаптивного дискретно-непрерывного фильтра. Из теории случайных процессов известно, что при оптимальном выборе параметров модели состояния невязка \tilde{Z}_K является последовательностью независимых гауссовских векторов с ковариационной матрицей ошибок P_K .

Это позволяет построить функцию правдоподобия $f(Z, \Theta)$, определяемую как вероятность поступления именно последовательности измерений Z_K при выбранном составе параметров модели состояния системы

$$f(Z, \Theta) = \prod_{K=1}^N [(2\pi)^m \det P]^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \tilde{Z}_K^T P^{-1} \tilde{Z}_K \right\}. \quad (2.21)$$

Следовательно, необходимо получить такие значения параметров Θ , при которых функция правдоподобия (2.21) будет максимальной. В качестве параметров Θ используются периоды корреляции и дисперсии случайных процессов движения и погрешностей измерений модели.

Для реализации адаптации параметров модели (2.5) методом максимального правдоподобия при расчетах удобнее пользоваться логарифмом функции правдоподобия

$$-L(Z, \Theta) = \frac{Nm}{2} \ln 2\pi + \frac{N}{2} \ln \det P + \frac{1}{2} \sum_{K=1}^N \tilde{Z}_K^T P^{-1} \tilde{Z}_K. \quad (2.22)$$

Из анализа этого выражения видно, что для минимизации функционала $L(Z, \Theta)$ необходимо оптимизировать последние слагаемые

$$L^* = \sum_{K=1}^N \tilde{Z}_K^T P^{-1} \tilde{Z}_K + N \ln \det P$$

Исходя из критерия оптимальности по минимуму суммы квадратов невязок, для минимизации L^* будем использовать метод Ньютона, который обеспечивает более высокую скорость сходимости последовательных приближений к решению по сравнению с другими методами оптимизации. Количество вычислений на одной итерации методом Ньютона больше, чем в градиентных методах из-за вычисления и обращения матриц вторых частных производных. Но для получения решения методом Ньютона требуется в десятки и сотни раз меньше итераций. Поэтому решение минимизационной задачи методом Ньютона более эффективно, так как квадратичная функция локально точнее

аппроксимирует минимизируемую функцию, чем линейная, лежащая в основе градиентных методов.

Итерационная процедура метода Ньютона имеет вид

$$\Theta_{m+1} = \Theta_m - I_2^{-1} I_1, \quad (2.23)$$

где Θ — вектор определяемых параметров; m — номер итерации; I_2 — матрица вторых частных производных; I_1 — вектор первых частных производных.

Частные производные вычисляются по формулам

$$\begin{aligned} \frac{\partial L^*}{\partial \Theta} &= \sum_{K=1}^N \tilde{Z}_K^T P^{-1} \frac{\partial \tilde{Z}_K}{\partial \Theta_i} + N \frac{\partial}{\partial \Theta_i} (\ln \det P), \\ \frac{\partial^2 L^*}{\partial \Theta_i \partial \Theta_j} &= \sum_{K=1}^N \frac{\partial \tilde{Z}_K}{\partial \Theta_i} P^{-1} \frac{\partial \tilde{Z}_K}{\partial \Theta_j} + \sum_{K=1}^N \tilde{Z}_K^T P^{-1} \frac{\partial^2 \tilde{Z}_K}{\partial \Theta_i \partial \Theta_j} + \\ &+ N \frac{\partial^2}{\partial \Theta_i \partial \Theta_j} (\ln \det P). \end{aligned} \quad (2.24)$$

Однако необходимость вычисления матрицы вторых частных производных приводит к сложностям вычислительного характера при реализации метода в практических алгоритмах.

Во-первых, вычисление элементов матрицы $\frac{\partial^2 \tilde{Z}_K}{\partial \Theta_i \partial \Theta_j}$ занимает значительную часть вычислений.

Во-вторых, в процессе вычислений матрица $\frac{\partial^2 L^*}{\partial \Theta_i \partial \Theta_j}$ в ряде случаев теряет положительную определенность и тогда в процедуре оптимизации необходимо применять специальные способы защиты от этого, что в итоге приводит к усложнению алгоритмов и, как следствие, к увеличению объема занимаемой памяти и времени вычислений.

Кроме того, для оптимизируемых параметров, когда их значения близки к оптимальным, элементы матрицы $\frac{\partial^2 \tilde{Z}_K}{\partial \Theta_i \partial \Theta_j}$ мало отличны от нуля. Учитывая это и используя результаты, приведенные в работе [67], рассчитываем матрицу вторых частных производных по сокращенной формуле

$$\frac{\partial^2 L^*}{\partial \Theta_i \partial \Theta_j} = \sum_{K=1}^N \frac{\partial \tilde{Z}_K^T}{\partial \Theta_i} P^{-1} \frac{\partial \tilde{Z}_K}{\partial \Theta_j} + N \frac{\partial^2}{\partial \Theta_i \partial \Theta_j} (\ln \det P). \quad (2.25)$$

Другим недостатком метода Ньютона является то, что если целевая функция не является сильно выпуклой и начальное приближение находится достаточно далеко от минимума, может наблюдаться явление расходимости. Градиентные методы по скорости решения задачи уступают методу Ньютона, но повышают надежность сходимости. Поэтому при решении оптимизационной задачи будем использовать комбинированный подход, в котором методу Ньютона, как более быстрому действующему, отдано предпочтение. Градиентный метод (в частности,

метод скорейшего спуска) используется лишь в случае нарушения нормальной сходимости процесса приближений. Критерием перехода к нему является увеличение суммы квадратов невязок измерений m -й итерации в сравнении с $(m - 1)$ -й. Это означает, что полученные на m -й итерации оценки искоемых параметров уводят от требуемого минимума квадратичной формы и в этом смысле являются неоптимальными. Используя оценки, полученные в $(m - 1)$ -й итерации, используем метод скорейшего спуска для определения направления движения к минимуму и вновь переходим к методу Ньютона.

В результате дифференцирования выражения (2.18) с учетом формулы (2.12) для матрицы H_0 получим

$$\frac{\partial \tilde{Z}_K}{\partial \Theta} = \frac{\partial a_2}{\partial \Theta} (Z_{K-1} - H_0) - \frac{\partial H^*}{\partial \Theta} M_{K-1} - H^* \frac{\partial M_{K-1}}{\partial \Theta}. \quad (2.26)$$

Отсюда видно, что необходимо рассчитать производную условного математического ожидания, для чего продифференцируем выражение (2.6). В результате получим

$$\begin{aligned} \frac{\partial M_K}{\partial \Theta} = \frac{\partial F}{\partial \Theta} M_{K-1} + F \frac{\partial M_{K-1}}{\partial \Theta} + \frac{\partial K_K}{\partial \Theta} P_K^{-1} \tilde{Z}_K + K_K P_K^{-1} \frac{\partial \tilde{Z}_K}{\partial \Theta} + \\ + K_K \frac{\partial P_K^{-1}}{\partial \Theta} \tilde{Z}_K. \end{aligned} \quad (2.27)$$

Здесь производная коэффициента усиления определяется дифференцированием уравнения (2.17):

$$\frac{\partial K_K}{\partial \Theta} = \frac{\partial GG^T}{\partial \Theta} H^T + \frac{\partial F}{\partial \Theta} D_{K-1} H^{*T} + F \frac{\partial D_{K-1}}{\partial \Theta} H^{*T} + F D_{K-1} \frac{\partial H^{*T}}{\partial \Theta}, \quad (2.28)$$

где производная матрицы измерений определяется дифференцированием выражения (2.13) и имеет вид

$$\frac{\partial H^*}{\partial \Theta} = H \frac{\partial F}{\partial \Theta} + \frac{\partial a_2}{\partial \Theta} H. \quad (2.29)$$

В выражении (2.25) осталось найти частные производные логарифма определителя ковариационной матрицы ошибок прогноза.

Находим первую производную $\frac{\partial}{\partial \Theta} (N \ln \det P)$. Определитель матрицы P представляет собой произведение ее собственных чисел:

$$\det P = \prod_{i=1}^K \lambda_i.$$

Тогда логарифм этого выражения равен сумме логарифмов собственных чисел

$$\ln \det P = \sum_{i=1}^N \ln \lambda_i,$$

что есть не что иное, как след матрицы $\ln P$, т. е.

$$\ln \det P = \text{sp}(\ln P).$$

Производная этого выражения имеет вид

$$\frac{\partial}{\partial \Theta} (\ln \det P) = \frac{\partial}{\partial \Theta} (\text{sp} \ln P).$$

Учитывая линейность оператора sp , получаем окончательное выражение для производной логарифма определителя ковариационной матрицы ошибок прогноза

$$\frac{\partial}{\partial \Theta} (\text{sp} \ln P) = \text{sp} \frac{\partial}{\partial \Theta} \ln P = \text{sp} \left(P^{-1} \frac{\partial P}{\partial \Theta} \right).$$

Для нахождения матрицы вторых частных производных продифференцируем первую производную и получим

$$\frac{\partial^2}{\partial \Theta_i \partial \Theta_j} (\ln \det P) = \frac{\partial}{\partial \Theta_i} \left[\text{sp} \left(P^{-1} \frac{\partial P}{\partial \Theta_j} \right) \right].$$

Из правила дифференцирования произведения видно, что в этом выражении необходимо вычислить производную обратной матрицы. Для этого воспользуемся тем свойством обратной матрицы, что ее умножение на прямую матрицу дает в результате единичную, т. е.

$$P^{-1} P = I.$$

Тогда

$$\frac{\partial}{\partial \Theta} (P^{-1} P) = 0,$$

или

$$\frac{\partial P^{-1}}{\partial \Theta} P + P^{-1} \frac{\partial P}{\partial \Theta} = 0.$$

Перенесем второе слагаемое в правую часть и умножим обе части на обратную матрицу справа

$$\frac{\partial P^{-1}}{\partial \Theta} P P^{-1} = - P^{-1} \frac{\partial P}{\partial \Theta} P^{-1}.$$

В результате получим выражение для производной обратной матрицы

$$\frac{\partial P^{-1}}{\partial \Theta} = - P^{-1} \frac{\partial P}{\partial \Theta} P^{-1}.$$

С учетом этого выражение для второй частной производной логарифма определителя ковариационной матрицы ошибок принимает следующий вид:

$$\frac{\partial^2}{\partial \Theta_i \partial \Theta_j} (\ln \det P) = \text{sp} \left\{ \left(- P^{-1} \frac{\partial P}{\partial \Theta_i} P^{-1} \right) \frac{\partial P}{\partial \Theta_j} + P^{-1} \frac{\partial^2 P}{\partial \Theta_i \partial \Theta_j} \right\}.$$

В свою очередь выражение для производной $\partial P / \partial \Theta$ определяется дифференцированием выражения (2.19):

$$\begin{aligned} \frac{\partial P_K}{\partial \Theta} &= \frac{\partial}{\partial \Theta} G G^T H^T + \frac{\partial}{\partial \Theta} b_2 b_2^T + \frac{\partial H^*}{\partial \Theta} D_{K-1} H^{*T} + H^* \frac{\partial D_{K-1}}{\partial \Theta} H^{*T} + \\ &+ H^* D_{K-1} \frac{\partial H^*}{\partial \Theta}. \end{aligned}$$

Производная $\partial^2 P / \partial \Theta_i \partial \Theta_j$ определяется численным методом по приращениям.

Полученные выражения для частных производных позволяют рассчитать элементы вектора I_1 матрицы I_2 и решить оптимизационную задачу (2.23) и тем самым найти решение уравнения максимального правдоподобия (2.22).

Однако разработанный таким образом адаптивный фильтр предполагает строго определенную модель состояния. При ее изменении возникает необходимость пересчета фильтра, в результате которого появляются громоздкие аналитические выражения для составляющих фильтра. Кроме того, такой алгоритм не позволяет использовать в процессе обработки дополнительную априорную информацию, получаемую по результатам обработки предыдущих экспериментов.

Проанализируем формулы расчета производных (2.26) — (2.29). В них входят производные матричных функций F , GG^T и D по адаптивным параметрам Θ . Вычислим эти производные.

Фундаментальная матрица $\Phi(t)$ представляет собой матричную экспоненту вида

$$\Phi(t) = e^{tA} = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!}$$

и является решением дифференциального уравнения

$$\frac{d\Phi}{dt} = A\Phi, \quad \Phi(0) = I. \quad (2.30)$$

Продифференцируем это уравнение по параметрам

$$\frac{\partial}{\partial \Theta} \frac{d\Phi}{dt} = \frac{\partial A}{\partial \Theta} \Phi + A \frac{\partial \Phi}{\partial \Theta}.$$

На основании теоремы Шварца поменяем порядок дифференцирования [23]

$$\frac{d}{dt} \left(\frac{\partial \Phi}{\partial \Theta} \right) = A \frac{\partial \Phi}{\partial \Theta} + \frac{\partial A}{\partial \Theta} \Phi \quad (2.31)$$

и, объединив уравнения (2.30) и (2.31), получим систему матричных дифференциальных уравнений

$$\begin{cases} \frac{d}{dt} \left(\frac{\partial \Phi}{\partial \Theta} \right) = A \frac{\partial \Phi}{\partial \Theta} + \frac{\partial A}{\partial \Theta} \Phi, \\ \frac{d}{dt} \Phi = A\Phi, \end{cases} \quad (2.32)$$

решением которой является матрица искомым производных.

Аналогично находим производную $\frac{\partial}{\partial \Theta} GG^T$. Матричная функция GG^T определяется выражением (2.11). В соответствии с теоремой 15.1, приведенной в работе [33], производная такой функции по времени вычисляется в виде

$$\frac{d}{dt} GG^T = AGG^T + GG^T A^T + BB^T. \quad (2.33)$$

Продифференцируем это уравнение по параметрам адаптации и поменяем порядок дифференцирования

$$\frac{d}{dt} \left(\frac{\partial}{\partial \Theta} GG^T \right) = \frac{\partial A}{\partial \Theta} GG^T + A \frac{\partial}{\partial \Theta} GG^T + \frac{\partial}{\partial \Theta} GG^T A^T + GG^T \frac{\partial A}{\partial \Theta} + \frac{\partial}{\partial \Theta} BB^T.$$

Перегруппировав это уравнение и объединив его с уравнением (2.33), получим систему матричных дифференциальных уравнений

$$\left\{ \begin{array}{l} \frac{d}{dt} \left(\frac{\partial}{\partial \Theta} GG^T \right) = \frac{\partial A}{\partial \Theta} GG^T + A \frac{\partial}{\partial \Theta} GG^T + \frac{\partial}{\partial \Theta} GG^T A^T + \\ + GG^T \frac{\partial A}{\partial \Theta} + \frac{\partial}{\partial \Theta} BB^T, \\ \frac{d}{dt} GG^T = AGG^T + GG^T A^T + BB^T, \end{array} \right. \quad (2.34)$$

решением которой является искомая производная.

Для расчета производной ковариационной матрицы ошибок фильтрации продифференцируем уравнение (2.20) и получим следующее уравнение:

$$\begin{aligned} \frac{\partial D}{\partial \Theta} q_1 D + D \frac{\partial q_1}{\partial \Theta} D + D q_1 \frac{\partial D}{\partial \Theta} + \frac{\partial D}{\partial \Theta} q_2 + D \frac{\partial q_2}{\partial \Theta} - \\ - \frac{\partial q_3}{\partial \Theta} D - q_3 \frac{\partial D}{\partial \Theta} - \frac{\partial q_4}{\partial \Theta} = 0, \end{aligned} \quad (2.35)$$

где

$$\begin{aligned} \frac{\partial q_1}{\partial \Theta} &= \frac{\partial C_1}{\partial \Theta} C_4 + C_1 \frac{\partial C_4}{\partial \Theta}, \\ \frac{\partial q_2}{\partial \Theta} &= \frac{\partial C_1^{-T}}{\partial \Theta}, \\ \frac{\partial q_3}{\partial \Theta} &= \frac{\partial C_1}{\partial \Theta} + \frac{\partial C_2}{\partial \Theta} C_1^{-T} C_4 + C_2 \frac{\partial C_1^{-T}}{\partial \Theta} C_4 + C_2 C_1^{-T} \frac{\partial C_4}{\partial \Theta}, \\ \frac{\partial q_4}{\partial \Theta} &= \frac{\partial C_2}{\partial \Theta} C_1^{-T} + C_2 \frac{\partial C_1^{-T}}{\partial \Theta}, \\ \frac{\partial C_1}{\partial \Theta} &= \frac{\partial F}{\partial \Theta} GG^T H^T C_3^{-1} H^{*T} + F \frac{\partial}{\partial \Theta} GG^T H^T C_3^{-1} H^{*T} + \\ &+ FGG^T H^T \frac{\partial C_3}{\partial \Theta} H^{*T} + FGG^T H^T C_3^{-1} \frac{\partial H^{*T}}{\partial \Theta}, \\ \frac{\partial C_2}{\partial \Theta} &= \frac{\partial}{\partial \Theta} GG^T - \frac{\partial C_3^{-1}}{\partial \Theta} HGG^T - C_3^{-1} H \frac{\partial}{\partial \Theta} GG^T, \\ \frac{\partial C_4}{\partial \Theta} &= \frac{\partial H^{*T}}{\partial \Theta} C_3^{-1} H^* + H^{*T} \frac{\partial C_3^{-1}}{\partial \Theta} H^* + H^{*T} C_3^{-1} \frac{\partial H^*}{\partial \Theta}. \end{aligned}$$

Таким образом, решение уравнений (2.32), (2.34), (2.35) позволяет полностью рассчитать фильтр с адаптацией по параметрам стохастической модели в непрерывном виде. Изменение функциональной формы модели не требует пересчета фильтра, так как в этом случае изменяются только элементы переходных матриц и весь расчет проводится по тем же алгоритмам синтеза фильтра. Такой подход к расчету адаптивных фильтров расширяет область их практического использования для фильтрации случайных процессов широкого класса.

В разработанном дискретно-непрерывном фильтре b_2 является параметром регуляризации, который адаптируется к реализующимся измерениям. Однако получаемые оценки являются субоптимальными, так как их вычисление зависит непосредственно от значения параметра регуляризации, что эквивалентно внесению некоторой дополнительной погрешности на вход фильтра, приводящей к увеличению результирующей ошибки фильтрации.

В результате обработки траекторной информации, получаемой по результатам летного эксперимента (ЛЭ), накапливаются различные статистические данные, в том числе и о значениях регуляризирующего параметра b_2 , который для простоты изложения и соответствия общепринятой терминологии обозначим через α . Это позволяет задать априорную плотность $\rho_0(\alpha)$ распределения параметра регуляризации на положительной полуоси $(0, \infty)$. Экспериментальные исследования показали, что наилучшая точность оценок при определении движения ЛА, привлекаемых к ЛЭ, достигается при $\alpha \in (10^{-8}, 10^{-2})$. Поэтому целесообразно ограничиться конечным интервалом (α_0, α_m) значений α и тогда, когда для решаемого класса задач не проводились численные эксперименты и расчеты. В этом случае плотность $\rho_0(\alpha)$ будем считать соответствующей равномерному распределению [4]. При поступлении реализации наблюдаемого процесса $Z(t)$, используя свойство математического ожидания, получаем

$$\begin{aligned}\hat{X}_0(t) &= M[X(t) | Z(t)] = M\{M[X(t) | Z(t), \alpha] | Z(t)\} = \\ &= \int_a^b \hat{X}_\alpha(t) \rho_t(\alpha) d\alpha,\end{aligned}$$

где $\hat{X}_\alpha(t)$ — регуляризованная оценка; $\rho_t(\alpha)$ — апостериорная плотность распределения.

Известно [33], что если некоторая случайная величина связана каким-либо образом с измерениями этой величины, то для нахождения оптимальной (в определенном смысле) оценки со статистической точки зрения необходимо восстановить апостериорную плотность распределения вероятности этой величины, зная которую, легко найти оценку методом максимального правдоподобия.

Выразим $\rho_t(\alpha)$ через $\rho_0(\alpha)$, $Z(t)$. По формуле Байеса получим

$$\rho_t(\alpha) = \rho[\alpha | Z(t)] = \frac{\rho[X(t), \alpha | Z(t)]}{\rho[X(t) | Z(t), \alpha]}. \quad (2.36)$$

По теореме Бьюси [3] находим

$$\rho[X(t), \alpha | Z(t)] = \frac{M_Z[\Psi_t(\alpha) | X(t), \alpha] \rho_0(\alpha) \rho[X(t)]}{M_Z[\Psi_t(\alpha)]}, \quad (2.37)$$

$$\begin{aligned} \Psi_t(\alpha) = & [(2\pi)^n \det S_\alpha(t)]^{-1} \exp \left\{ \int_0^t X^\tau(\tau) H^\tau(\tau) S_\alpha^{-1}(\tau) Z(\tau) d\tau - \right. \\ & \left. - \frac{1}{2} \int_0^t X^\tau(\tau) H^\tau(\tau) S_\alpha^{-1}(\tau) X(\tau) d\tau \right\}, \end{aligned}$$

где M_Z — математическое ожидание относительно $Z(t)$; $\rho[X(t)]$ — плотность распределения вектора.

Применим теорему Бьюси к знаменателю и получим

$$\rho[X(t) | Z(t), \alpha] = \frac{M_Z[\Psi_t(\alpha) | X(t), \alpha] \rho[X(t)]}{M_Z[\Psi_t(\alpha) | \alpha]}. \quad (2.38)$$

Подставляя выражения (2.37) и (2.38) в формулу (2.36), получаем

$$\rho_t(\alpha) = \frac{M_Z[\Psi_t(\alpha) | \alpha] \rho_0(\alpha)}{M[\Psi_t(\alpha)]}. \quad (2.39)$$

Используя свойства математических ожиданий, записываем (2.39) в следующем виде:

$$\rho_t(\alpha) = \frac{M_Z[\Psi_t(\alpha) | \alpha] \rho_0(\alpha)}{\int_0^b M_Z[\Psi_t(\alpha) | \alpha] \rho_0(\alpha) d\alpha},$$

где функция

$$\begin{aligned} M_Z[\Psi_t(\alpha) | \alpha] = & [(2\pi)^n \det S_\alpha(t)]^{-1} \exp \left\{ -\frac{1}{2} \int_0^t \{ Z(\tau) - H(\tau) \times \right. \\ & \left. \times \hat{X}_\alpha(\tau) \}^\tau S_\alpha^{-1}(\tau) [Z(\tau) - H(\tau) \hat{X}_\alpha(\tau)] d\tau \right\} \end{aligned} \quad (2.40)$$

есть не что иное, как функция правдоподобия $\Phi_t(\alpha)$.

Тогда окончательно выражение для апостериорной плотности распределения параметра регуляризации принимает вид

$$\rho_t(\alpha) = \frac{\Phi_t(\alpha) \rho_0(\alpha)}{\int_{\alpha_0}^{\alpha_m} \Phi_t(\alpha) \rho_0(\alpha) d\alpha}. \quad (2.41)$$

Таким образом, оптимальная в среднеквадратическом смысле линейная оценка $\hat{X}_0(t)$ состояния системы $X(t)$ по наблюдениям $Z(t)$ вычисляется по формуле

$$\hat{X}_0(t) = \int_{\alpha}^{\alpha_m} \hat{X}_\alpha(t) \rho_t(\alpha) d\alpha. \quad (2.42)$$

Корреляционная матрица ошибок имеет вид

$$P(t) = \int_0^b \{P_{\alpha}(t) + [\hat{X}_0(t) - \hat{X}_{\alpha}(t)][\hat{X}_0(t) - \hat{X}_{\alpha}(t)]\} \rho_t(\alpha) d\alpha, \quad (2.43)$$

где $\rho_t(\alpha)$ определяется выражением (2.41).

Вычисляемые регуляризованные оценки $\hat{X}_{\alpha}(\alpha)$ не зависят друг от друга при различных значениях параметра регуляризации из интервала $\alpha \in [\alpha_0, \alpha_m]$. Это позволяет распараллелить вычислительный процесс по α , задавшись сеткой опорных значений параметра регуляризации $\alpha_0, \alpha_1 = \alpha_0 + \Delta\alpha, \alpha_2 = \alpha_1 + \Delta\alpha$ и т. д. до $\alpha_j = \alpha_{j-1} + \Delta\alpha = \alpha_m$, и получить оптимальную в среднеквадратическом смысле и соответствующую с условным математическим ожиданием оценку в темпе поступления измерений.

На рис. 1 представлена структурная схема такого фильтра с распараллеливанием по параметру регуляризации, положенного в основу изобретения [88]. Он содержит N регуляризованных фильтров Калмана 1 [89], первый сумматор 2, N основных перемножителей 3, N делителей 4, N канальных перемножителей 5, N интеграторов 6, N делителей на минус два 7, N вычислителей экспоненты 8, первый умножитель на постоянную величину 9, второй сумматор 10, N умножителей 11, второй умножитель на постоянную величину 12, N блоков деления 13 и генератор тактовых импульсов 14.

Каждый регуляризованный фильтр Калмана (РФК), структурная схема которого приведена на рис. 2, содержит вычитатель 1, первый сумматор 2, умножитель 3, второй сумматор 4, интегратор 5, матричные умножители 6, 7 и генератор стационарного случайного сигнала 8.

На вход адаптивного фильтра поступает сигнал $Z(t)$ и попадает на все регуляризованные фильтры Калмана 1. Каждый РФК 1 стабилизируется случайным процессом с различной интенсивностью. Первый РФК имеет параметр регуляризации α_0 , второй — $\alpha_1 = \alpha_0 + \Delta\alpha$, третий — $\alpha_2 = \alpha_1 + \Delta\alpha$ и т. д., или в общем виде $\alpha_N = \alpha_{N-1} + \Delta\alpha$.

Значения α_0 и $\Delta\alpha$ задаются заранее с таким расчетом, чтобы оптимальное $\alpha_{\text{опт}}$ для решаемой задачи попадало в интервал $[\alpha_0, \alpha_N]$. На выходе вычитателя 1 (см. рис. 2), который является разностным выходом РФК, образуется невязка

$$\tilde{Z}_i = Z(t) - H\hat{X}_i(t),$$

где $\hat{X}_i(t)$ — оценка, получаемая на выходе i -го РФК; H — коэффициент перехода от оценки к измерению; i — номер канала.

Значения невязки $\tilde{Z}_i(t)$ поступают на соответствующие входы делителей 4, вычисляющие отношение невязки и регуляризованной дисперсии $\tilde{Z}_i(t)/S_{\alpha_i}(t)$, и на канальный перемножитель 5, на выходе которых получаются отношения $\tilde{Z}_i^2(t)/S_{\alpha_i}$. Эти значения поступают последовательно на интеграторы 6, делители 7 на минус два и вычислители 8 экспоненты, на выходе которых образуются значения отноше-

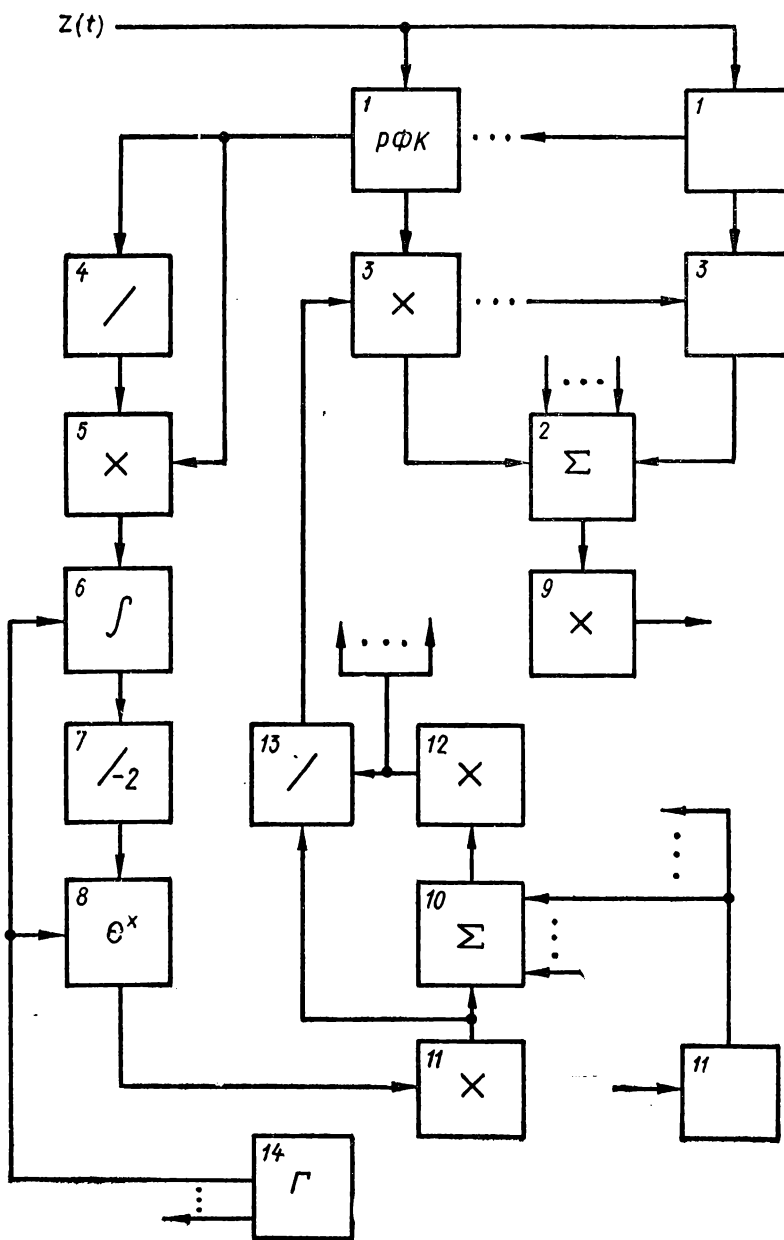


Рис. 1

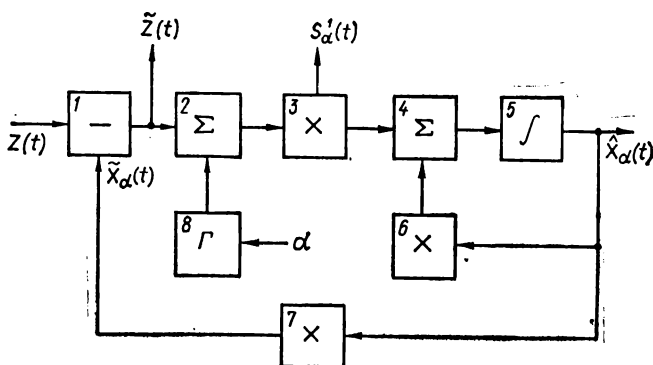


Рис. 2

ний правдоподобия

$$\Phi(\alpha_i) = \exp \left\{ -\frac{1}{2} \int_0^t \frac{\tilde{z}_i^2(\tau)}{S_{\alpha_i}(\tau)} d\tau \right\},$$

которые поступают на умножители 11 и на их выходе получают произведения отношений правдоподобия и априорной плотности распределения параметра регуляризации $\Pi_i = P_0(\alpha_i) \Phi(\alpha_i)$. Произведения Π_i с выходов умножителей 11 подаются на входы «делимое» блоков 13 деления и на второй сумматор 10, образуя сумму

$$S_m = \sum_{i=1}^N \Pi_i = \sum_{i=1}^N P_0(\alpha_i) \Phi(\alpha_i),$$

которая умножается на постоянную величину $\Delta\alpha$ в умножителе 12 и подается на входы «делитель» блока 13 деления. На выходах блоков 13 образуются значения апостериорной плотности распределения параметра регуляризации

$$\rho_i(\alpha_i) = \frac{\Pi_i}{S_m} = \frac{\rho_0(\alpha_i) \Phi(\alpha_i)}{\Delta\alpha \sum_{i=1}^N \rho_0(\alpha_i) \Phi(\alpha_i)},$$

которые являются весовыми коэффициентами для соответствующих оценок $\hat{X}_i(t)$ и перемножаются с ними в основных перемножителях 3. Значения получаемых на выходах основных перемножителей 3 произведений $\rho_i(\alpha_i) X_i(t)$ суммируются в первом сумматоре 2, и сумма

$$S_m = \sum_{i=1}^N \rho_i(\alpha_i) \hat{X}_i(t)$$

перемножается с величиной $\Delta\alpha$ в умножителе 9. При поступлении таковых импульсов от генератора 14 на выходе фильтра вычисляется оценка

$$\hat{X}_{\text{опт}}(t) = \Delta\alpha \sum_{i=1}^N \rho_i(\alpha_i) X_i(t).$$

В результате на основе использования апостериорной информации о параметре регуляризации получена не зависящая от этого параметра оптимальная в среднеквадратическом смысле и совпадающая с условным математическим ожиданием оценка исследуемого процесса.

При необходимости может быть использована субоптимальная оценка, получаемая на информационном выходе каждого из РФК.

Полученный результат позволяет эффективно использовать в практике обработки ТИ перспективные многопроцессорные вычислительные системы.

§ 3. Алгоритм адаптивной дискретно-непрерывной фильтрации траекторных измерений

Для непосредственного синтеза вычислительного алгоритма адаптивной дискретно-непрерывной фильтрации результатов ТИ необходимо разработать следующие процедуры:

1) алгоритмы преобразования непрерывных СДУ вида (2.4), (2.5) в рекуррентные уравнения (2.8), (2.9); в основе этих алгоритмов лежит расчет элементов переходных матриц F и G ;

2) алгоритмы расчета частных производных элементов этих матриц по параметрам адаптации для решения оптимизационной задачи (2.23);

3) алгоритм расчета ковариационной матрицы ошибок фильтрации решением матричного уравнения Риккати (2.20);

4) алгоритм расчета производной ковариационной матрицы ошибок фильтрации по параметрам адаптации.

В основу этих алгоритмов положены современные численные методы линейной алгебры, использующие свойства матричной экспоненциальной функции и матрицы проектора [5].

Матрицы перехода играют центральную роль в динамике систем, описываемых линейными дифференциальными уравнениями. В настоящее время математические свойства этих матриц достаточно хорошо исследованы для теоретических целей. Однако использование чисто теоретических результатов в физических задачах требует конкретного численного расчета матрицы перехода системы. Как отмечалось в § 2, в нашей задаче переходная матрица Φ , получившая название фундаментальной, представляет собой решение однородного матричного дифференциального уравнения с постоянными коэффициентами

$$\frac{d}{dt} \Phi(t) = A\Phi(t), \quad \Phi(0) = I. \quad (3.1)$$

Нетрудно показать, что решением уравнения (3.1) является матричная экспонента.

В общем виде матричная экспонента представляет собой бесконечный матричный ряд

$$e^{At} = \sum_{k=0}^{\infty} \frac{(At)^k}{k!}, \quad (3.2)$$

где A — квадратная матрица.

Продифференцируем экспоненту (3.2)

$$\begin{aligned}\frac{d}{dt}(e^{At}) &= \sum_{k=0}^{\infty} \frac{d}{dt} \left[\frac{(At)^k}{k!} \right] = A + A^2t + \dots + \frac{A^{k+1}t^k}{k!} = \\ &= A \sum_{k=0}^{\infty} \frac{(At)^k}{k!} = Ae^{At},\end{aligned}\quad (3.3)$$

т. е. $\frac{d}{dt}(e^{At}) = Ae^{At}$.

Отсюда следует, что матричная экспонента есть не что иное, как решение уравнения (3.1), т. е. матричная функция (3.2) является искомой фундаментальной матрицей.

Однако непосредственное использование матричной экспоненты для расчета фундаментальной матрицы нецелесообразно. Это обусловлено тем, что при больших численных значениях элементов фундаментальной матрицы для сходимости ряда (3.2) к этим значениям необходимо большое число слагаемых, что может приводить к переполнению разрядной сетки ЭВМ.

Кроме того, при больших значениях отдельных слагаемых могут получаться малые разности, так как ряд знакопеременный. Это может вызвать искажение результатов вычислений за счет округления этих разностей до нуля.

В то же время при малых значениях t ряд быстро сходится.

Учитывая все сказанное, для расчета фундаментальной матрицы воспользуемся свойством матричной экспоненты

$$e^{A(t+s)} = e^{At}e^{As},$$

что аналогично

$$e^{At}e^{At} = e^{2At} = e^{A2t}.$$

Таким образом, рассчитав матричную экспоненту для малого интервала t , когда ряд (3.2) быстро сходится последовательным удвоением, можно быстро рассчитать матричную экспоненту для любого интервала времени. Для выбора начального интервала t_n используем свойства нормы матрицы

$$\|A\|^2 = \sum_{ij} a_{ij}^2,$$

состоящие в том, что

$$\|AB\| \leq \|A\| \|B\|, \quad \|A+B\| \leq \|A\| + \|B\|.$$

Это позволяет оценить остаточный член при конечном суммировании по формуле (3.2)

$$\left\| \sum_{k=m}^{\infty} \frac{(At_n)^k}{k!} \right\| \leq \frac{\|At_n\|^m}{m!},$$

где m — число просуммированных членов.

В работе [8] показано, что, выбрав норму $\|At_n\|$ меньше $1/2$, уже при числе слагаемых $m = 16$ остаточный член имеет порядок 10^{-40} и этой остаточной суммой можно пренебречь.

Таким образом, для численного расчета фундаментальной матрицы с помощью матричной экспоненты (3.2) расчет начального интервала t_n будем проводить по формуле

$$t_n = \frac{\Delta t}{2^n}, \quad (3.4)$$

где n определяется следующим образом:

$$n = \operatorname{argmin}_m \left(\left\| A \frac{\Delta t}{2^m} \right\| \leq \frac{1}{2} \right). \quad (3.5)$$

Ограничившись 16 членами в матричном ряде (3.2), вычисляем матричную экспоненту на начальном интервале

$$e_0^{At} = e^{At_n} \sum_{k=0}^{15} \frac{(At_n)^k}{k!} + \varepsilon_{16}, \quad \varepsilon_{16} \approx 10^{-40}. \quad (3.6)$$

Теперь последовательным удвоением n раз рассчитываем матричную экспоненту на полном интервале

$$e_{i+1}^{At} = e_i^{At} e_i^{At}, \quad i = 0, 1, \dots, n. \quad (3.7)$$

На рис. 3 показана блок-схема алгоритма расчета фундаментальной матрицы.

В блоке 1 осуществляется ввод исходных данных в виде элементов матрицы A , значение интервала Δt , на котором необходимо рассчитать фундаментальную матрицу, и устанавливаются начальные значения переменных. В блоках 2—4 рассчитываются число разбиений n исходного интервала Δt по формуле (3.5) и начальный интервал суммирования t_n по формуле (3.4).

В блоках 5—7 в цикле по k от 0 до 15 вычисляется матричная экспонента на начальном интервале суммированием по формуле (3.6). В блоках 8—10 последовательным умножением n раз рассчитывается искомая матрица по формуле (3.7), а в блоке 11 реализуется вывод результатов расчета.

Расчет производной фундаментальной матрицы можно свести к аналогичному алгоритму. Для этого в качестве показателя степени матричной экспоненты используем матрицу Ψ системы уравнений (2.32):

$$\Psi = \begin{vmatrix} A & \partial A / \partial Q \\ 0 & A \end{vmatrix}. \quad (3.8)$$

Эта матрица блочно-верхнетреугольная. Исходя из этого любая матрица Ψ^k сохраняет свойство верхнего треугольника. Поэтому матричная экспонента $e^{\Psi t}$ также будет блочно-верхнетреугольной и ее можно записать в следующем виде:

$$e^{\Psi t} = \begin{vmatrix} e_{11} & e_{12} \\ 0 & e_{22} \end{vmatrix}.$$

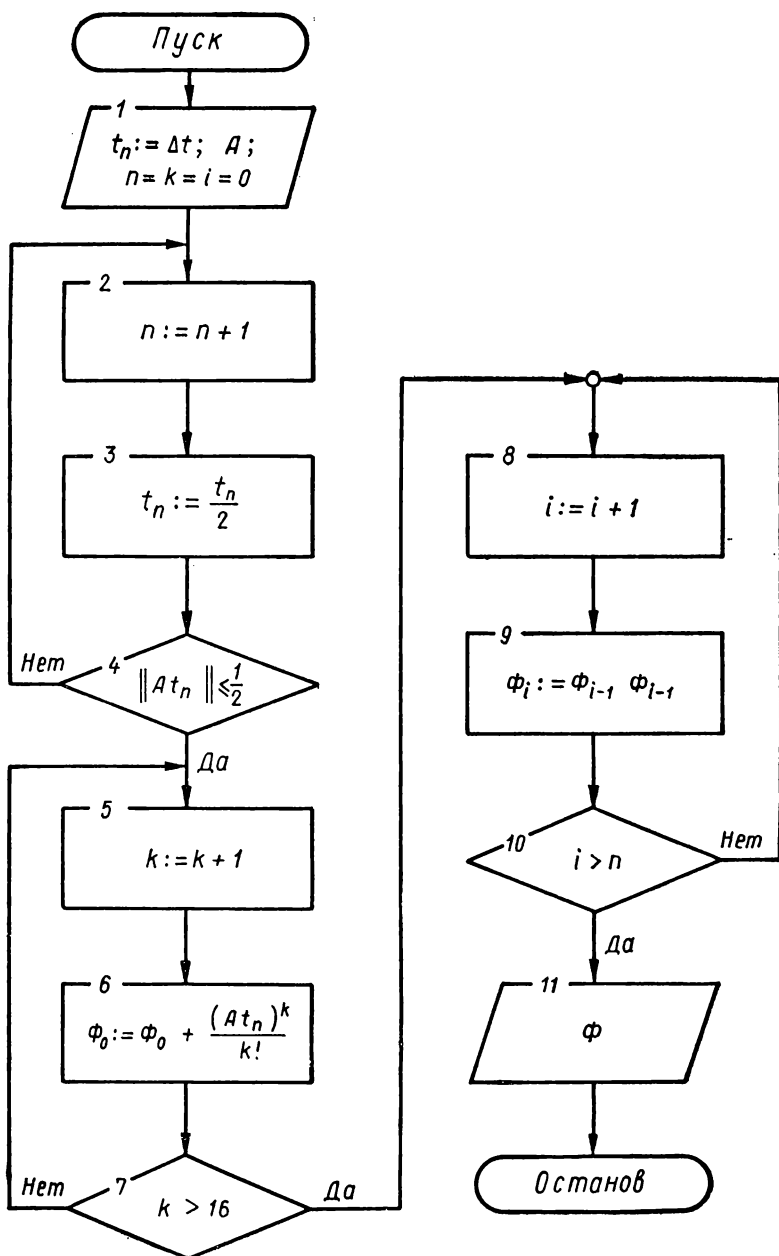


Рис. 3

Используя выражение (3.3), получаем

$$\begin{aligned} \frac{d}{dt} (e^{\Psi t}) &= \Psi e^{\Psi t} = \begin{vmatrix} A & \partial A / \partial Q \\ 0 & A \end{vmatrix} \begin{vmatrix} e_{11} & e_{12} \\ 0 & e_{22} \end{vmatrix} = \\ &= \begin{vmatrix} Ae_{11} & Ae_{12} + \partial A / \partial Q e_{22} \\ 0 & Ae_{22} \end{vmatrix}, \end{aligned}$$

откуда следует система дифференциальных уравнений

$$\begin{aligned} \frac{d}{dt} e_{11} &= Ae_{11}, \\ \frac{d}{dt} e_{12} &= Ae_{12} + \partial A / \partial Q e_{22}, \\ \frac{d}{dt} e_{22} &= Ae_{22}, \\ e_{11}(0) &= e_{22}(0) = I, \\ e_{12}(0) &= 0. \end{aligned}$$

Сравним эти уравнения с (2.29) и из теоремы о единственности решения систем обыкновенных уравнений можно сделать вывод, что элемент e_{12} является искомой производной фундаментальной матрицы. Ее расчет производится по алгоритму, описанному на рис. 3, проведя в нем следующие изменения: вычисление степеней матрицы Ψ осуществляется для отдельных компонентов по рекуррентным соотношениям

$$\begin{aligned} (\Psi t)^k &= \begin{bmatrix} \Psi_{11}(k) & \Psi_{12}(k) \\ 0 & \Psi_{22}(k) \end{bmatrix}, \quad \Psi_{11}(k) = \Psi_{22}(k), \\ (\Psi t)^k &= \Psi t (\Psi t)^{k-1} = \begin{bmatrix} At & \partial A / \partial Q^t \\ 0 & At \end{bmatrix} \begin{bmatrix} \Psi_{11}(k-1) & \Psi_{12}(k-1) \\ 0 & \Psi_{22}(k-1) \end{bmatrix} = \\ &= \begin{bmatrix} At\Psi_{11}(k-1) & At\Psi_{12}(k-1) + \frac{\partial A}{\partial Q} t\Psi_{22}(k-1) \\ 0 & At\Psi_{22}(k-1) \end{bmatrix}. \end{aligned}$$

Отсюда получаем систему рекуррентных уравнений

$$\begin{aligned} \Psi_{11}(k) &= \Psi_{22}(k) = At\Psi_{11}(k-1), \\ \Psi_{12}(k) &= At\Psi_{12}(k-1) + \frac{\partial A}{\partial Q} t\Psi_{22}(k-1). \end{aligned} \quad (3.9)$$

[При удвоении интервала t , после преобразований

$$\begin{aligned} e^{\Psi 2t} &= \begin{bmatrix} e_{11}(2t) & e_{12}(2t) \\ 0 & e_{22}(2t) \end{bmatrix} = e^{\Psi t} e^{\Psi t} = \\ &= \begin{bmatrix} e_{11}(t) e_{11}(t) & e_{11}(t) e_{12}(t) + e_{12}(t) e_{22}(t) \\ 0 & e_{22}(t) e_{22}(t) \end{bmatrix} \end{aligned}$$

используется система уравнений

$$\begin{aligned} e_{11}(2t) &= e_{22}(2t) = e_{11}(t), \\ e_{12}(2t) &= e_{11}(t)e_{12}(t) + e_{12}(t)e_{22}(t). \end{aligned} \quad (3.10)$$

На рис. 4 показана блок-схема алгоритма расчета производной по параметрам адаптации фундаментальной матрицы.

В блоке 1 осуществляется ввод исходных данных и устанавливаются начальные значения переменных; в 2 формируется матрица Ψ в соответствии с выражением (3.8); в 3—5 реализуется расчет начального интервала.

В блоках 6—9 в цикле по k от 0 до 15 рассчитываются степени матрицы $(t_n \Psi)^k$ по формуле (3.9) (блок 7) и значение матрицы Ψ на начальном интервале по формуле матричной экспоненты (3.6), в которой $A = \Psi$ (блок 8); в 10 рассчитывается значение производной фундаментальной матрицы на начальном интервале.

В блоках 11—13 по схеме последовательного удвоения рассчитывается производная фундаментальной матрицы на полном интервале по формуле (3.30); в 14 осуществляется вывод результатов расчета.

Вычисление значений элементов переходной матрицы шума, задаваемой выражением (2.11)

$$G(t)G^T(t) = \Phi(t) \left[\int_0^t \Phi^{-1}(\tau) \mathcal{L}(\tau) \mathcal{L}^T(\tau) \Phi^{-T}(\tau) d\tau \right] \Phi^T(t),$$

непосредственно интегрированием, — процесс довольно сложный и труднореализуемый вычислительными средствами. Для расчета матрицы GG^T воспользуемся свойствами матричной экспоненциальной функции, используемыми ранее.

Возьмем матрицу

$$\gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ 0 & \gamma_{22} \end{bmatrix}$$

с произвольными блоками γ_{ij} , для которой матричная экспонента

$$e^{\gamma t} = \begin{bmatrix} e_{11} & e_{12} \\ 0 & e_{22} \end{bmatrix}$$

удовлетворяет следующему свойству:

$$\frac{d}{dt} e^{\gamma t} = \gamma t e^{\gamma t} = \begin{bmatrix} \gamma_{11} e_{11} & \gamma_{11} e_{12} + \gamma_{12} e_{22} \\ 0 & \gamma_{22} e_{22} \end{bmatrix}.$$

Отсюда получаем систему матричных уравнений

$$\frac{d}{dt} e_{11} = \gamma_{11} e_{11}, \quad e_{11}(0) = I, \quad (3.11)$$

$$\frac{d}{dt} e_{12} = \gamma_{11} e_{12} + \gamma_{12} e_{22}, \quad e_{12}(0) = 0, \quad (3.12)$$

$$\frac{d}{dt} e_{22} = \gamma_{22} e_{22}, \quad e_{22}(0) = I. \quad (3.13)$$

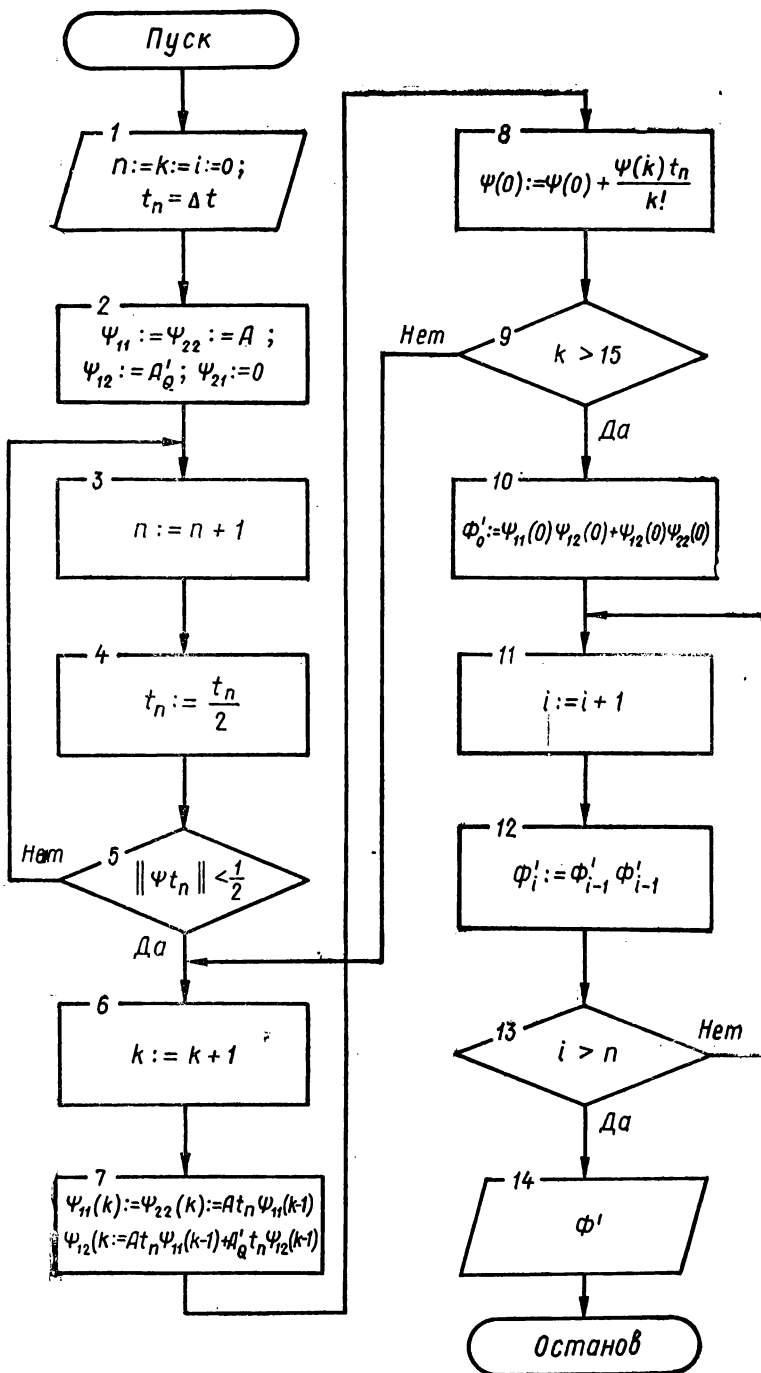


Рис. 4

Решение уравнения (3.11) имеет вид

$$e_{11}(t) = e^{\gamma_{11}t}.$$

Решение уравнения (3.13)

$$e_{22}(t) = e^{\gamma_{22}t}$$

подставим в выражение (3.12) и получим неоднородное матричное дифференциальное уравнение

$$\frac{d}{dt} e_{12} = \gamma_{11}e_{11} + \gamma_{12}e^{\gamma_{22}t}.$$

Решение такого уравнения при наличии неоднородной компоненты $\gamma_{12}e^{\gamma_{22}t}$ определяется выражением

$$e_{12} = e^{\gamma_{11}t} \left[\int_0^t (e^{\gamma_{11}\tau})^{-1} \gamma_{12} e^{\gamma_{22}\tau} d\tau \right]. \quad (3.14)$$

Если сформировать матрицу γ следующим образом:

$$\gamma_{11} = \mathcal{F}, \quad \gamma_{22} = -\mathcal{F}^T, \quad \gamma_{12} = \mathcal{L}\mathcal{L}^T,$$

то

$$e^{\gamma_{11}t} = \Phi(t), \quad e^{\gamma_{22}t} = \Phi^{-1}(t)$$

и выражение (3.14) примет вид

$$e_{12}(t) = \Phi(t) \left[\int_0^t \Phi^{-1}(\tau) \mathcal{L}(\tau) \mathcal{L}^T(\tau) \Phi^{-T}(\tau) d\tau \right]. \quad (3.15)$$

Сравним полученное выражение с (2.11). Умножим правую часть выражения (3.15) на $\Phi^T(t)$, что эквивалентно умножению левой части (3.15) на e_{11}^T , т. е.

$$e_{12}e_{11}^T = \Phi(t) \left[\int_0^t \Phi^{-1}(\tau) \mathcal{L}(\tau) \mathcal{L}^T(\tau) \Phi^{-T}(\tau) d\tau \right] \Phi^T(t). \quad (3.16)$$

Правые части уравнений (2.11) и (3.16) равны, следовательно, равны и левые части

$$GG^T = e_{12}e_{11}^T. \quad (3.17)$$

Таким образом, для вычисления матрицы GG^T достаточно рассчитать элементы e_{11} и e_{12} матричной экспоненты $e^{\gamma t}$, у которой квадратная матрица γ имеет вид

$$\gamma = \begin{bmatrix} \mathcal{F} & \mathcal{L}\mathcal{L}^T \\ 0 & -\mathcal{F}^T \end{bmatrix} \quad (3.18)$$

по алгоритму, аналогичному алгоритму, представленному на рис. 3. Согласно (3.6) элементы e_{11} и e_{12} определяются удвоением интервала t по формулам

$$\begin{aligned} e_{11}(2t) &= e_{11}(t) e_{11}(t), \\ e_{12}(2t) &= e_{11}(t) e_{12}(t) + e_{12}(t) e_{22}(t). \end{aligned}$$

Этот алгоритм можно упростить.

Умножив (3.17) справа на $e_{11}^T(t)$, получим

$$GG^T e^{-T}(t) = e_{12}(t) e_{11}^T(t) e_{11}^{-T}(t) = e_{12}(t).$$

Подставим этот результат в выражение для $e_{12}(2t)$

$$e_{12}(2t) = GG^T e_{12}^{-T}(2t) = e_{11}(t) GG^T e_{11}^{-T}(t) + GG^T e_{11}^{-T}(t) e_{22}(t).$$

Умножая это выражение справа на $e_{11}^T(2t)$

$$GG^T e_{11}^{-T}(2t) e_{11}^T(2t) = e_{11}(t) GG^T e_{11}^{-T}(t) e_{11}^T(2t) + GG^T e_{11}^{-T}(t) e_{22}(t) e_{11}^T(2t),$$

и учитывая, что $e_{11}^T(2t) = e_{11}^T(t) e_{11}^T(t)$, $e_{22}(t) = e_{11}^{-T}(t)$, окончательно получаем формулу удвоения

$$GG^T(2t) = e_{11}(t) GG^T(t) e_{11}^T(t) + GG^T(t). \quad (3.19)$$

На рис. 5 показана блок-схема алгоритма расчета переходной матрицы шума.

В блоке 1 осуществляется ввод исходных данных и устанавливаются начальные значения переменных; в 2 формируется матрица γ по формуле (3.18); в 3—5 реализуется расчет начального интервала в соответствии с выражениями (3.4), (3.5).

В блоках 6—9 в цикле по k от 0 до 15 рассчитываются степени матрицы $(t_n \gamma)^k$ аналогично формуле (3.9) и значение матрицы γ на начальном интервале по формуле (3.2); в 10 вычисляется значение искомой матрицы на начальном интервале по формуле (3.17); в 11—13 рассчитывается переходная матрица шума на полном интервале по схеме удвоения (3.19); в 14 осуществляется вывод результатов расчетов.

В § 2 было показано, что производная переходной матрицы шума $\frac{\partial}{\partial Q} GG^T$ является решением системы уравнений (2.34)

$$\left\{ \begin{aligned} \frac{d}{dt} \left(\frac{\partial}{\partial Q} GG^T \right) &= A \frac{\partial}{\partial Q} GG^T + \frac{\partial}{\partial Q} GG^T A^T + \frac{\partial A}{\partial Q} GG^T + \\ &+ GG^T \frac{\partial A^T}{\partial Q} + \frac{\partial}{\partial Q} BB^T, \\ \frac{d}{dt} GG^T &= AGG^T + GG^T A^T + BB^T. \end{aligned} \right.$$

Запишем эту систему в виде матричного дифференциального уравнения

$$\begin{aligned} \frac{d}{dt} \begin{bmatrix} \frac{\partial}{\partial Q} GG^T & GG^T \\ GG^T & 0 \end{bmatrix} &= \begin{bmatrix} A & \frac{\partial A}{\partial Q} \\ 0 & A \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial Q} GG^T & GG^T \\ GG^T & 0 \end{bmatrix} + \\ &+ \begin{bmatrix} \frac{\partial}{\partial Q} GG^T & GG^T \\ GG^T & 0 \end{bmatrix} \begin{bmatrix} A^T & 0 \\ \frac{\partial A}{\partial Q} & A^T \end{bmatrix} + \begin{bmatrix} \frac{\partial}{\partial Q} BB^T & BB^T \\ BB^T & 0 \end{bmatrix}. \end{aligned}$$

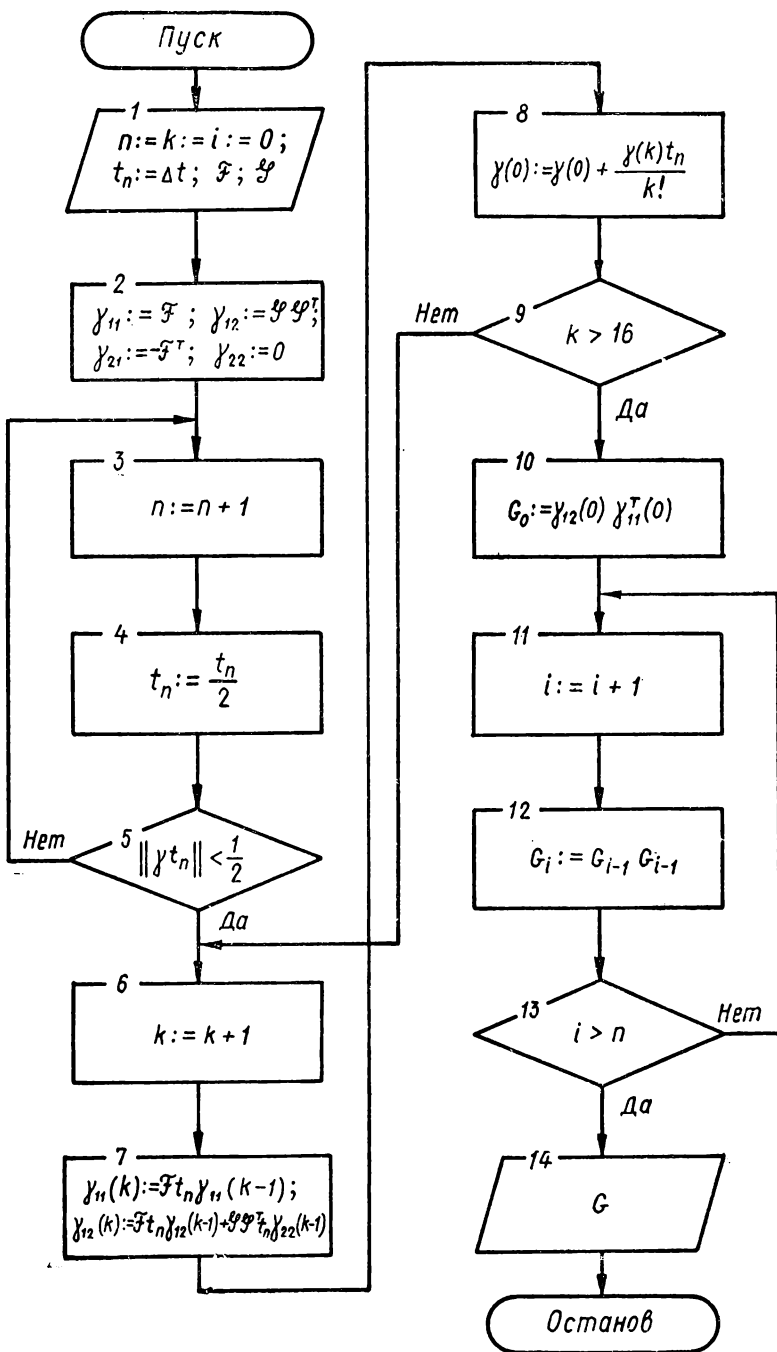


Рис. 5

Тогда искомая производная определяется решением уравнения

$$\frac{d}{dt} \tilde{G}G^T = \tilde{A}\tilde{G}G^T + \tilde{G}G^T\tilde{A}^T + \tilde{B}B^T,$$

в котором

$$\begin{aligned}\tilde{G}G^T &= \begin{bmatrix} \frac{\partial}{\partial Q} GG^T & GG^T \\ GG^T & 0 \end{bmatrix}, \\ \tilde{A} &= \begin{bmatrix} A & \frac{\partial A}{\partial Q} \\ 0 & A \end{bmatrix}, \\ \tilde{B}B^T &= \begin{bmatrix} \frac{\partial}{\partial Q} BB^T & BB^T \\ BB^T & 0 \end{bmatrix}.\end{aligned}$$

Согласно уравнению (3.12) это решение выражается формулой

$$\frac{d}{dt} \tilde{G}G^T = e^{\tilde{A}t} \left[\int_0^t e^{-\tilde{A}\tau} \tilde{B}B^T (e^{-\tilde{A}\tau})^{-T} d\tau \right] (e^{\tilde{A}t})^T,$$

что позволяет рассчитывать $\frac{\partial}{\partial Q} GG^T$ по алгоритму, представленному на рис. 3, используя вместо матрицы γ расширенную блочную матрицу

$$A^* = \begin{bmatrix} \tilde{A} & \tilde{B}B^T \\ 0 & -\tilde{A}^T \end{bmatrix},$$

которая в развернутом виде принимает следующую форму:

$$A^* = \begin{bmatrix} A & \frac{\partial A}{\partial Q} & \frac{\partial}{\partial Q} BB^T & BB^T \\ 0 & A & BB^T & 0 \\ & & -A^T & 0 \\ 0 & & \frac{\partial A^T}{\partial Q} & -A^T \end{bmatrix}. \quad (3.20)$$

Расчет степеней матрицы A^*t для суммирования

$$(tA^*)^k = tA^*(tA^*)^{k-1} = \begin{bmatrix} \Psi_{11} & \Psi_{12} & \Psi_{13} & \Psi_{14} \\ 0 & \Psi_{22} & \Psi_{23} & 0 \\ & 0 & \Psi_{33} & 0 \\ & & \Psi_{43} & \Psi_{44} \end{bmatrix}$$

для каждого элемента осуществляется по следующим рекуррентным соотношениям:

$$\begin{aligned}
 \Psi_{11}(k) &= A\Psi_{11}(k-1), \\
 \Psi_{12}(k) &= A\Psi_{12}(k-1) + \frac{\partial A}{\partial Q} \Psi_{22}(k-1), \\
 \Psi_{13}(k) &= A\Psi_{13}(k-1) + \frac{\partial A}{\partial Q} \Psi_{23}(k-1) + \\
 &\quad + \frac{\partial BB^T}{\partial Q} \Psi_{33}(k-1) + BB^T \Psi_{43}(k-1), \\
 \Psi_{14}(k) &= A\Psi_{14}(k-1) + BB^T \Psi_{44}(k-1), \\
 \Psi_{22}(k) &= A\Psi_{22}(k-1), \\
 \Psi_{23}(k) &= A\Psi_{23}(k-1) + BB^T \Psi_{33}(k-1), \\
 \Psi_{33}(k) &= -A^T \Psi_{33}(k-1), \\
 \Psi_{43}(k) &= -\frac{\partial A^T}{\partial Q} \Psi_{33}(k-1) - A^T \Psi_{43}(k-1), \\
 \Psi_{44}(k) &= -A^T \Psi_{44}(k-1).
 \end{aligned} \tag{3.21}$$

Обозначим рассчитанную матричную экспоненту на начальном интервале

$$\Phi = e^{A^* t_n} = \sum_{k=0}^{16} \frac{(A^* t_n)^k}{k!} = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}. \tag{3.22}$$

Тогда расчет $\frac{\partial}{\partial Q} GG^T$ на начальном интервале проводится по формуле

$$\gamma_0 = \Phi_{12} \Phi_{11}^T = \begin{bmatrix} \Psi_{13} & \Psi_{14} \\ \Psi_{23} & 0 \end{bmatrix} \begin{bmatrix} \Psi_{11} & \Psi_{12} \\ 0 & \Psi_{22} \end{bmatrix}^T = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, \tag{3.23}$$

где $\gamma_{11} = \Psi_{13} \Psi_{11}^T + \Psi_{14} \Psi_{12}^T$; $\gamma_{12} = \Psi_{14} \Psi_{22}^T$; $\gamma_{21} = \Psi_{23} \Psi_{11}^T$.

В итоге рассчитываем производную переходной матрицы шума по формуле

$$\frac{\partial}{\partial Q} GG^T(2t) = \Phi_{11}(t) \gamma(t) \Phi_{11}^T(t) + \gamma(t), \quad t = 0, 1, \dots, n. \tag{3.24}$$

На рис. 6 показана блок-схема алгоритма расчета производной переходной матрицы шума по параметрам адаптации.

В блоке 1 осуществляется ввод исходных данных и устанавливаются начальные значения переменных; в 2 формируется матрица $C = -A^*$ согласно формуле (3.20); в 3—5 вычисляется начальный интервал для суммирования.

В блоках 6—9 в цикле по k от 0 до 15 рассчитываются степени матрицы $(t_n \Psi)^k$ по рекуррентным соотношениям (3.21) (блок 7) и значение

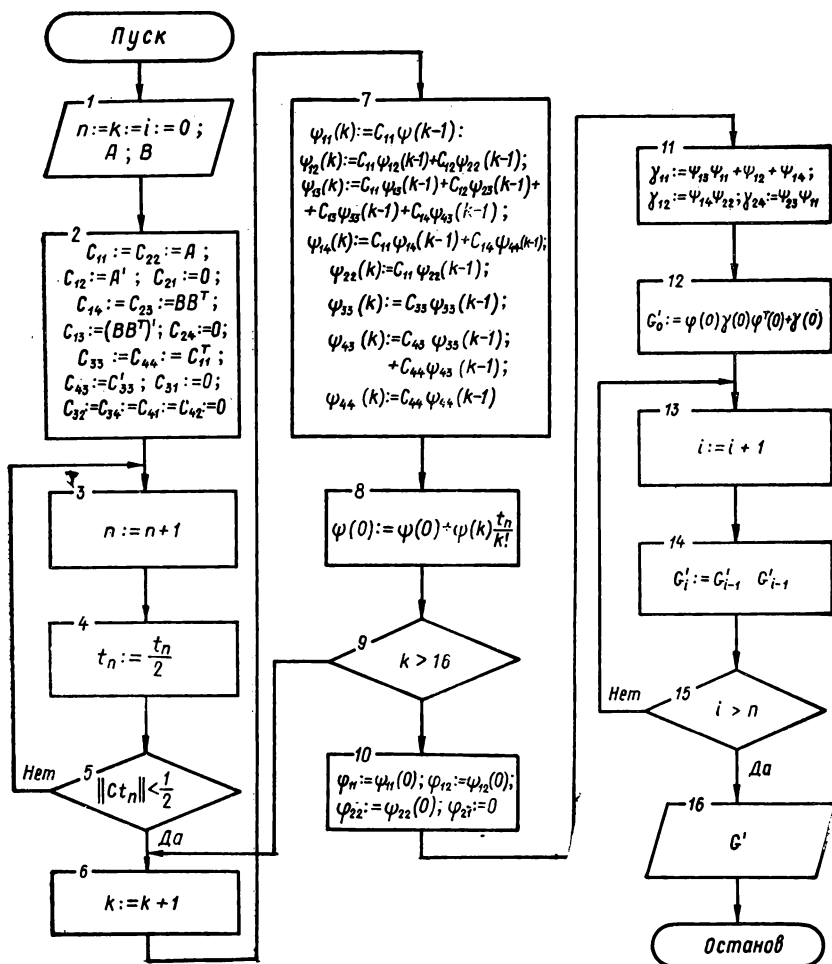


Рис. 6

матрицы Ψ на начальном интервале (блок 8); в 10—11 по формулам (3.22) и (3.23) определяются матрицы ϕ и γ на начальном интервале; в 12 рассчитывается значение производной на начальном интервале по формуле (3.24).

В блоках 13—15 вычисляется значение производной матрицы шума на полном интервале; в 16 осуществляется вывод результатов.

Для расчета ковариационной матрицы ошибок фильтрации необходимо решить матричное уравнение Риккати (2.20):

$$Dq_1D + Dq_2 - q_3D - q_4 = 0. \quad (3.25)$$

Для этого используем свойство компонент, входящих в это уравнение. Одним из таких свойств является то, что для них справедливы следующие тождества, которые легко проверяются непосредственным

перемножением:

$$\begin{aligned} & \begin{bmatrix} I & 0 \\ D & I \end{bmatrix} \begin{bmatrix} q_2 + q_1 D & q_1 \\ 0 & q_3 - Dq_1 \end{bmatrix} \begin{bmatrix} I & 0 \\ D & I \end{bmatrix} = \begin{bmatrix} q_2 & q_1 \\ q_3 & q_4 \end{bmatrix}, \\ & \begin{bmatrix} I & -\hat{q} \\ 0 & I \end{bmatrix} \begin{bmatrix} q_2 + q_1 D & 0 \\ 0 & q_3 - Dq_1 \end{bmatrix} \begin{bmatrix} I & \hat{q} \\ 0 & I \end{bmatrix} = \begin{bmatrix} q_2 + q_1 D & q_1 \\ 0 & q_3 - Dq_1 \end{bmatrix}, \end{aligned} \quad (3.26)$$

где I — единичная матрица; \hat{q} — решение уравнения

$$q_1 = -\hat{q}(q_4 - Dq_1) + (q_2 + q_1 D)\hat{q}.$$

Эти матричные уравнения позволяют получить факторизованное представление гамильтониана системы

$$\mathcal{H} = \begin{bmatrix} q_2 & q_1 \\ q_4 & q_3 \end{bmatrix} = \rho \begin{bmatrix} q_2 + q_1 D & 0 \\ 0 & q_4 - Dq_1 \end{bmatrix} \rho^{-1}, \quad (3.27)$$

где

$$\begin{aligned} \rho &= \begin{bmatrix} I & 0 \\ D & I \end{bmatrix} \begin{bmatrix} I & -\hat{q} \\ 0 & I \end{bmatrix} = \begin{bmatrix} I & -\hat{q} \\ D & -D\hat{q} + I \end{bmatrix}, \\ \rho^{-1} &= \begin{bmatrix} I & \hat{q} \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -D & I \end{bmatrix} = \begin{bmatrix} I - \hat{q}D & \hat{q} \\ -D & I \end{bmatrix}. \end{aligned}$$

Если к матрице \mathcal{H} применить преобразование, результат которого можно представить в каноническом виде

$$S(\mathcal{H}) = \begin{bmatrix} S_{11} & S_{22} \\ S_{21} & S_{22} \end{bmatrix} = \rho \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \rho^{-1}, \quad (3.28)$$

то, проведя перемножения входящих в формулу (3.28) матриц

$$S(\mathcal{H}) = \begin{bmatrix} I & 0 \\ D & 0 \end{bmatrix} \begin{bmatrix} I - \hat{q}D & \hat{q} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I - \hat{q}D & \hat{q} \\ D(I - \hat{q}D) & D\hat{q} \end{bmatrix},$$

в итоге получим

$$D = S_{21}S_{11}^{-1} = S_{12}S_{22}^{-1}. \quad (3.29)$$

Таким образом, решение уравнения Риккати (3.25) заключается в определении элементов матрицы S по гамильтониану \mathcal{H} .

Гамильтониан системы обладает свойством, которое позволяет сконструировать следующее преобразование. Оно заключается в том, что в непрерывном случае собственные числа матрицы $(q_3 - Dq_1)$ положительны, а матрицы $(q_2 + q_1 D)$ отрицательны; в дискретном случае характеристические корни матриц лежат соответственно внутри единичного круга и вне его.

В этом случае матрица $S(\mathcal{H})$ в выражении (3.28) представляет собой матрицу-проектор на пространство собственных векторов и собственных чисел, лежащих вне единичного круга.

Представим матрицу E в виде

$$E\{A\} = CAC^{-1}, \quad (3.30)$$

где E, A, C — матрицы.

Для нахождения этого проектора используем следующие свойства матрицы E .

$$1. E\{A\} - I = E\{A - I\},$$

где I — единичная матрица.

Доказательство. В результате очевидных преобразований получим

$$\begin{aligned} E\{A\} - I &= CAC^{-1} - ICC^{-1} = CAC^{-1} - CIC^{-1} = \\ &= C\{AC^{-1} - IC^{-1}\} = C\{A - I\}C^{-1}, \end{aligned}$$

что в соответствии с выражением (3.30) есть не что иное, как $E\{A - I\}$, что и требовалось доказать.

$$2. [E\{A\} - I][E\{A\} + I]^{-1} = E\{(A - I)(A + I)^{-1}\}.$$

Доказательство. Используя представление (3.30), запишем

$$[E\{A\} - I][E\{A\} + I]^{-1} = [CAC^{-1} - I][CAC^{-1} + I]^{-1}.$$

Исходя из свойства 1, получим

$$\begin{aligned} [CAC^{-1} - I][CAC^{-1} + I]^{-1} &= [C(A - I)C^{-1}][C(A + I)C^{-1}]^{-1} = \\ &= [C(A - I)C^{-1}][(C^{-1})^{-1}(A + I)^{-1}C^{-1}], \end{aligned}$$

так как $(ABC)^{-1} = A^{-1}B^{-1}C^{-1}$, если A, B, C — матрицы.

Раскроем квадратные скобки и в результате получим

$$\begin{aligned} [C(A - I)C^{-1}][(C^{-1})^{-1}(A + I)^{-1}C^{-1}] &= [C(A - I)(A + I)^{-1}C^{-1}] = \\ &= E\{(A - I)(A + I)^{-1}\} \end{aligned}$$

в соответствии с (3.30), что и требовалось доказать.

$$3. E\{A\} + E^{-1}\{A\} = E\{A + A^{-1}\}.$$

Доказательство. Для доказательства этого свойства используем представление (3.30) и правило обращения матриц. Получим

$$\begin{aligned} E\{A\} + E^{-1}\{A\} &= [CAC^{-1} + (CAC^{-1})^{-1}] = [CAC^{-1} + \\ &+ (C^{-1})^{-1}A^{-1}C^{-1}] = CAC^{-1} + CA^{-1}C^{-1} = C(A + A^{-1})C^{-1} = \\ &= E\{A + A^{-1}\}, \end{aligned}$$

что и требовалось доказать.

Доказательства свойств 4 и 5 следуют непосредственно из свойств собственных чисел матриц [27], поэтому приведем эти свойства без доказательства.

4. Умножение матрицы на любое положительное число не меняет структуры собственных векторов этой матрицы.

5. Для выражения вида $E^* \{A\} = \frac{1}{2} [mA + nA^{-1}]$ значения собственных векторов матриц A и E преобразуются по формуле

$$\lambda_k(E) = \frac{1}{2} [m\lambda_k(A) + n\lambda_k^{-1}(A)],$$

где $\lambda_k(E)$, $\lambda_k(A)$ — k -е собственные числа матрицы; m , n — положительные числа.

Таким образом, после преобразования

$$S^*(\mathcal{H}) = (\mathcal{H} - I)(\mathcal{H} + I)^{-1}$$

собственные векторы матрицы сохраняются, их собственные числа изменяются таким образом, что те из них, которые были вне единичного круга, станут положительными, а те, которые находились внутри единичного круга, — отрицательными. Из свойств 2—5 следует, что итерационная процедура

$$S_k = \frac{1}{2} [(\det S_{k-1})^{-1} S_{k-1} + (\det S_{k-1}) S_{k-1}^{-1}] \quad (3.31)$$

при начальных условиях

$$S_0 = S^*(\mathcal{H}) \quad (3.32)$$

будет сходиться к стационарной матрице

$$S_{ст} = \rho \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \rho^{-1}. \quad (3.33)$$

Используя свойство 1, получаем окончательное выражение для искомой матрицы проектора (3.28)

$$S(\mathcal{H}) = \frac{1}{2} [S_{ст} + I]. \quad (3.34)$$

На рис. 7 показана блок-схема алгоритма расчета ковариационной матрицы ошибок фильтрации.

В блоке 1 осуществляется ввод исходных данных (расчет q_1 — q_4 и установка начальных значений переменных); в 2 формируется гамильтониан системы (3.27).

В блоках 3—7 реализуется итерационная процедура расчета по формулам (3.32), (3.31) до сходимости решения к стационарной матрице $S_{ст}$ (3.33); в 8 по формуле (3.34) рассчитывается матрица-проектор, а в 9 — значение искомой матрицы по формуле (3.29).

Расчет производной ковариационной матрицы ошибок фильтрации по параметрам адаптации можно провести по аналогичному алгоритму. Исследуем уравнение (3.25). С этой целью сделаем ряд необ-

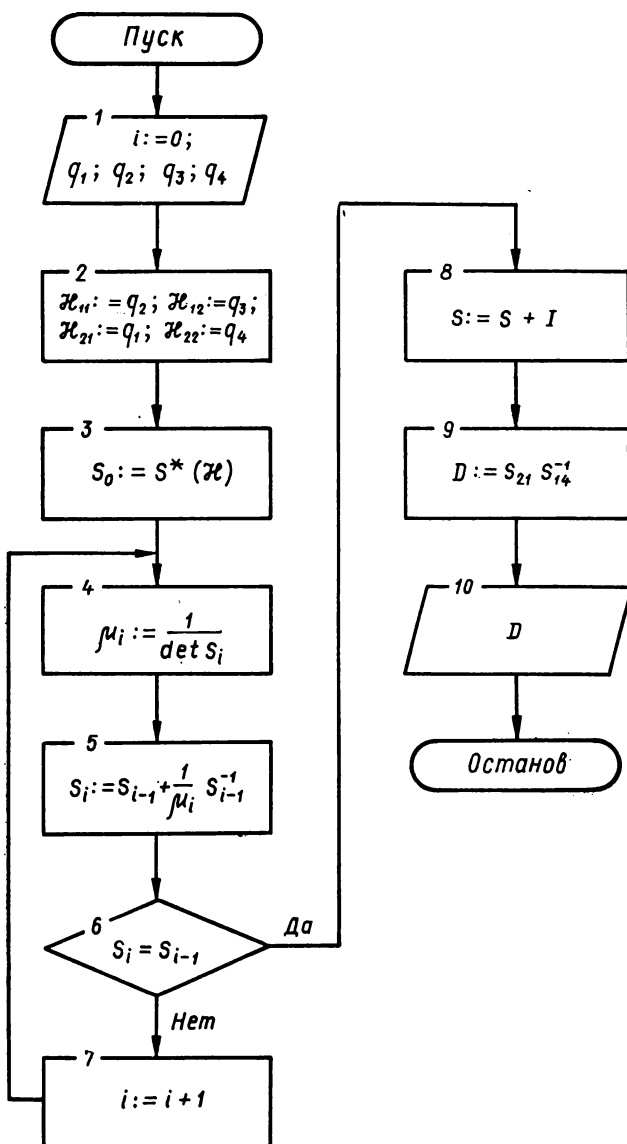


Рис. 7

ходимых преобразований

$$\frac{\partial D}{\partial q} q_1 D + D \frac{\partial q_1}{\partial Q} + D q_1 \frac{\partial D}{\partial Q} + \frac{\partial D}{\partial Q} q_2 + D \frac{\partial q_2}{\partial Q} - \frac{\partial q_3}{\partial Q} -$$

$$- q_3 \frac{\partial D}{\partial Q} - \frac{\partial q_4}{\partial Q} = 0.$$

Перегруппировав это уравнение, получим

$$\left(\frac{\partial}{\partial Q} q_1 D + \frac{\partial D}{\partial Q} q_2\right) + \left(D q_1 \frac{\partial D}{\partial Q} - q_3 \frac{\partial D}{\partial Q}\right) + \\ + \left(D \frac{\partial q_1}{\partial Q} + D \frac{\partial q_2}{\partial Q} - \frac{\partial q_3}{\partial Q} D - \frac{\partial q_4}{\partial Q}\right) = 0$$

или

$$\frac{\partial D}{\partial Q} (q_1 D + q_2) + (D q_1 - q_3) \frac{\partial D}{\partial Q} + \\ + \left(D \frac{\partial q_1}{\partial Q} D + D \frac{\partial q_2}{\partial Q} - \frac{\partial q_3}{\partial Q} D - \frac{\partial q_4}{\partial Q}\right) = 0.$$

Обозначим

$$\begin{aligned} \bar{q}_2 &= q_1 D + q_2, \\ \bar{q}_3 &= -D q_1 + q_3, \\ \bar{q}_4 &= -D \frac{\partial q_1}{\partial Q} - D \frac{\partial q_2}{\partial Q} + \frac{\partial q_3}{\partial Q} D + \frac{\partial q_4}{\partial Q}. \end{aligned} \quad (3.35)$$

В результате получим матричное уравнение для $\frac{\partial D}{\partial Q}$, аналогичное уравнению (2.20), в котором $q_1 = 0$:

$$\frac{\partial D}{\partial Q} \bar{q}_2 - \bar{q}_3 \frac{\partial D}{\partial Q} - \bar{q}_4 = 0.$$

Решая это уравнение так же, как и для D , будем использовать алгоритм, полученный для расчета ковариационной матрицы, предварительно пересчитав \bar{q}_2 , \bar{q}_3 , \bar{q}_4 по формулам (3.34).

Так как $q_1 = 0$, то гамильтониан такой системы имеет нижнетреугольную структуру. Обратная к этой матрица также имеет форму нижней треугольной [10]. Поэтому после каждой итерации матрица S_k в выражении (3.31) будет нижнетреугольной. В результате итерационного процесса матрица S_k в итоге сходится к стационарной матрице

$$S_{\text{ст}} = \begin{bmatrix} I & 0 \\ \nu & 0 \end{bmatrix}. \quad (3.36)$$

Согласно выражению (3.29)

$$\frac{\partial D}{\partial Q} = S_{21} S_{11}^{-1}.$$

Из формулы (3.36) следует, что $S_{11} = I$. Поэтому элемент ν является искомой производной, т. е.

$$\frac{\partial D}{\partial Q} = S_{21}. \quad (3.37)$$

На рис. 8 представлена блок-схема алгоритма расчета производной ковариационной матрицы ошибок фильтрации по параметрам адаптации.

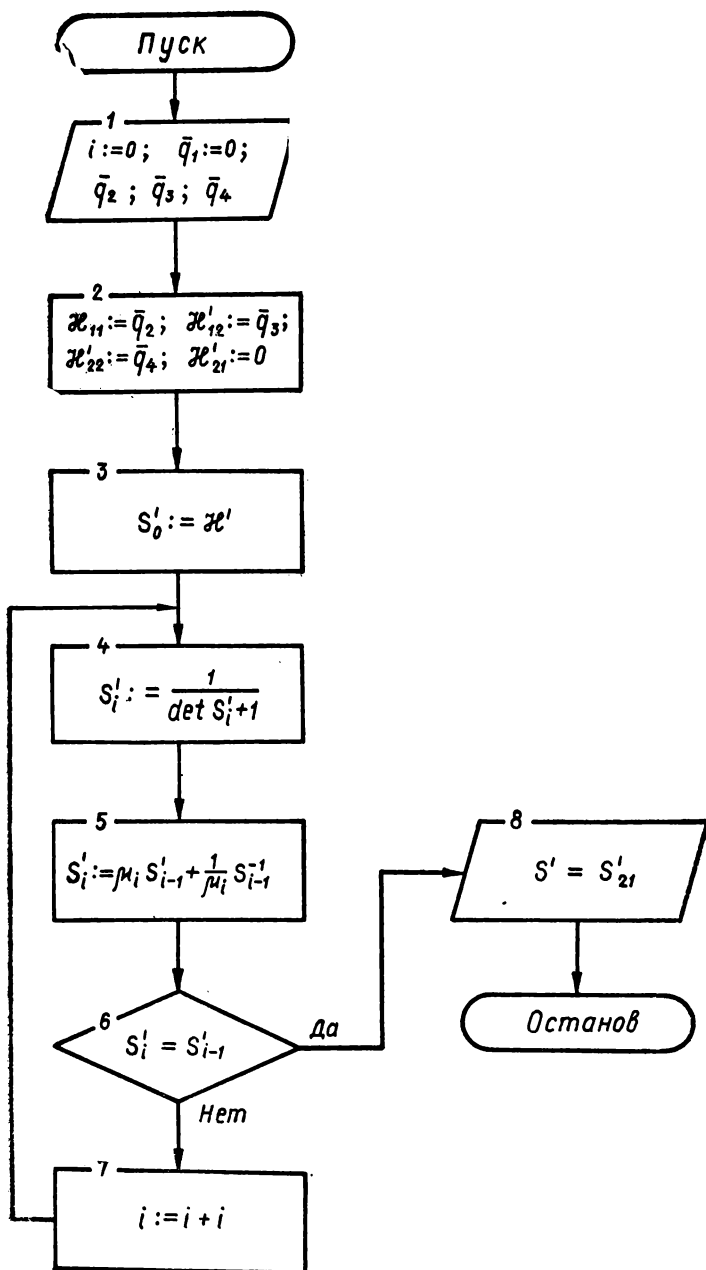


Рис. 8

В блоке 1 по формуле (3.35) рассчитываются $\bar{q}_2, \bar{q}_3, \bar{q}_4$ и устанавливаются начальные значения переменных; в 2 формируется гамильтониан системы.

В блоках 3—7 реализуется расчет стационарной матрицы $S_{ст}$, элемент S_{21} которой определяет искомую производную; в 8 осуществляется вывод результатов.

Разработанные алгоритмы расчета составляющих адаптивного дискретно-непрерывного фильтра придают ему свойство универсальности и расширяют область применения такого фильтра в практике исследования случайных процессов при обработке ТИ.

Разработанные вычислительные алгоритмы для расчета фильтра позволяют на основе обобщения полученных теоретических результатов построить схему оценивания в виде алгоритма адаптивной дискретно-непрерывной фильтрации результатов ТИ, обобщенная блок-схема которого показана на рис. 9.

В блоке 1 осуществляется ввод исходных данных. В качестве исходных данных задаются:

размерности переходных матриц \mathcal{F}, \mathcal{L} в СДУ процесса и матриц α, β в СДУ погрешностей измерений (2.4) в зависимости от состава вектора измерений, привлекаемого к обработке;

детерминированные компоненты в матрицах $\mathcal{F}, \mathcal{L}, \alpha, \beta$, не зависящие от параметров адаптации;

набор параметров адаптации и границы областей допустимых значений этих параметров;

дискретность измерений;

режим обработки: если обработка осуществляется в темпе поступления измерения, то задается интервал адаптации; если проводится послеполютная обработка, то задается объем выборки измерений;

начальные условия фильтрации;

начальные приближения для параметров адаптации.

В блоке 2 значения параметров адаптации сравниваются с наложенными на диапазон их возможного изменения ограничениями. В случае выхода величины параметров адаптации за пределы ограничений им присваивается ближайшее граничное значение.

В блоке 3 проводятся подготовительные вычисления, включающие расчет частных производных $f'(X_0)$ в модели измерений (2.1), расчет недетерминированных компонент α, β матриц \mathcal{F} и \mathcal{L} через параметры адаптации по формуле (2.4) и их производных по параметрам адаптации.

В блоке 4 формируются переходные матрицы A и B расширенной модели (2.5). В блоке 5 рассчитываются элементы дискретных матриц F и G , а также элементы a_1, a_2, b_1, b_2 и их производные по параметрам адаптации. На их основе формируются дискретные аналоги (2.8), (2.9) непрерывной модели (2.5).

В блоке 6 формируется частично наблюдаемый процесс (2.15) на основе расчета матрицы измерений H^* по формуле (2.13), ее производной по параметрам адаптации и шумовой составляющей по формуле (2.14). В блоке 7 рассчитывается ковариационная матрица ошибок фильтрации и ее производная по параметрам адаптации, а в блоке 8 —

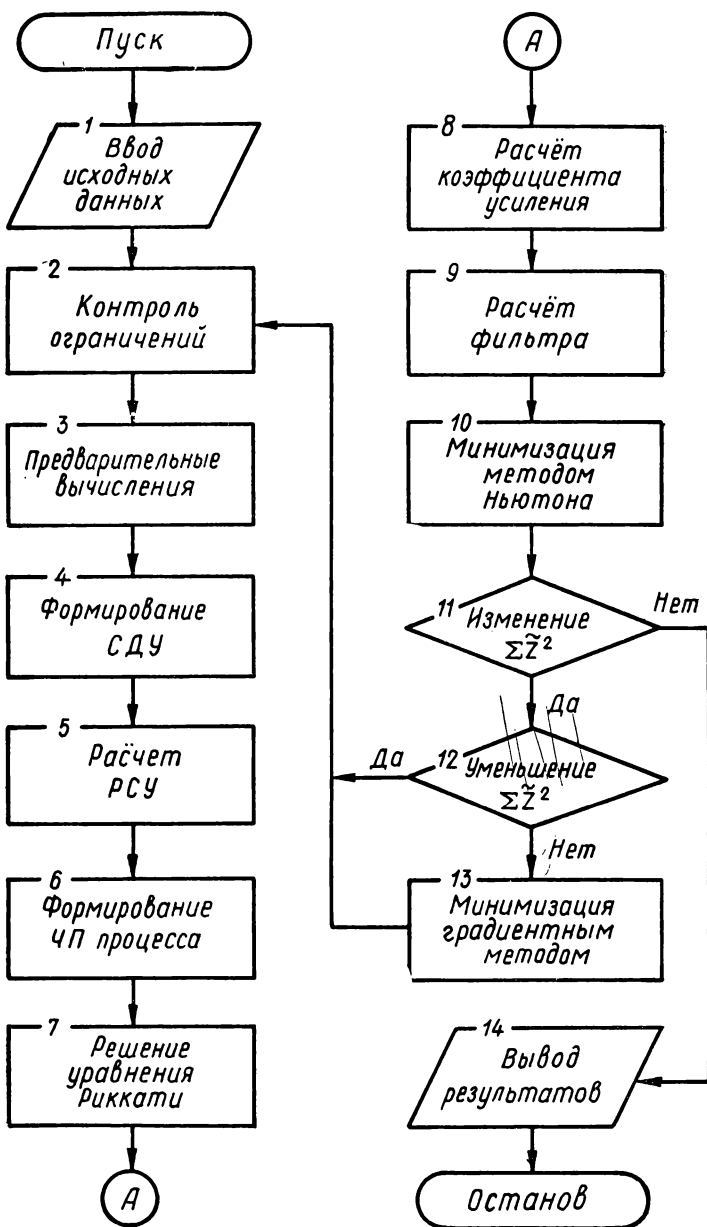


Рис. 9

коэффициент усиления фильтра по формуле (2.17) и его производная по параметрам адаптации в соответствии с выражением (2.28).

В блоке 9 в цикле поступления измерений рекуррентно рассчитываются: матрица измерений H_{0k}^* по формуле (2.12); ковариационная матрица ошибок прогноза P_k по формуле (2.19); оценки фильтруемого процесса M_k по формуле (2.16); невязки измерений Z_k по формуле (2.18); сумма квадратов этих невязок; производные вычислительных составляющих фильтра по параметрам адаптации в соответствии с выражениями (2.26) — (2.29).

В блоке 10 рассчитываются вектор частных производных по параметрам адаптации (2.24) и матрица вторых частных производных (2.25), а также реализуется вычислительная схема минимизации суммы квадратов невязок измерений методов Ньютона по формуле (2.23). В результате определяются новые значения параметров адаптации.

В блоке 11 текущая сумма квадратов невязок сравнивается с результатом, полученным на предыдущей итерации, при изменении величины этой суммы осуществляется переход к блоку 12.

В блоке 12 анализируется характер изменения текущей суммы квадратов невязок. В случае ее уменьшения переходим к блоку 2, в котором рассчитанные на этой итерации параметры проверяются на соответствие наложенным на них ограничениям. При увеличении суммы квадратов невязок, что свидетельствует о расхождении процесса минимизации методом Ньютона, переходим к блоку 13.

В блоке 13 реализуется вычислительная схема минимизации градиентным методом и осуществляется переход к блоку 2.

Итерационный процесс продолжается до тех пор, пока сумма квадратов невязок не стане минимальной.

В блоке 14 осуществляется вывод результатов в форме, удобной для потребителя.

Для численного расчета оптимальной в среднеквадратическом смысле оценки $\hat{X}(t)$, не зависящей от значений параметра регуляризации, следует рассчитать апостериорную плотность распределения этого параметра $p_t(\alpha)$, для чего необходимо знать функцию правдоподобия $\Phi_t(\alpha)$, которая вычисляется по формуле (2.40). Учтя дискретность измерений Δt , запишем интеграл в этой формуле суммой по формуле трапеций. Тогда приняв, что за интервал времени от 0 до t (пределы интегрирования) поступило N измерений, запишем функцию правдоподобия следующим образом:

$$\begin{aligned} \Phi_t(\alpha) &= [(2\pi)^M \det S_\alpha(t)]^{-1} \exp \left\{ -\frac{1}{2} \int_0^t \tilde{Z}^T(\tau) S_\alpha^{-1}(\tau) \tilde{Z}(\tau) d\tau \right\} = \\ &= \exp \left\{ -\frac{\Delta t}{4} \tilde{Z}^T(0) S_\alpha^{-1}(0) \tilde{Z}(0) - \right. \\ &\quad \left. - \frac{\Delta t}{2} \sum_{j=1}^{N-1} \tilde{Z}^T(j) S_\alpha^{-1}(j) \tilde{Z}(j) - \frac{\Delta t}{4} \tilde{Z}^T(N) S_\alpha^{-1}(N) \tilde{Z}(N) \right\}, \quad (3.38) \end{aligned}$$

где $Z(j) = Z(j) - H(j) \hat{X}_\alpha(j)$; $\hat{X}_\alpha(j)$ — регуляризованная оценка,

получаемая из РФК на j -м шаге с фиксированным значением параметра регуляризации.

Предположим, что априорная плотность распределения параметра регуляризации $\rho_0(\alpha)$ имеет максимум в некоторой точке $\alpha_{\text{опт}}$ и в области этого максимума она описывается нормальным законом. Тогда на выбранном интервале $\alpha \in [\alpha_0, \alpha_m]$ функция $\rho_0(\alpha)$ имеет вид

$$\rho_0(\alpha) = \frac{1}{\sqrt{2\pi}\sigma_\alpha} \exp \left\{ -\frac{1}{2} \frac{(\alpha - \alpha_{\text{опт}})^2}{\sigma_\alpha^2} \right\}, \quad (3.39)$$

где σ_α — среднее квадратическое отклонение α от $\alpha_{\text{опт}}$.

Исходя из этого и используя выражение (3.37), рассчитываем $\rho_t(\alpha)$ по формуле (2.41). Для этого также заменяем суммой по формуле трапеций интеграл в знаменателе, учитывая дискретность $\Delta\alpha$.

Получим

$$S_m = \int_a^b \Phi_t(\alpha) \rho_0(\alpha) d\alpha = \frac{\Delta\alpha}{2} \Phi_t(a) \rho_0(b) + \Delta\alpha \sum_{i=1}^{n-1} \Phi_t(\alpha_i) \rho_0(\alpha_i) + \\ + \frac{\Delta\alpha}{2} \Phi_t(a) \rho_0(b), \quad (3.40)$$

где $n = \frac{\alpha_m - \alpha_0}{\Delta\alpha}$ — число узлов в сетке значений параметра регуляризации на интервале $\alpha \in [\alpha_0, \alpha_m]$.

Тогда можно вычислить значения апостериорной плотности в узлах α_i :

$$\rho_t(\alpha_0 = a) = \frac{\Phi_t(\alpha_0) \rho_0(\alpha_0)}{S_m}; \\ \rho_t(\alpha_1 = \alpha_0 + \Delta\alpha) = \frac{\Phi_t(\alpha_1) \rho_0(\alpha_1)}{S_m}; \\ \vdots \\ \rho_t(\alpha_n = b) = \frac{\Phi_t(\alpha_n) \rho_0(\alpha_n)}{S_m}. \quad (3.41)$$

Таким образом, в соответствии с выражением (2.42) можно вычислить оценку

$$\hat{X}_{\text{опт}}(t) = \int_a^b \hat{X}_\alpha(t) \rho_t(\alpha) d\alpha = \frac{\Delta\alpha}{2} \hat{X}_{\alpha_0}(t) \rho_t(\alpha_0) + \\ + \Delta\alpha \sum_{i=1}^{n-1} \hat{X}_{\alpha_i}(t) \rho_t(\alpha_i) + \frac{\Delta\alpha}{2} \hat{X}_{\alpha_n}(t) \rho_t(\alpha_n), \quad (3.42)$$

где $\hat{X}_{\alpha_i}(t)$ — оценка, получаемая из РФК.

Корреляционная матрица погрешностей оптимальной оценки вычисляется в соответствии с формулой (2.43)

$$P(t) = \int_a^b \{ P_\alpha(t) + [\hat{X}_0(t) - X_\alpha(t)] [\hat{X}_0(t) - \hat{X}_\alpha(t)]^T \} \rho_t(\alpha) d\alpha \approx$$

$$\begin{aligned} &\approx \frac{\Delta\alpha}{2} \{P_{\alpha_0}(t) + [\hat{X}_0(t) - \hat{X}_{\alpha_0}(t)] [\hat{X}_0(t) - \hat{X}_{\alpha_0}(t)]^T\} \rho_t(\alpha_0) + \\ &+ \Delta\alpha \sum_{i=1}^{n-1} \{P_{\alpha_i}(t) + [\hat{X}_0(t) - \hat{X}_{\alpha_i}(t)] [\hat{X}_0(t) - \hat{X}_{\alpha_i}(t)]^T\} \rho_t(\alpha_i) + \\ &+ \frac{\Delta\alpha}{2} \{P_{\alpha_n}(t) + [\hat{X}_0(t) - \hat{X}_{\alpha_n}(t)] [\hat{X}_0(t) - \hat{X}_{\alpha_n}(t)]^T\} \rho_t(\alpha_n), \end{aligned}$$

где $P_\alpha(t)$ — апостериорная дисперсия РФК.

На рис. 10 приведена блок-схема разработанного алгоритма, реализуемого на однопроцессорном вычислителе. В качестве исходных данных в блоке 1 вводятся: граничные значения интервала $[\alpha_0, \alpha_m]$ для параметра регуляризации α , шаг дискретизации этого параметра $\Delta\alpha$, корреляционные матрицы погрешностей измерений и оценок, начальные значения вектора оцениваемых параметров $\sigma_\alpha, \alpha_{\text{опт}}$.

В блоке 2 устанавливается начальное значение параметра регуляризации, равное нижней границе интервала его изменения α_0 ; в 3 для текущего значения параметра регуляризации вычисляется априорная плотность распределения по формуле (3.39).

В блоке 4 рассчитывается субоптимальная оценка $\hat{X}_\alpha(t)$ РФК для текущего значения параметра регуляризации, а в 5 — функция правдоподобия $\Phi_t(\alpha)$ по формуле (3.38).

В блоке 6 значение текущего параметра регуляризации изменяется на величину $\Delta\alpha$; в 7 осуществляется проверка выхода текущего значения за верхнюю границу интервала его изменения. Если $\alpha < \alpha_m$ (рассчитаны еще не все субоптимальные оценки $\hat{X}_\alpha(t)$), то осуществляется переход к блоку 3.

В блоке 8 по известным значениям априорной плотности и функции правдоподобия для текущих узлов $\alpha_0, \alpha_1 = \alpha_0 + \Delta\alpha, \dots, \alpha_i = \alpha_{i-1} + \Delta\alpha = \alpha_m$ рассчитывается интеграл по формуле (3.40); в 9 по формуле (3.41) вычисляются значения апостериорной плотности распределения параметра регуляризации в узлах α_i .

В блоке 10 определяются оптимальная оценка $\hat{X}_{\text{опт}}(t)$ и корреляционная матрица ошибок $P(t)$ по формулам (3.42), (3.43); в 11 осуществляется вывод результатов.

§ 4. Некоторые вопросы коррекции систематических погрешностей

Многообразный практический опыт статистической обработки измерительной информации показывает, что часто эта измерительная информация наряду с обычными случайными (шумовыми) погрешностями содержит и некоторую неизвестную систематическую погрешность, которая в течение сеанса измерений либо остается постоянной, либо регулярно меняется по какому-нибудь закону. При этом в разных сеансах измерений она может оставаться той же или может каким-либо образом меняться. Иногда такие погрешности превосходят, причем значительно, случайные погрешности, содержащиеся в

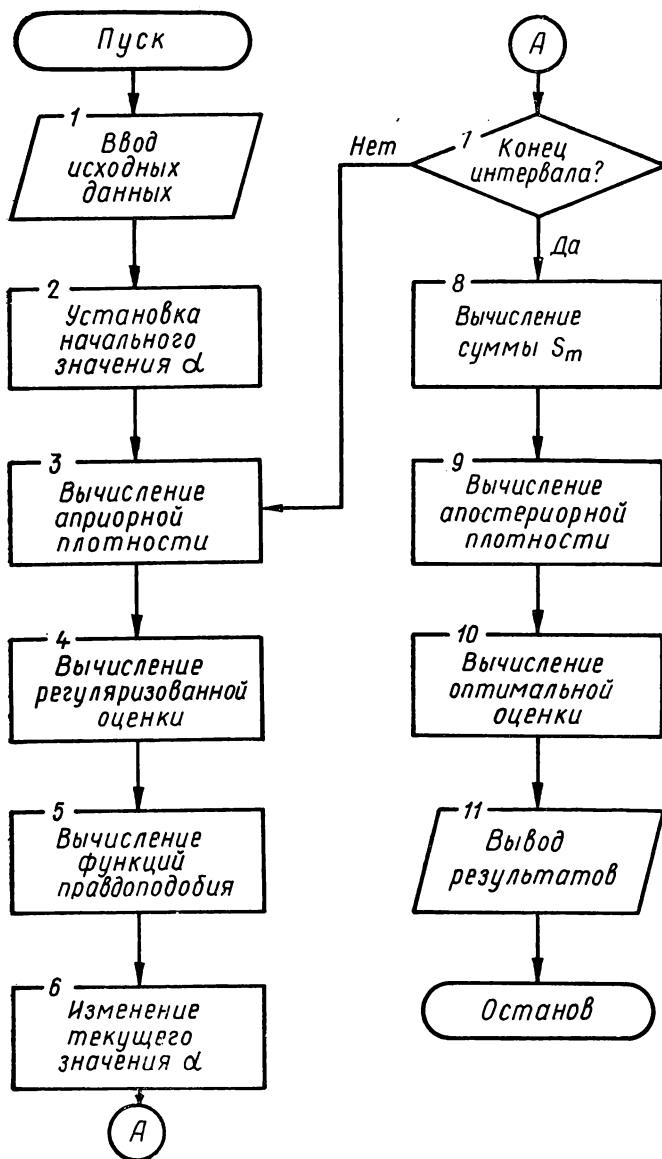


Рис. 10

обрабатываемых измерениях, и недопустимо искажают результаты обработки. Поэтому борьба с такими погрешностями, их подавление или, как чаще говорят, коррекция систематических погрешностей в процессах обработки измерительной информации приобретает весьма важное значение.

Традиционно математические методы коррекции систематических погрешностей измерительной информации основываются на использовании в той или иной форме так называемых структурно избыточных измерений, когда протекающий физический процесс измеряется одновременно многими однотипными или разнотипными приборами, каждый из которых может иметь свою собственную систематическую погрешность. Методы такого типа позволяют скорректировать (уменьшить) систематические погрешности, в сущности, до некоторого среднестатистического уровня, достигаемого на совокупности реализовавшихся систематических погрешностей.

Существуют, однако, методы иного типа, основанные на использовании так называемых контрольных условий: часто те или иные характеристики наблюдаемого (измеряемого) физического процесса подчинены некоторым условиям, связям (например, должны удовлетворять заданным дифференциальным или иным уравнениям). Тогда эти условия, связи можно трансформировать в условия, которым должны удовлетворять результаты идеальных измерений. Реальные же измерения в таком случае перестанут тождественно удовлетворять этим уравнениям и по возникающим расхождениям можно судить о характере и величинах погрешностей, содержащихся в анализируемой измерительной информации.

Ниже рассматриваются некоторые теоретические вопросы, связанные с таким методом коррекции систематических погрешностей.

Одномерный процесс. Прямые измерения. Линейные связи. Начнем с простейшего случая обработки измерений. Пусть в течение интервала времени $[\alpha, \beta]$ протекает некоторый физический процесс, описываемый функцией $x(t)$, которая считается неизвестной, и пусть в дискретные моменты времени

$$t_1, t_2, \dots, t_N \quad (4.1)$$

из $[\alpha, \beta]$ измеряются значения этой функции. Предположим, что погрешности, содержащиеся в результатах измерений в каждый момент времени t_i , представляют собой сумму двух слагаемых ξ_i и η_i . При этом погрешности $\xi_1, \xi_2, \dots, \xi_N$ являются случайными нормально распределенными величинами, каждая из которых имеет нулевое математическое ожидание и заданную дисперсию

$$\sigma_i^2 = M \{ \xi_i^2 \}, \quad i = 1, 2, \dots, N,$$

причем

$$M \{ \xi_i \xi_j \} = 0, \quad i \neq j, \quad i, j = 1, 2, \dots, N.$$

Случайные величины $\xi_1, \xi_2, \dots, \xi_N$ в дальнейшем будем рассматривать как значения некоторого гауссовского процесса $\xi(t)$ в моменты времени (4.1).

Относительно величин $\eta_1, \eta_2, \dots, \eta_N$ предполагаем, что они представляют собой так называемую систематическую погрешность измерений и являются значениями некоторой известной (детерминированной или случайной) функции $\eta(t)$ в моменты времени (4.1). Считаем, что $\eta(t)$ аппроксимируется выражением

$$\zeta(t) = \sum_{k=1}^m c_k \Phi_k(t), \quad (4.2)$$

где коэффициенты c_k — неизвестные величины, а

$$\Phi_1(t), \Phi_2(t), \dots, \Phi_m(t) \quad (4.3)$$

являются известными линейно независимыми функциями, заданными на интервале $[\alpha, \beta]$. В дальнейшем будут сформулированы дополнительные требования к этим функциям. Погрешность такой аппроксимации

$$\kappa(t) = \eta(t) - \zeta(t)$$

будет зависеть от того, какими выбраны функции (4.3), от их числа m и от коэффициентов c_k . Считая $\kappa(t)$ малой величиной, пока пренебрегаем ею, т. е. в качестве систематической погрешности измерений будем рассматривать лишь ее учитываемую долю — функцию $\zeta(t)$. В конце этого раздела будет показан характер влияния неучтенной величины $\kappa(t)$ на точность оценивания систематической погрешности.

Таким образом, результаты измерений имеют структуру

$$y(t_i) = x(t_i) + \xi_i + \sum_{k=1}^m c_k \Phi_k(t_i), \quad i = 1, 2, \dots, N. \quad (4.4)$$

Введем некоторые обозначения, необходимые для дальнейшего изложения. Пусть E — N -мерное евклидовое пространство со скалярным произведением

$$(e', e'')_E = \sum_{i=1}^N \frac{e'_i e''_i}{\sigma_i^2}$$

и нормой

$$\|e'\|_E = \sqrt{(e', e')_E}$$

его элементов $e' = (e'_1, e'_2, \dots, e'_N)$ и $e'' = (e''_1, e''_2, \dots, e''_N)$. Условимся произвольной вещественной скалярной функции $u(t)$, определенной на интервале $[\alpha, \beta]$, ставить в соответствие в пространстве E вектор $\bar{u} = (u(t_1), u(t_2), \dots, u(t_N))$. Тогда измерения (4.4) можно записать в векторной форме

$$\bar{y} = \bar{x} + \bar{\xi} + \sum_{k=1}^m c_k \bar{\Phi}_k. \quad (4.5)$$

При сглаживании этих измерений поступают по-разному. В простейшем случае их сглаживают по методу наименьших квадратов какими-либо линейными формами, игнорируя наличие в измерениях систематических погрешностей. При этом, как известно, систематическая погрешность практически не поддается фильтрации. В других случаях систематическая погрешность аппроксимируется какими-либо

функциями с неизвестными параметрами, которые включаются в число параметров, подлежащих статистической оценке по полученным измерениям.

Ниже рассмотрим метод коррекции систематических погрешностей, основанный на использовании априорной или дополнительной информации о наблюдаемом процессе.

Выясним прежде всего, как сказывается на результате сглаживания измерений (4.4) по методу наименьших квадратов наличие в этих измерениях систематической погрешности вида (4.2). С этой целью сначала рассмотрим обычную задачу сглаживания измерений в предположении, что они не содержат систематической погрешности и имеют, следовательно, структуру

$$z(t_i) = x(t_i) + \xi_i, \quad i = 1, 2, \dots, N, \quad (4.6)$$

или в векторной форме

$$\bar{z} = \bar{x} + \bar{\xi}. \quad (4.6a)$$

В силу предположений о свойствах случайных погрешностей $\xi_1, \xi_2, \dots, \xi_N$ сглаживать измерения (4.6) будем по методу наименьших квадратов линейной формой относительно заранее выбранной системы функций

$$\omega_1(t), \omega_2(t), \dots, \omega_n(t), \quad (4.7)$$

линейно независимых на множестве точек (4.1), т. е. таких, что равенства

$$\alpha_1 \omega_1(t_i) + \alpha_2 \omega_2(t_i) + \dots + \alpha_n \omega_n(t_i) = 0, \quad i = 1, 2, \dots, N,$$

выполняются лишь при $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$. Помимо этого обычного требования в дальнейшем на функции (4.7) будет наложено еще одно требование. В остальном эти функции могут быть произвольными, в частности, степенными (они приводят к полиномиальному сглаживанию), тригонометрическими (они обуславливают сглаживание тригонометрическими полиномами) или какими-либо иными.

Системе функций (4.7) по введенному выше правилу в пространстве E соответствует система линейно независимых векторов

$$\bar{\omega}_1, \bar{\omega}_2, \dots, \bar{\omega}_n. \quad (4.8)$$

В соответствии с методом наименьших квадратов будем искать такую линейную форму

$$z^*(t) = \sum_{i=1}^n b_i^* \omega_i(t),$$

которая удовлетворяет условию минимума

$$\|\bar{z} - \bar{z}^*\|_E^2 = \min_{b_1, \dots, b_n} \left\| \bar{z} - \sum_{i=1}^n b_i \bar{\omega}_i \right\|_E^2.$$

Из этого условия обычным способом находим значения коэффициентов

$$b_j^* = \sum_{p=1}^n W_{jp}(\bar{\omega}_p, \bar{z})_E,$$

где W_{jp} , $j, p = 1, 2, \dots, n$, — элементы матрицы W , обратной к матрице Грама Ω , порожденной векторами (4.8). Поскольку эти векторы линейно независимы, то матрица $W = \Omega^{-1}$ существует.

Заметим, что если векторы (4.8) ортонормированы, то

$$b_j^* = (\bar{\omega}_j, \bar{z})_E, \quad j = 1, 2, \dots, n.$$

Для дальнейшего нам важно выяснить структуру сглаживающей функции $z^*(t)$. С этой целью введем в рассмотрение величины

$$x_i^* = \sum_{j=1}^n W_{ij} (\bar{\omega}_j, \bar{x})_E,$$

$$\xi_i^* = \sum_{j=1}^n W_{ij} (\bar{\omega}_j, \bar{\xi}), \quad i = 1, 2, \dots, n.$$

Замечание. Поскольку в (4.6а) векторы \bar{x} и $\bar{\xi}$ порознь нам не известны (известна лишь их сумма — вектор \bar{z}), то величины x_i^* и ξ_i^* на практике определить, конечно, нельзя. Ниже они используются только для теоретического (качественного) анализа структуры сглаживающих функций и не входят ни в какие практические методы или алгоритмы.

С помощью величин x_i^* и ξ_i^* образуем функции

$$x^*(t) = \sum_{i=1}^n x_i^* \omega_i(t),$$

$$\xi^*(t) = \sum_{i=1}^n \xi_i^* \omega_i(t),$$

которые будут наилучшими аппроксимациями измеряемой функции $x(t)$ и гауссовского случайного процесса $\xi(t)$, причем наилучшими по такому же критерию, какой использован выше при сглаживании измерений. Следовательно, функция $z^*(t)$ представляет собой сумму двух функций

$$z^*(t) = x^*(t) + \xi^*(t),$$

первая из которых аппроксимирует идеальную измеряемую функцию $x(t)$, а вторая является статистической оценкой математического ожидания случайного процесса $\xi(t)$, которое в данном случае равно нулю (в силу сделанного выше предположения о случайных погрешностях $\xi_1, \xi_2, \dots, \xi_N$). Следовательно, погрешность

$$r(t) = x(t) - z^*(t) \quad (4.9)$$

есть сумма погрешностей

$$r_x(t) = x(t) - x^*(t)$$

и

$$r_\xi(t) = 0 - \xi^*(t).$$

Из теории приближения функций и из математической статистики известно, что при достаточно большом числе N точек (4.1) и при удачном

(с точки зрения качества аппроксимации функции $x(t)$) выборе системы функций (4.7) погрешность (4.9) будет достаточно малой величиной. Подчеркнем этот факт, поскольку он нам понадобится в дальнейшем.

Перейдем теперь к задаче сглаживания измерений (4.4), содержащих систематическую погрешность (4.2). Сглаживать их будем, как и измерения (4.6), по методу наименьших квадратов относительно той же системы функций (4.7), игнорируя при этом наличие в измерениях систематической погрешности (4.2). Иными словами, в соответствии с методом наименьших квадратов найдем такую линейную форму

$$y^*(t) = \sum_{i=1}^n a_i^* \omega_i(t), \quad (4.10)$$

коэффициенты которой $a_1^*, a_2^*, \dots, a_n^*$ определяются из условия минимума

$$\|\bar{y} - \bar{y}^*\|_E^2 = \min_{a_1, \dots, a_n} \left\| \bar{y} - \sum_{i=1}^n a_i \bar{\omega}_i \right\|_E^2.$$

Как и в предыдущем случае, находим

$$a_i^* = \sum_{j=1}^n W_{ij} (\bar{\omega}_j, \bar{y})_E, \quad i = 1, 2, \dots, n.$$

Если векторы (4.8) будут ортонормированными, то

$$a_i^* = (\bar{\omega}_i, \bar{y})_E, \quad i = 1, 2, \dots, n.$$

Выясним теперь структуру оценок $a_1^*, a_2^*, \dots, a_n^*$ и сглаживающей функции $y^*(t)$. Нетрудно показать, что

$$a_i^* = b_i^* + \sum_{k=1}^m c_k \Phi_{ik}^*,$$

где

$$\Phi_{ik}^* = \sum_{j=1}^n W_{ij} (\bar{\omega}_j, \bar{\Phi}_k), \quad i = 1, 2, \dots, n; \quad k = 1, 2, \dots, m. \quad (4.11)$$

Следовательно,

$$y^*(t) = z^*(t) + \sum_{k=1}^m c_k \Phi_k^*(t) \quad (4.12)$$

или

$$y^*(t) = x^*(t) + \xi^*(t) + \zeta^*(t), \quad (4.13)$$

где

$$\Phi_k^*(t) = \sum_{i=1}^n \Phi_{ik}^* \omega_i(t), \quad (4.14)$$

$$\zeta^*(t) = \sum_{k=1}^m c_k \Phi_k^*(t). \quad (4.15)$$

Таким образом, если измерения, содержащие систематическую погрешность, сглаживать линейной формой по методу наименьших квадратов и не учитывать эту погрешность, то сглаживающая функция $y^*(t)$ будет включать помимо функций $x^*(t)$ и $\xi^*(t)$, о которых уже шла речь выше, еще и функцию $\zeta^*(t)$, аппроксимирующую погрешность $\zeta(t)$.

В частном случае, когда в качестве функций (4.3), используемых для описания систематической погрешности, берутся (при $m \leq n$) функции из системы (4.7), с помощью которой строятся сглаживающие линейные формы, то в этом случае, как нетрудно видеть, будет иметь место равенство $\zeta^*(t) = \zeta(t)$. Из этого равенства следует, что в данном частном случае систематическая погрешность $\zeta(t)$ полностью воспроизводится в сглаживающей функции $y^*(t)$.

Итак, обычное сглаживание измерений линейными формами по методу наименьших квадратов практически не дает возможности «отфильтровать» систематическую погрешность, содержащуюся в сглаживаемых измерениях. Следовательно, возникает проблема коррекции либо непосредственно измерений $y(t_i)$, либо функции $y^*(t)$, сглаживающей эти измерения. Иными словами, возникает проблема приближенного определения неизвестных коэффициентов c_1, c_2, \dots, c_m в выражении (4.2) для систематической погрешности.

Один из путей решения этой проблемы состоит в специальном использовании той или иной априорной или дополнительной информации о поведении наблюдаемого процесса. Ниже в рамках этого направления предлагаются методы коррекции измерений, основанные на использовании различного рода соотношений (связей), которым может удовлетворять измеряемый процесс. Природа таких соотношений (связей) может быть самой разнообразной: геометрической, интегральной, кинематической, динамической, комбинированной. Соотношения (связи) могут быть как линейными, так и нелинейными, как скалярными, так и векторными. В этом параграфе мы начинаем с простейшего случая, когда функция $x(t)$, описывающая наблюдаемый одномерный процесс, удовлетворяет линейному операторному (в общем случае) соотношению. В дальнейшем будут рассмотрены и другие возможности.

Итак, предположим, что функция $x(t)$, описывающая наблюдаемый процесс, является элементом некоторого линейного функционального пространства \mathcal{X} . На множестве D этого пространства, содержащем функцию $x(t)$, определен линейный оператор L , переводящий произвольную функцию $u(t) \in D$ в евклидово или гильбертово пространство H со скалярным произведением $(h, g)_H$ и нормой $\|h\|_H = \sqrt{(h, h)_H}$ его элементов h и g . Пусть во множестве $L(D) \subset H$ задан элемент h и пусть функция $x(t)$ удовлетворяет условию

$$Lx = h. \quad (4.16)$$

Возможна альтернатива: либо это условие позволяет однозначно определить функцию $x(t)$, либо оно не позволяет сделать этого. В первом случае задача статистической оценки функции $x(t)$ теряет смысл, и поэтому его рассматривать не будем.

Отметим, что оператор L определяется «физикой» каждой конкретной решаемой задачи. В рассматриваемом ниже способе от него помимо линейности требуется лишь, чтобы он удовлетворял еще одному условию, которое будет сформулировано позже. В остальном этот оператор может быть произвольным. Трудность здесь заключается в том, как в каждой конкретной задаче найти этот оператор, выписать его.

В некоторых случаях может оказаться, что функция, описывающая наблюдаемый процесс, удовлетворяет нескольким соотношениям типа (4.16):

$$L^i x = h^i, \quad i = 1, 2, \dots, \mu. \quad (4.17)$$

Нетрудно, однако, показать, что случай с несколькими соотношениями легко сводится к случаю с одним соотношением. В самом деле, не нарушая общности, предположим, что каждый из операторов L^i определен на одном и том же множестве пространства \mathfrak{X} . Однако действовать эти операторы могут, вообще говоря, в различные евклидовы или гильбертовы пространства. Пусть, скажем, L^i действует в (евклидово или гильбертово) пространство H^i со скалярным произведением $(h^i, g^i)_{H^i} \quad \forall h^i, g^i \in H^i$ и нормой $\|h^i\|_{H^i} = \sqrt{(h^i, h^i)_{H^i}} \quad \forall h^i \in H^i$. Рассмотрим декартово произведение пространств H^1, \dots, H^μ :

$$H = H^1 \times H^2 \times \dots \times H^\mu$$

со скалярным произведением

$$(h, g)_H = \sum_{i=1}^{\mu} (h^i, g^i)_{H^i} \quad \forall h, g \in H$$

и нормой

$$\|h\|_H = \sqrt{(h, h)_H} \quad \forall h \in H,$$

где $h = (h^1, h^2, \dots, h^\mu)$, $g = (g^1, g^2, \dots, g^\mu)$, $h^i, g^i \in H^i$, $i = 1, 2, \dots, \mu$.

Введем далее оператор

$$L = (L^1, L^2, \dots, L^\mu)$$

с областью определения D , действующий из \mathfrak{X} в H , положив

$$Lu = (L^1 u, L^2 u, \dots, L^\mu u) \quad \forall u(t) \in D.$$

Тогда систему соотношений (4.17) можно записать одним соотношением (4.16). Таким образом, случай с несколькими соотношениями сведен к случаю с одним соотношением.

Итак, пусть функция $x(t)$ удовлетворяет соотношению (4.16). Из выражения (4.9) следует, что

$$z^*(t) = x(t) - r(t).$$

Подставив $z^*(t)$ в равенство (4.12), придем к соотношению

$$y^*(t) = x(t) - r(t) + \sum_{k=1}^m c_k \Phi_k^*(t),$$

из которого находим

$$x(t) = y^*(t) - \sum_{k=1}^m c_k \Phi_k^*(t) + r(t).$$

Поскольку (идеальная) измеряемая функция $x(t)$ должна удовлетворять условию (4.16), то получаем тождество

$$L\left(y^*(t) - \sum_{k=1}^m c_k \Phi_k^*(t) + r(t)\right) - h = 0.$$

В этом тождестве неопределенными остаются коэффициенты c_1, c_2, \dots, c_m и погрешность $r(t)$. Считая величину $r(t)$ малой и пренебрегая ею, получаем ненулевую невязку

$$\Delta = L\left(y^*(t) - \sum_{k=1}^m c_k \Phi_k^*(t)\right) - h \quad (4.18)$$

в пространстве H .

В силу линейности оператора L получаем

$$\Delta = Ly^*(t) - h - \sum_{k=1}^m c_k \Phi_k^*(t) = F - \sum_{k=1}^m c_k \Phi_k, \quad (4.18a)$$

где

$$F = Ly^*(t) - h, \quad (4.19)$$

$$\Phi_k = L\Phi_k^*(t), \quad k = 1, 2, \dots, m. \quad (4.20)$$

Поскольку оператор L и элемент $h \in H$ заданы заранее, функция $y^*(t)$ найдена выше путем сглаживания измерений, а функция $\Phi_k^*(t)$ — путем сглаживания значений функций $\Phi_k(t)$ в точках (4.1), то F и Φ_k являются известными элементами в пространстве H . Следовательно, невязка Δ есть функция от неизвестных параметров c_1, c_2, \dots, c_m , входящих в выражение (4.2) для систематической погрешности. В качестве приближенных значений этих параметров возьмем такие значения $c_1^*, c_2^*, \dots, c_m^*$, которые доставляют минимум квадрату нормы невязки:

$$\begin{aligned} \|\Delta\|_H^2 &= \left\| F - \sum_{k=1}^m c_k \Phi_k \right\|_H^2 = \|F\|_H^2 - 2 \sum_{k=1}^m c_k (\Phi_k, F)_H + \\ &+ \sum_{k_1=1}^m \sum_{k_2=1}^m c_{k_1} c_{k_2} (\Phi_{k_1}, \Phi_{k_2})_H. \end{aligned} \quad (4.21)$$

Для определения $c_1^*, c_2^*, \dots, c_m^*$ из условия минимума

$$\|\Delta\|_H^2 = \min$$

обычным путем получаем систему линейных уравнений (в векторной форме)

$$\Phi C^* = f, \quad (4.22)$$

где

$$\Phi = \begin{pmatrix} (\Phi_1, \Phi_1)_H & (\Phi_1, \Phi_2)_H & \dots & (\Phi_1, \Phi_m)_H \\ (\Phi_2, \Phi_1)_H & (\Phi_2, \Phi_2)_H & \dots & (\Phi_2, \Phi_m)_H \\ \dots & \dots & \dots & \dots \\ (\Phi_m, \Phi_1)_H & (\Phi_m, \Phi_2)_H & \dots & (\Phi_m, \Phi_m)_H \end{pmatrix},$$

Φ — матрица Грама, порожденная векторами (4.20); C^* — вектор-столбец с компонентами $c_1^*, c_2^*, \dots, c_m^*$; f — вектор-столбец с компонентами $(\Phi_1, F)_H, (\Phi_2, F)_H, \dots, (\Phi_m, F)_H$.

Если матрица Φ имеет обратную матрицу $V = \Phi^{-1}$ и если $V_{kj}, k, j = 1, 2, \dots, m$, суть элементы матрицы V , то

$$C^* = Vf \quad (4.23)$$

или

$$c_k^* = \sum_{j=1}^m V_{kj} (\Phi_j, F)_H, \quad k = 1, 2, \dots, m. \quad (4.23a)$$

Теперь можно найти приближенное значение

$$\tilde{\zeta}(t) = \sum_{k=1}^m c_k^* \Phi_k(t) \quad (4.24)$$

для систематической погрешности $\zeta(t)$ и приближенное значение

$$\tilde{\zeta}^*(t) = \sum_{k=1}^m c_k^* \Phi_k^*(t) \quad (4.25)$$

для функции $\zeta^*(t)$, входящей в состав сглаживающей функции $y^*(t)$, а затем ввести поправки (коррекцию) либо в измерения (4.4):

$$\tilde{y}(t_i) = y(t_i) - \tilde{\zeta}(t_i), \quad i = 1, 2, \dots, N, \quad (4.26)$$

либо в сглаживающую функцию $y^*(t)$:

$$\tilde{y}(t) = y^*(t) - \tilde{\zeta}^*(t). \quad (4.27)$$

Выясним теперь структуры скорректированных измерений и сглаживающей функции $\tilde{y}^*(t)$. С этой целью преобразуем выражение (4.23a) для c_k^* :

$$\begin{aligned} c_k^* &= \sum_{j=1}^m V_{kj} (\Phi_j, F)_H = \sum_{j=1}^m V_{kj} (\Phi_j, Ly^*(t) - h)_H = \\ &= \sum_{j=1}^m V_{kj} \left(\Phi_j, L \left[z^*(t) + \sum_{p=1}^m c_p \Phi_p^*(t) \right] - h \right)_H = \\ &= \sum_{j=1}^m V_{kj} \left(\Phi_j, Lz^*(t) - h + \sum_{p=1}^m c_p L\Phi_p^*(t) \right)_H = \\ &= \sum_{j=1}^m V_{kj} \left(\Phi_j, \sum_{p=1}^m c_p \Phi_p - L(x(t) - z^*(t)) \right)_H = \\ &= \sum_{p=1}^m c_p \sum_{j=1}^m V_{kj} (\Phi_j, \Phi_p)_H - \sum_{j=1}^m V_{kj} (\Phi_j, Lr(t))_H = c_k + r_k, \end{aligned}$$

где

$$r_k = \sum_{j=1}^m V_{kj} (\Phi_j, Lr(t))_H \quad (4.28)$$

является методической погрешностью, возникающей при замене c_k его приближенным значением c_k^* . Эта методическая погрешность порождена (в условиях, рассматриваемых в настоящем параграфе) погрешностью $r(t)$, определяемой выражением (4.9). Таким образом,

$$c_k^* = c_k - r_k, \quad k = 1, 2, \dots, m.$$

Тогда в силу (4.24) получаем

$$\tilde{\zeta}(t) = \sum_{k=1}^m c_k^* \Phi_k(t) = \sum_{k=1}^m c_k \Phi_k(t) - \sum_{k=1}^m r_k \Phi_k(t) = \zeta(t) - \rho(t),$$

где

$$\rho(t) = \sum_{k=1}^m r_k \Phi_k(t). \quad (4.29)$$

Поэтому скорректированные измерения (4.26) имеют структуру

$$\begin{aligned} \tilde{y}(t_i) &= y(t_i) - \tilde{\zeta}(t_i) = x(t_i) + \xi_i + \zeta(t_i) - \zeta(t_i) + \rho(t_i) = \\ &= x(t_i) + \xi_i + \rho(t_i), \end{aligned} \quad (4.30)$$

т. е. систематическая погрешность $\zeta(t)$ при корректировке измерений исключается полностью, зато вместо нее появляется методическая погрешность $\rho(t)$, имеющая вид (4.29).

Аналогично скорректированная сглаживающая функция (4.27) имеет структуру

$$\tilde{y}^*(t) = x^*(t) + \xi^*(t) + \rho^*(t), \quad (4.31)$$

где

$$\rho^*(t) = \sum_{k=1}^m r_k \Phi_k^*(t), \quad (4.32)$$

т. е. систематическая погрешность $\zeta^*(t)$, содержащаяся в сглаживающей функции $y^*(t)$, при корректировке исключается полностью, зато вместо нее в новой сглаживающей функции появляется методическая погрешность $\rho^*(t)$ вида (4.32).

Выясним теперь условия, при которых возможна описанная выше коррекция, т. е. возможно определение величин c_1^*, \dots, c_m^* . Иными словами, выясним условия, при которых решение системы (4.22) существует и единственно, т. е. при которых существует матрица $V = \Phi^{-1}$.

Поскольку Φ является матрицей Грама, порожденной векторами (4.20), то для ее невырожденности необходима и достаточна линейная независимость элементов (4.20). Следовательно, достаточно выяснить условия, при которых эти элементы будут линейно независимыми.

Заметим прежде всего, что

$$\begin{aligned}\Phi_k &= L\varphi_k^*(t) = L\left(\sum_{j=1}^n \varphi_j^* \omega_j(t)\right) = \sum_{j=1}^n \varphi_{kj}^* L\omega_j(t) = \\ &= \sum_{j=1}^n \left(\sum_{p=1}^n W_{jp}(\bar{\omega}_p, \bar{\varphi}_k)_E\right) L\omega_j(t).\end{aligned}\quad (4.33)$$

Покажем сначала, что для линейной независимости элементов (4.20) необходимо, чтобы было $m \leq n$, т. е. чтобы число функций (4.3) в выражении (4.2), аппроксимирующем систематическую погрешность $\xi(t)$, не превосходило числа функций (4.7), с помощью которых осуществляется сглаживание измерений и среднеквадратическая аппроксимация значений тех или иных функций в точках (4.1). В самом деле, если $m > n$, то m n -мерных векторов

$$\Psi_l = \begin{pmatrix} (\bar{\omega}_1, \bar{\varphi}_l)_E \\ (\bar{\omega}_2, \bar{\varphi}_l)_E \\ \vdots \\ (\bar{\omega}_n, \bar{\varphi}_l)_E \end{pmatrix}, \quad l = 1, 2, \dots, m, \quad (4.34)$$

всегда будут линейно зависимыми. Следовательно, в этом случае (т. е. при $m > n$) найдутся такие постоянные $\alpha_1, \dots, \alpha_m$, что

$$\alpha_1^2 + \alpha_2^2 + \dots + \alpha_m^2 > 0$$

и

$$\sum_{k=1}^m \alpha_k \Psi_k = 0.$$

Последнее равенство эквивалентно системе равенств

$$\sum_{k=1}^m \alpha_k (\bar{\omega}_p, \bar{\varphi}_k)_E = 0, \quad p = 1, 2, \dots, n.$$

Поэтому для тех же $\alpha_1, \alpha_2, \dots, \alpha_m$ получаем в силу (4.33)

$$\sum_{k=1}^m \alpha_k \Phi_k = \sum_{j=1}^n L\omega_j(t) \sum_{p=1}^n W_{jp} \sum_{k=1}^m \alpha_k (\bar{\omega}_p, \bar{\varphi}_k)_E = 0,$$

что и означает линейную зависимость элементов (4.20) между собой.

Таким образом, для невырожденности матрицы Φ в (4.22) необходимо выполнение условия $m \leq n$.

Итак, пусть $m \leq n$. Предположим теперь, что система функций (4.3) выбрана так, что порожденные ею векторы (4.34) линейно независимы. Такие системы существуют. Например, в качестве функций (4.3) можно взять все или часть функций (4.7). Покажем, что если векторы (4.34) линейно независимы, то одновременно с ними линейно незави-

симыми будут и n -мерные векторы

$$\Phi_k^* = \begin{pmatrix} \Phi_{1k}^* \\ \Phi_{2k}^* \\ \vdots \\ \Phi_{nk}^* \end{pmatrix}, \quad k = 1, 2, \dots, m, \quad (4.35)$$

компоненты которых определяются равенствами (4.11). Предположим противное, т. е. предположим, что векторы (4.35) линейно зависимы. Тогда найдутся такие постоянные $\beta_1, \beta_2, \dots, \beta_m$, что

$$\beta_1^2 + \beta_2^2 + \dots + \beta_m^2 > 0$$

и

$$\sum_{k=1}^m \beta_k \Phi_k^* = 0.$$

Но в силу (4.11) векторы Φ_k^* и Ψ_k связаны соотношением

$$\Omega \Phi_k^* = \Psi_k,$$

где Ω — матрица Грама, порожденная векторами (4.8) в евклидовом пространстве E . Поэтому для тех же постоянных $\beta_1, \beta_2, \dots, \beta_m$ получаем

$$\sum_{k=1}^m \beta_k \Psi_k = \sum_{k=1}^m \beta_k \Omega \Phi_k^* = \Omega \left(\sum_{k=1}^m \beta_k \Phi_k^* \right) = 0,$$

что невозможно, поскольку по сделанному выше предположению векторы $\Psi_1, \Psi_2, \dots, \Psi_m$ линейно независимы.

Итак, из линейной независимости векторов (4.34) следует линейная независимость векторов (4.35).

Далее сделаем следующее предположение. Пусть система функций (4.7) и линейный оператор L таковы, что

$$L\omega_1(t), L\omega_2(t), \dots, L\omega_n(t) \quad (4.36)$$

образуют в пространстве H систему линейно независимых элементов. При этом предположении и при условии линейной независимости векторов (4.34) элементы $\Phi_1, \Phi_2, \dots, \Phi_m$, определяемые равенствами (4.20), будут линейно независимыми.

Предположим противное. Тогда найдутся такие постоянные

$$\gamma_1, \gamma_2, \dots, \gamma_m, \text{ что } \gamma_1^2 + \gamma_2^2 + \dots + \gamma_m^2 > 0 \text{ и } \sum_{k=1}^m \gamma_k \Phi_k = 0.$$

Но тогда в силу равенства (4.33) получаем

$$\sum_{k=1}^m \gamma_k \Phi_k = \sum_{j=1}^n \left(\sum_{k=1}^m \gamma_k \Phi_k^* \right) L\omega_j(t) = 0. \quad (4.37)$$

А так как по сделанному выше предположению элементы (4.36) линейно независимы, то равенство (4.37) возможно лишь при условии, что

$$\sum_{k=1}^m \gamma_k \Phi_{kj}^* = 0, \quad j = 1, 2, \dots, n,$$

или, в силу (4.35),

$$\sum_{k=1}^m \gamma_k \Phi_k^* = 0,$$

что невозможно, поскольку по предположению векторы (4.35) линейно независимы. Следовательно, элементы $\Phi_1, \Phi_2, \dots, \Phi_m$ линейно независимы, что и обеспечивает существование матрицы $V = \Phi^{-1}$.

Таким образом, суммируя сказанное, сформулируем условия, при которых существует матрица $V = \Phi^{-1}$, т. е. условия, при которых возможно определение приближенных значений c_1^*, \dots, c_m^* неизвестных параметров c_1, \dots, c_m :

\mathcal{L}_1) должно выполняться неравенство $m \leq n$;

\mathcal{L}_2) функции $\omega_1(t), \omega_2(t), \dots, \omega_n(t)$ и функции $\varphi_1(t), \varphi_2(t), \dots, \varphi_m(t)$ должны быть такими, чтобы векторы $\Psi_1, \Psi_2, \dots, \Psi_m$, определяемые равенствами (4.34), были линейно независимыми между собой;

\mathcal{L}_3) функции $\omega_1(t), \omega_2(t), \dots, \omega_n(t)$ и линейный оператор L должны быть такими, чтобы определяемые ими элементы (4.36) в пространстве H были линейно независимыми.

Выше мы исходили из предположения, что неучитываемая доля $\kappa(t)$ фактической систематической погрешности $\eta(t)$ является малой величиной, и пренебрегли ею. В действительности же $\kappa(t)$ не равна нулю и оказывает определенное влияние на точность описанного выше способа коррекции систематической погрешности.

Действительно, более точная, чем (4.4), структура измерений имеет вид

$$y(t_i) = x(t_i) + \xi_i + \sum_{k=1}^m c_k \Phi_k(t_i) + \kappa(t_i), \quad i = 1, 2, \dots, N.$$

Произведя сглаживание этих измерений так, как это описано выше, получим сглаженную функцию $y^*(t)$, которая будет иметь следующую структуру:

$$y^*(t) = x^*(t) + \xi^*(t) + \sum_{k=1}^m c_k \Phi_k^*(t) - r_\kappa(t),$$

где функции $x^*(t), \xi^*(t), \Phi_k^*(t)$ определены выше, а

$$r_\kappa(t) = - \sum_{i=1}^n \kappa_i^* \omega_i(t) = - \kappa^*(t),$$

$$\kappa_i^* = \sum_{j=1}^n W_{ij}(\bar{\omega}_j, \bar{\kappa})_E.$$

Здесь $\bar{\kappa} = (\kappa(t_1), \dots, \kappa(t_N))$; $r_\kappa(t)$ — остаточная погрешность в сглаживающей функции $y^*(t)$, которая возникает в результате замены (ап-

проксимации) реальной систематической погрешности $\eta(t)$ функцией $\zeta(t)$ вида (4.2).

Далее, при определении приближенных значений c_k^* величин c_k поступим точно так же, как и ранее. Получим

$$x(t) = y^*(t) - \sum_{k=1}^m c_k \Phi_k^*(t) + l(t),$$

где

$$l(t) = r(t) + r_{\kappa}(t) = r_x(t) + r_{\xi}(t) + r_{\kappa}(t).$$

Выражение для c_k^* будет таким же, как и в (4.23). Однако погрешность $c_k - c_k^* = r_k$ на этот раз имеет вид

$$r_k = \sum_{j=1}^m V_{kj}(\Phi_j, Ll(t))_H = \delta_{xk} + \delta_{\xi k} + \delta_{\kappa k}, \quad (4.38)$$

где

$$\delta_{xk} = \sum_{j=1}^m V_{kj}(\Phi_j, Lr_x(t))_H,$$

$$\delta_{\xi k} = \sum_{j=1}^m V_{kj}(\Phi_j, Lr_{\xi}(t))_H,$$

$$\delta_{\kappa k} = \sum_{j=1}^m V_{kj}(\Phi_j, Lr_{\kappa}(t))_H.$$

Здесь δ_{xk} — доля в погрешности $c_k - c_k^*$, порожденная погрешностью аппроксимации $r_x(t) = x(t) - x^*(t)$; $\delta_{\xi k}$ — доля, порожденная погрешностью $r_{\xi}(t) = 0 - \xi^*(t)$; $\delta_{\kappa k}$ — доля, порожденная погрешностью $\kappa(t) = \eta(t) - \zeta(t)$. Таким образом, в рассматриваемом случае погрешность r_k зависит от $r_x(t)$, $r_{\xi}(t)$ и $\kappa(t)$ линейно.

Скорректированные измерения и скорректированная сглаживающая функция теперь имеют соответственно структуру

$$\hat{y}(t_i) = x(t_i) + \xi_i + \kappa(t_i) + \rho(t_i),$$

$$\hat{y}^*(t) = x^*(t) + \xi^*(t) + \kappa^*(t) + \rho^*(t),$$

где $\rho(t) = \sum_{k=1}^m r_k \Phi_k(t)$, $\rho^*(t) = \sum_{k=1}^m r_k \Phi_k^*(t)$, т. е. $\rho(t)$ и $\rho^*(t)$ определяются здесь такими же выражениями, как и в (4.29) и (4.32), с той, однако, разницей, что теперь $r_k = \delta_{xk} + \delta_{\xi k} + \delta_{\kappa k}$, а в (4.29) и (4.32) было $r_k = \delta_{xk} + \delta_{\xi k}$.

Таким образом, при корректировке измерений учитываемая доля $\zeta(t)$ реальной систематической погрешности $\eta(t)$ полностью устраняется. При этом вместо нее возникает погрешность $\rho(t)$, порожденная погрешностями $r_x(t)$, $r_{\xi}(t)$ и $r_{\kappa}(t)$, и, следовательно, остаточная систематическая погрешность в скорректированных измерениях представляет собой алгебраическую сумму $\kappa(t_i) + \rho(t_i)$.

Аналогично при корректировке сглаживающей функции $y^*(t)$ учитываемая доля $\zeta^*(t)$ систематической погрешности $\eta^*(t) = \zeta^*(t) +$

$+ \kappa^*(t)$, содержащейся в $y^*(t)$, полностью устраняется. При этом вместо нее возникает погрешность $\rho^*(t)$, порожденная погрешностями $r_x(t)$, $r_{\xi}(t)$ и $r_{\kappa}(t)$, и, следовательно, остаточная систематическая погрешность, содержащаяся в новой сглаживающей функции $\tilde{y}^*(t)$, представляет собой алгебраическую сумму $\kappa^*(t) + \rho^*(t)$.

Отсюда следует, что повысить точность коррекции систематической погрешности можно путем пропорционального уменьшения всех трех погрешностей $r_k(t)$, $r_{\xi}(t)$ и $\kappa(t)$ (а тем самым и $r_{\kappa}(t)$).

Источником погрешности при вычислении величин c_k^* в рассматриваемом случае могут явиться также погрешности в задании элемента h и оператора L . Пусть наблюдаемая функция $x(t)$ удовлетворяет условию (4.16). Предположим, что при практическом задании элемента h вместо него задан элемент $\tilde{h} = h + \delta h$ с погрешностью $\delta h \in H$. Тогда погрешность r_k в (4.38) примет вид

$$r_k = \delta_{xk} + \delta_{\xi k} + \delta_{\kappa k} + \delta_{hk},$$

где

$$\delta_{hk} = - \sum_{j=1}^m V_{kj}(\Phi_j, \delta h)_H.$$

Таким образом, погрешности $r_k = c_k - c_k^*$, $k = 1, 2, \dots, m$, зависят от погрешности δh также линейно.

Гораздо сложнее характер зависимости погрешности r_k от погрешности в задании оператора L . Пусть при практическом задании оператора L вместо него задан оператор $\tilde{L} = L + \delta L$. Тогда в силу (4.19) и (4.20)

$$\tilde{F} = \tilde{L}y^*(t) - h = Ly^*(t) - h + \delta Ly^*(t) = F + \delta F,$$

$$\tilde{\Phi}_k = \tilde{L}\varphi_k^*(t) = L\varphi_k^*(t) + \delta L\varphi_k^*(t) = \Phi_k + \delta\Phi_k.$$

В результате приближенной будет матрица $\tilde{\Phi}$ системы (4.22), а следовательно, и ее обратная матрица \tilde{V} , причем погрешность в этой обратной матрице будет нелинейно зависеть от погрешности δL . Следовательно, и погрешность $r_k = c_k - c_k^*$ будет нелинейно (причем весьма сложно) зависеть от погрешности δL .

Замечание 1. Если систематическая погрешность $\eta(t)$ содержит в себе такую аддитивную составляющую, скажем, $\gamma(t)$, которая порождает вектор $\bar{\gamma} = (\gamma(t_1), \gamma(t_2), \dots, \gamma(t_N))$, ортогональный ко всем векторам (4.8), т. е. если

$$(\bar{\omega}_i, \bar{\gamma})_E = 0, \quad i = 1, 2, \dots, n,$$

то

$$\gamma_i^* = \sum_{j=1}^n W_{ij}(\bar{\omega}_j, \bar{\gamma})_E = 0, \quad i = 1, 2, \dots, n,$$

$$\gamma^*(t) = \sum_{i=1}^n \gamma_i^* \omega_i(t) \equiv 0.$$

Следовательно, в этом случае наличие в измерениях систематической погрешности указанного типа $\gamma(t)$ никак не влияет на сглаживающую функцию $y^*(t)$. Это означает, что при сглаживании измерений по методу наименьших квадратов с помощью линейных форм, образованных из функций (4.7), те составляющие систематической погрешности, которые ортогональны в указанном выше смысле векторам (4.8), не влияют на точность сглаживающей функции $y^*(t)$.

Замечание 2. Как нетрудно проверить, изложенный метод коррекции легко распространяется на случай, когда вместо процесса $x(t)$ наблюдается некоторый другой процесс, линейно связанный с $x(t)$. Случай нелинейных связей будет рассмотрен ниже.

Многомерный процесс. Прямые измерения. Здесь предполагается, что наблюдаемый в течение некоторого интервала времени $[\alpha, \beta]$ физический процесс описывается вектор-функцией $X(t) = (x_1(t), x_2(t), \dots, x_v(t))$, которая является неизвестной. Предполагается также, что в дискретные моменты времени

$$t_{i1}, t_{i2}, \dots, t_{iN_i} \quad (4.39)$$

из $[\alpha, \beta]$ измеряются значения i -й компоненты $x_i(t)$ вектор-функции $X(t)$ ($i = 1, 2, \dots, v$). Индекс i в (4.39) обозначает, что значения компоненты $x_i(t)$ измеряются не обязательно синхронно с другими компонентами.

Предполагается, что погрешность, содержащаяся в результате измерения компоненты $x_i(t)$ в момент времени t_{ij} , представляет собой сумму двух слагаемых $\xi_{ij} + \eta_{ij}$. При этом погрешности ξ_{ij} ($i = 1, 2, \dots, v; j = 1, 2, \dots, N_i$) считаются случайными нормально распределенными величинами, каждая из которых имеет нулевое математическое ожидание и заданную дисперсию $M\{\xi_{ij}^2\} = \sigma_{ij}^2$, причем $M\{\xi_{ip}, \xi_{kq}\} = 0$, если выполнено хотя бы одно из двух условий $i \neq k$ и $p \neq q$. Для каждого $i = 1, 2, \dots, v$ случайные величины $\xi_{i1}, \dots, \xi_{iN_i}$ рассматриваются как значения некоторого гауссовского случайного процесса в моменты времени (4.39).

Относительно величин $\eta_{i1}, \dots, \eta_{iN_i}$ предполагается, как и в одномерном случае, что они представляют собой так называемую систематическую погрешность измерений и являются значениями некоторой функции $\eta_i(t)$ в моменты времени (4.39). При этом $\eta_i(t)$ может быть как случайной, так и неизвестной детерминированной функцией. Считаем, что функция $\eta_i(t)$ аппроксимируется выражением

$$\zeta_i(t) = \sum_{k=1}^{m_i} c_{ik} \varphi_{ik}(t), \quad (4.40)$$

где коэффициенты c_{ik} — неизвестные величины, а

$$\varphi_{i1}(t), \varphi_{i2}(t), \dots, \varphi_{im_i}(t) \quad (4.41)$$

являются линейно независимыми функциями, заданными на отрезке $[\alpha, \beta]$. В дальнейшем будут сформулированы дополнительные условия, которым должны удовлетворять эти функции. Через x_i обозначим

погрешность

$$\kappa_i(t) = \eta_i(t) - \xi_i(t), \quad (4.42)$$

возникающую при аппроксимации $\eta_i(t)$ выражением (4.40). Ясно, что эта погрешность зависит от того, какими выбраны функции (4.41), от их числа m_i и коэффициентов $c_{i1}, c_{i2}, \dots, c_{im_i}$.

Таким образом, результаты измерений компоненты $x_i(t)$ (они обозначены $y_i(t_{ij})$) будут иметь следующую структуру:

$$y_i(t_{ij}) = x_i(t_{ij}) + \xi_{ij} + \sum_{k=1}^{m_i} c_{ik} \varphi_{ik}(t_{ij}) + \kappa_i(t_{ij}). \quad (4.43)$$

По аналогии с предыдущим для каждого $i = 1, 2, \dots, v$ введем в рассмотрение N_i -мерное евклидово пространство E_i со скалярным произведением

$$(\bar{u}_i, \bar{V}_i)_{E_i} = \sum_{p=1}^{N_i} \frac{u_{ip} V_{ip}}{\sigma_{ip}^2}$$

его элементов $\bar{u}_i = (u_{i1}, \dots, u_{iN_i})$ и $\bar{V}_i = (V_{i1}, \dots, V_{iN_i})$. Тогда $\|\bar{u}_i\|_{E_i} = \sqrt{(\bar{u}_i, \bar{u}_i)_{E_i}}, \bar{u}_i \in E_i$.

Условимся произвольной вещественной скалярной функции $u(t)$, определенной на отрезке $[\alpha, \beta]$, ставить в пространстве E_i в соответствие вектор

$$\bar{u}_i = (u(t_{i1}), u(t_{i2}), \dots, u(t_{iN_i})).$$

Тогда измерения (4.43) можно записать в векторной форме

$$\bar{y}_i = \bar{x}_i + \bar{\xi}_i + \sum_{k=1}^{m_i} c_{ik} \bar{\varphi}_{ik} + \bar{\kappa}_i. \quad (4.43a)$$

Прежде чем ставить и решать задачу определения приближенных значений c_{ik}^* неизвестных параметров — коэффициентов c_{ik} , т. е. задачу приближенного определения систематических погрешностей, измерения (4.43) необходимо сгладить. Сглаживать их будем покомпонентно, т. е. отдельно по каждой компоненте $x_i(t)$, причем сглаживать так же, как это делалось в одномерном случае. Особенностью здесь может быть только то, что для каждого $i = 1, 2, \dots, v$ может выбираться своя система функций

$$\omega_{i1}(t), \omega_{i2}(t), \dots, \omega_{in_i}(t), \quad (4.44)$$

линейно независимых на множестве точек (4.39). Этой системе функций в пространстве E_i ставится в соответствие система линейно независимых векторов

$$\bar{\omega}_{i1}, \bar{\omega}_{i2}, \dots, \bar{\omega}_{in_i}. \quad (4.44a)$$

Произведя сглаживание, найдем

$$y_i^*(t) = \sum_{j=1}^{n_i} a_{ij}^* \omega_{ij}(t),$$

где

$$a_{ij}^* = \sum_{p=1}^{n_i} W_{ijp} (\bar{\omega}_{ip}, \bar{y}_i)_{E_i},$$

W_{ijp} ($j, p = 1, 2, \dots, n_i$) — элементы матрицы W_i , обратной к матрице Грама Ω_i , порожденной векторами (4.44а).

Как и раньше, устанавливаем, что сглаживающая функция $y_i^*(t)$ фактически имеет структуру

$$y_i^*(t) = x_i^*(t) + \xi_i^*(t) + \zeta_i^*(t) + \bar{\kappa}_i^*(t),$$

где

$$x_i^*(t) = \sum_{j=1}^{n_i} x_{ij}^* \omega_{ij}(t), \quad x_{ij}^* = \sum_{p=1}^{n_i} W_{ijp} (\bar{\omega}_{ip}, \bar{x}_i)_{E_i},$$

$$\xi_i^*(t) = \sum_{j=1}^{n_i} \xi_{ij}^* \omega_{ij}(t), \quad \xi_{ij}^* = \sum_{p=1}^{n_i} W_{ijp} (\bar{\omega}_{ip}, \bar{\xi}_i)_{E_i},$$

$$\zeta_i^*(t) = \sum_{k=1}^{m_i} c_{ik} \varphi_{ijk}(t),$$

$$\varphi_{ik}^*(t) = \sum_{j=1}^{n_i} \varphi_{ijk}^* \omega_{ij}(t), \quad \varphi_{ijk}^* = \sum_{p=1}^{n_i} W_{ijp} (\bar{\omega}_{ip}, \bar{\varphi}_{ik})_{E_i}.$$

Здесь $x_i^*(t)$ — наилучшая аппроксимация идеальной измеряемой компоненты $x_i(t)$ в смысле критерия

$$\|\bar{x}_i^* - x_i\|_{E_i}^2 = \min_{\alpha_{i1}, \dots, \alpha_{in_i}} \left\| \sum_{j=1}^{n_i} \alpha_{ij}^* \bar{\omega}_{ij} - \bar{x}_i \right\|_{E_i}^2; \quad (4.45)$$

$\xi_i^*(t)$ — статистическая оценка математического ожидания случайного процесса $\xi_i(t)$, которое в данном случае равно нулю (иными словами, $\xi_i^*(t)$ — наилучшая аппроксимация конкретной реализации случайного процесса $\xi_i(t)$ в смысле такого же критерия, как и (4.45)); $\zeta_i^*(t)$ — среднеквадратическая аппроксимация (по такому же критерию) учитываемой доли $\zeta_i(t)$ систематической погрешности $\eta_i(t)$; $\varphi_{ik}^*(t)$ — среднеквадратическая аппроксимация функции $\varphi_{ik}(t)$; $\bar{\kappa}_i^*(t)$ — среднеквадратическая аппроксимация неучитываемой доли $\kappa_i(t)$ систематической погрешности $\eta_i(t)$.

Замечание. По аналогии с одномерным случаем подчеркнем, что поскольку в (4.43а) векторы \bar{x}_i , $\bar{\xi}_i$, $\bar{\kappa}_i$ и постоянные c_{ik} нам не известны, то функции $x_i^*(t)$, $\xi_i^*(t)$, $\zeta_i^*(t)$ и $\bar{\kappa}_i^*(t)$ определить, конечно, нельзя. Они здесь используются только для качественного анализа структуры сглаживающих функций и не входят ни в какие практические методы и алгоритмы.

Введем в рассмотрение погрешности

$$r_{ix}(t) = x_i(t) - x_i^*(t),$$

$$r_{i\xi}(t) = 0 - \xi_i^*(t).$$

Тогда

$$x_i(t) - y_i^*(t) = r_{ix}(t) + r_{i\xi}(t) - \sum_{k=1}^{m_i} c_{ik} \Phi_{ik}^*(t) - \kappa_i^*(t).$$

Отсюда получаем

$$x_i(t) = y_i^*(t) - \sum_{k=1}^{m_i} c_{ik} \Phi_{ik}^*(t) + r_{ix}(t) + r_{i\xi}(t) - \kappa_i^*(t). \quad (4.46)$$

Как и в одномерном случае, возникает проблема приближенного определения неизвестных коэффициентов c_{ik} в выражении (4.40) для систематических погрешностей. Для решения этой проблемы предположим, что вектор-функция $X(t)$, описывающая наблюдаемый процесс, является элементом некоторого векторного линейного функционального пространства \mathfrak{X}_v . На множестве D_v определен линейный оператор L , переводящий произвольную вектор-функцию $U(t) = (u_1(t), \dots, u_v(t))$ из D_v в элемент гильбертового пространства H со скалярным произведением $(g, h)_H$ и нормой $\|h\|_H = \sqrt{(h, h)_H}$ его элементов h, g , причем имеет место равенство

$$LU = \sum_{i=1}^v L_i u_i(t), \quad (4.47)$$

где линейный оператор L_i действует на i -ю компоненту $u_i(t)$ вектор-функции $U(t)$, переводя ее в соответствующий элемент пространства H .

Пусть во множестве $L(D_v) \in H$ задан элемент h и пусть известно, что вектор-функция $X(t)$ удовлетворяет условию

$$LX = h, \quad (4.48)$$

причем предполагается, что это условие не позволяет однозначно определить $X(t)$ (в противном случае, как уже было отмечено выше, задача статистической оценки компонент вектор-функции $X(t)$ теряет смысл).

Подставляя в равенство (4.48) вектор-функцию $X(t)$, компоненты которой имеют вид (4.46), и учитывая выражение (4.47), получаем тождество

$$\sum_{i=1}^v L_i \left(y_i^*(t) - \sum_{k=1}^{m_i} c_{ik} \Phi_{ik}^*(t) + r_{ix}(t) + r_{i\xi}(t) - \kappa_i^*(t) \right) - h = 0.$$

Считая, далее, погрешности $r_{ix}(t)$, $r_{i\xi}(t)$ и $\kappa_i^*(t)$ малыми и пренебрегая ими, получаем «невязку» в пространстве H

$$\Delta = F - \sum_{i=1}^v \sum_{k=1}^{m_i} c_{ik} \Phi_{ik},$$

где

$$F = \sum_{i=1}^v L_i y_i^*(t) - h, \quad (4.49)$$

$$\Phi_{ik} = L_i \Phi_{ik}^*(t), \quad k = 1, 2, \dots, m_i, \quad i = 1, 2, \dots, v. \quad (4.50)$$

Эта невязка является функцией от неизвестных параметров (F и Φ_{ik} — известные элементы в H). В качестве их приближенных значений возьмем такие значения \dot{c}_{ik} , которые минимизируют квадрат нормы невязки: $\|\Delta\|_H^2 = \min$.

В результате для определения \dot{c}_{ik}^* получим линейную систему

$$\Phi C = f, \quad (4.51)$$

в которой Φ — матрица Грама, порожденная элементами Φ_{ik} (т. е. квадратная матрица, элементами которой являются скалярные произведения $(\Phi_{ik}, \Phi_{pq})_H$, $i, p = 1, 2, \dots, v$; $k = 1, 2, \dots, m_i$; $q = 1, 2, \dots, m_p$); C — вектор-столбец неизвестных параметров c_{ik} ; f — вектор-столбец с компонентами $(\Phi_{ik}, F)_H$, $k = 1, 2, \dots, m_i$; $i = 1, 2, \dots, v$.

Пусть матрица Φ имеет обратную матрицу $V = \Phi^{-1}$ с элементами V_{ikpq} (элемент V_{ikpq} стоит в матрице V на таком же месте, что и элемент $(\Phi_{ik}, \Phi_{pq})_H$ в матрице Φ). Тогда

$$C^* = Vf$$

или в компонентах

$$\dot{c}_{ik} = \sum_{p=1}^v \sum_{q=1}^{m_p} V_{ikpq} (\Phi_{pq}, F)_H; \quad k = 1, 2, \dots, m_i, \quad i = 1, 2, \dots, v.$$

Теперь можно найти приближенное значение

$$\tilde{\xi}_i(t) = \sum_{k=1}^{m_i} \dot{c}_{ik} \Phi_{ik}(t), \quad i = 1, 2, \dots, v, \quad (4.52)$$

для систематической погрешности $\xi_i(t)$ и приближенное значение

$$\tilde{\xi}_i^*(t) = \sum_{k=1}^{m_i} \dot{c}_{ik}^* \Phi_{ik}^*(t), \quad i = 1, 2, \dots, v, \quad (4.53)$$

для функции $\xi_i^*(t)$, входящей в состав сглаживающей функции $y_i^*(t)$, а затем внести поправки либо в измерения (4.43):

$$\tilde{y}_i(t_{ij}) = y_i(t_{ij}) - \tilde{\xi}_i(t_{ij}), \quad j = 1, 2, \dots, N_i; \quad i = 1, 2, \dots, v, \quad (4.54)$$

либо в сглаживающие функции $y_i^*(t)$:

$$\tilde{y}_i^*(t) = y_i^*(t) - \tilde{\xi}_i^*(t), \quad i = 1, 2, \dots, v. \quad (4.55)$$

Выясним теперь, как и в одномерном случае, структуру скорректированных измерений и скорректированных сглаживающих функций. С этой целью заметим прежде всего, что

$$\dot{c}_{ik}^* = c_{ik} - \rho_{ik},$$

где

$$\begin{aligned} \rho_{ik} &= \sum_{p=1}^v \sum_{q=1}^{m_p} V_{ikpq} (\Phi_{pq}, \tau)_H; \\ \tau &= \sum_{j=1}^v L_j \tau_j(t) \in H; \\ \tau_j(t) &= r_{jx}(t) + r_{j\xi}(t) - \dot{x}_j^*(t). \end{aligned} \quad (4.56)$$

Величина ρ_{ik} представляет собой погрешность, возникающую при замене параметра c_{ik} его приближенным значением c_{ik}^* . Эта величина порождена, как видно из (4.56), погрешностями $r_{jx}(t)$, $r_{j\xi}(t)$ и $\kappa_j^*(t)$, о которых речь уже шла выше, причем, как легко видеть,

$$\rho_{ik} = \delta_{xik} + \delta_{\xi ik} + \delta_{\kappa ik}, \quad (4.57)$$

где

$$\begin{aligned} \delta_{xik} &= \sum_{p=1}^v \sum_{q=1}^{m_p} V_{ikpq} \left(\Phi_{pq}, \sum_{j=1}^v L_j r_{jx}(t) \right)_H; \\ \delta_{\xi ik} &= \sum_{p=1}^v \sum_{q=1}^{m_p} V_{ikpq} \left(\Phi_{pq}, \sum_{j=1}^v L_j r_{j\xi}(t) \right)_H; \\ \delta_{\kappa ik} &= - \sum_{p=1}^v \sum_{q=1}^{m_p} V_{ikpq} \left(\Phi_{pq}, \sum_{j=1}^v L_j \kappa_j^*(t) \right)_H. \end{aligned}$$

Следовательно, ρ_{ik} — сумма трех погрешностей, каждая из которых линейно зависит от своей образующей погрешности.

Таким образом, найденные в (4.52) приближенные значения погрешностей $\zeta_i(t)$ имеют вид

$$\tilde{\zeta}_i(t) = \zeta_i(t) - \rho_i(t), \quad \text{где } \rho_i(t) = \sum_{k=1}^{m_i} \rho_{ik} \Phi_{ik}(t).$$

Скорректированные измерения (4.54) имеют фактически структуру

$$\hat{y}_i(t_{ij}) = x_i(t_{ij}) + \xi_{ij} + \rho_i(t_{ij}) + \kappa_i(t_{ij}),$$

т. е. учитываемая доля $\zeta_i(t)$ систематической погрешности $\eta_i(t)$ при описанной корректировке измерений полностью исключается, зато вместо нее появляется методическая погрешность $\rho_i(t)$.

Аналогично найденные в (4.53) приближенные значения для функций $\zeta_i^*(t)$ имеют вид

$$\tilde{\zeta}_i^*(t) = \zeta_i^*(t) - \rho_i^*(t),$$

где

$$\rho_i^*(t) = \sum_{k=1}^{m_i} \rho_{ik} \Phi_{ik}^*(t).$$

Следовательно, скорректированные сглаживающие функции (4.55) имеют фактически структуру

$$\tilde{y}_i^*(t) = x_i^*(t) + \xi_i^*(t) + \rho_i^*(t) + \kappa_i^*(t), \quad i = 1, 2, \dots, v,$$

т. е. погрешности $\zeta_i^*(t)$ при корректировке сглаживающих функций $y_i^*(t)$ исключаются полностью, зато вместо них в новых сглаживающих функциях $\tilde{y}_i^*(t)$ появляются методические погрешности $\rho_i^*(t)$.

Замечание. Как и в одномерном случае, устанавливается, что источником погрешности при вычислении величин c_{ik}^* могут явиться так-

же погрешности δh в задании элемента h и погрешность δL в задании оператора L в условии (4.48). При этом у ρ_{ik} в (4.57) появится четвертое слагаемое

$$\delta_{h_{ik}} = - \sum_{p=1}^v \sum_{q=1}^{m_p} V_{ikpq} (\Phi_{pq}, \delta h)_H.$$

От δL погрешность ρ_{ik} будет зависеть весьма сложно.

Выясним теперь условия, при которых возможна коррекция систематических погрешностей в рассматриваемом многомерном случае, т. е. возможно определение величин c_{ik}^* . Иными словами, выясним условия, при которых решение системы (4.51) существует и единственно, т. е. при которых существует матрица $V = \Phi^{-1}$.

Поскольку матрица Φ является матрицей Грама, порожденной векторами Φ_{ik} вида (4.50), то для существования обратной матрицы Φ^{-1} необходима и достаточна линейная независимость этих элементов. Поэтому для обеспечения существования единственного решения системы (4.51) достаточно обеспечить линейную независимость элементов (4.50). С этой целью заметим прежде всего, что

$$\Phi_{ik} = \sum_{j=1}^{n_i} \varphi_{ijk}^* L_i \omega_{ij}(t) = \sum_{j=1}^{n_i} \left(\sum_{p=1}^{n_i} W_{ijp} (\bar{\omega}_{ip}, \bar{\varphi}_{ik})_{E_i} \right) L_i \omega_{ij}(t). \quad (4.58)$$

Покажем, что для линейной независимости элементов (4.50) необходимо, чтобы для каждого $i = 1, 2, \dots, v$ было $m_i \leq n_i$. В самом деле, если для некоторого $i \in \{1, 2, \dots, v\}$ будет $m_i > n_i$, то m_i -мерные векторы

$$\Psi_{ij} = \begin{bmatrix} (\bar{\omega}_{i1}, \bar{\varphi}_{ij})_{E_i} \\ (\bar{\omega}_{i2}, \bar{\varphi}_{ij})_{E_i} \\ \vdots \\ (\bar{\omega}_{in_i}, \bar{\varphi}_{ij})_{E_i} \end{bmatrix}, \quad j = 1, 2, \dots, m_i, \quad (4.59)$$

обязательно будут линейно зависимыми. Следовательно, в этом случае найдутся такие постоянные $\alpha_1, \alpha_2, \dots, \alpha_{m_i}$, не равные одновременно нулю, что

$$\sum_{k=1}^{m_i} \alpha_k (\bar{\omega}_{ij}, \bar{\varphi}_{ik})_{E_i} = 0, \quad j = 1, 2, \dots, n_i.$$

Поэтому для тех же $\alpha_1, \alpha_2, \dots, \alpha_{m_i}$ получаем

$$\sum_{j=1}^{n_i} \alpha_j \Phi_{ij} = \sum_{j=1}^{n_i} \left(\sum_{p=1}^{n_i} W_{ijp} \sum_{k=1}^{m_i} \alpha_k (\bar{\omega}_{ip}, \bar{\varphi}_{ik})_{E_i} \right) L_i \omega_{ij}(t) = 0.$$

Это означает, что элементы $\Phi_{i1}, \Phi_{i2}, \dots, \Phi_{im_i}$ линейно зависимы. Но тогда линейно зависимыми будут и все элементы (4.50).

Итак, пусть $m_i \leq n_i$ для каждого $i = 1, 2, \dots, v$. Возьмем теперь системы функций (4.41) и (4.44) такими, чтобы векторы

$$\Psi_{i1}, \Psi_{i2}, \dots, \Psi_{im_i}, \quad (4.60)$$

определенные равенством (4.59), были линейно независимыми. Такие системы существуют. В самом деле, считая систему (4.44) выбранной, в качестве функций (4.41) можно взять все (если $m_i = n_i$) или часть (если $m_i < n_i$) функций (4.44). В этом случае, как нетрудно видеть, векторы (4.60) будут линейно независимыми.

Покажем, что если векторы (4.60) линейно независимы, то линейно независимыми будут также n_i -мерные векторы

$$\Phi_{ij}^* = \begin{pmatrix} \Phi_{i1j}^* \\ \vdots \\ \Phi_{in_ij}^* \end{pmatrix}, \quad j = 1, 2, \dots, m_i, \quad (4.61)$$

компоненты которых определяются равенством

$$\Phi_{ijk}^* = \sum_{p=1}^{n_i} W_{ijp} (\bar{\omega}_{ip}, \bar{\Phi}_{ik})_{E_i}. \quad (4.62)$$

Чтобы доказать это, предположим противное. Тогда найдутся такие постоянные $\beta_1, \beta_2, \dots, \beta_{m_i}$, не равные одновременно нулю, что

$$\sum_{i=1}^{m_i} \beta_i \Phi_{ii}^* = 0.$$

Но поскольку в силу (4.59) и (4.62)

$$\Omega_i \Phi_{ii}^* = \Psi_{ii}, \quad i = 1, 2, \dots, m_i,$$

где Ω_i — матрица Грама, порожденная векторами (4.44а), то для тех же $\beta_1, \beta_2, \dots, \beta_{m_i}$ получим

$$\sum_{i=1}^{m_i} \beta_i \Psi_{ii} = \Omega_i \left(\sum_{i=1}^{m_i} \beta_i \Phi_{ii}^* \right) = 0,$$

что невозможно, так как векторы (4.60) по предположению линейно независимы.

Итак, из линейной независимости векторов (4.60) следует линейная независимость векторов (4.61).

Далее сделаем следующее предположение. Пусть системы функций (4.44) (при $i = 1, 2, \dots, v$) таковы, что совокупность элементов

$$L_i \omega_{i1}(t), L_i \omega_{i2}(t), \dots, L_i \omega_{in_i}(t), \quad i = 1, 2, \dots, v, \quad (4.63)$$

образует в пространстве H систему линейно независимых векторов. Тогда при $m_i \leq n_i$ и при условии линейной независимости векторов (4.60) (отдельно для каждого $i = 1, 2, \dots, v$) элементы (4.50) будут в совокупности линейно независимыми. Предположим противное. Тогда найдутся такие постоянные γ_{ik} , не равные одновременно нулю, что

$$\sum_{i=1}^v \sum_{k=1}^{m_i} \gamma_{ik} \Phi_{ik} = 0.$$

Но в силу (4.58) получаем

$$\sum_{i=1}^v \sum_{k=1}^{m_i} \gamma_{ik} \Phi_{ik} = \sum_{i=1}^v \sum_{j=1}^{n_i} \left(\sum_{k=1}^{m_i} \gamma_{ik} \Phi_{ijk}^* \right) L_i \omega_{ij}(t) = 0.$$

А так как по предположению элементы (4.63) линейно независимы, то это равенство возможно лишь при условии

$$\sum_{k=1}^{m_i} \gamma_{ik} \Phi_{ijk}^* = 0, \quad j = 1, 2, \dots, n_i; \quad i = 1, 2, \dots, v,$$

или, что эквивалентно, при условии

$$\sum_{k=1}^{m_i} \gamma_{ik} \Phi_{ik}^* = 0, \quad i = 1, 2, \dots, v.$$

Однако последнее равенство невозможно, потому что для каждого $i = 1, 2, \dots, v$, как показано выше, векторы $\Phi_{i1}, \dots, \Phi_{im_i}^*$ линейно независимы. Следовательно, элементы (4.50) линейно независимы, что и обеспечивает существование матрицы $V = \Phi^{-1}$.

Таким образом, суммируя сказанное, сформулируем условия, при которых возможно определение приближенных значений c_{ik}^* неизвестных параметров c_{ik} :

\mathcal{L}_1) для каждого $i = 1, 2, \dots, v$ должно выполняться неравенство $m_i \leq n_i$;

\mathcal{L}_2) для каждого фиксированного $i \in \{1, 2, \dots, v\}$ функции $\omega_{i1}(t), \omega_{i2}(t), \dots, \omega_{im_i}(t)$ и функции $\Phi_{i1}(t), \Phi_{i2}(t), \dots, \Phi_{im_i}(t)$ должны быть такими, чтобы векторы $\Psi_{i1}, \Psi_{i2}, \dots, \Psi_{im_i}$, определяемые равенствами (4.59), были линейно независимыми между собой;

\mathcal{L}_3) функции $\omega_{ij}(t), j = 1, 2, \dots, n_i; i = 1, 2, \dots, v$, и линейный оператор L , удовлетворяющий равенству (4.47), должны быть такими, чтобы определяемые ими элементы (4.63) в пространстве H были линейно независимыми.

Многомерный процесс. Косвенные измерения. Линейные связи. Здесь задача коррекции систематической погрешности измерений, рассмотренная выше, а также в работе [91], распространяется и на случай косвенных измерений. Предполагается, что в моменты (4.39) измеряются не компоненты вектор-функции $X(t)$, описывающей наблюдаемый физический процесс, а компоненты $q_i(t)$ некоторой другой (тоже неизвестной) вектор-функции $Q(t) = (q_1(t), \dots, q_v(t))$, связанной с $X(t)$ линейным соотношением

$$X^T(t) = \mathcal{A}Q^T(t),$$

где \mathcal{A} — заданная невырожденная матрица порядка $v \times v$.

Относительно структуры результатов измерений $y_i(t_{ij})$ сохраняются прежние предположения:

$$y_i(t_{ij}) = q_i(t_{ij}) + \xi_{ij} + \sum_{k=1}^{m_i} c_{ik} \Phi_{ik}(t_{ij}) + \kappa_i(t_{ij}).$$

Чтобы сделать возможным определение величин c_{ik}^* , снова предположим, что вектор-функция $X(t)$ удовлетворяет условию (4.48). Тогда нетрудно показать, что измеряемая вектор-функция $Q(t)$ удовлетворяет условию

$$\sum_{i=1}^{\nu} B_i q_i(t) = h,$$

или в векторной форме

$$BQ = h,$$

где $B = (B_1, \dots, B_\nu)$ — линейный оператор, действующий в пространстве H , причем каждая его компонента B_i есть линейный оператор вида

$$B_i = \sum_{j=1}^{\nu} L_j \mathcal{A}_{ij},$$

где \mathcal{A}_{ij} ($i, j = 1, 2, \dots, \nu$) — элементы матрицы \mathcal{A} . Тем самым задача определения приближенных значений параметров c_{ik} в случае косвенных измерений сводится к аналогичной задаче для случая прямых измерений.

МЕТОДЫ АНАЛИЗА ДАННЫХ ПРИ ИССЛЕДОВАНИИ РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА

§ 1. Вводные замечания

Обычно отдельную единицу наблюдения независимо от ее конкретной природы характеризуют термином «объект». Свойства объектов определяются с помощью некоторого набора признаков и их значений (измерений). Данные — это зафиксированные результаты измерения признаков, выражающие свойства рассматриваемого множества объектов.

Говорят о многомерных статистических методах, когда данные представляют собой совокупность результатов измерений, относящихся к некоторому объекту, и все эти измерения рассматриваются в статистическом анализе одновременно для целого ряда объектов и оцениваются совместно. При этом основной анализа должна быть математическая модель рассматриваемого явления, а данные привлекаются лишь для установления адекватности этой модели, оценки ее параметров и т. п. (критерий χ^2 , дисперсионный анализ, метод максимального правдоподобия, робастное оценивание и т. д.).

При статистическом оценивании в предположении определенной математической модели речь идет о том, чтобы по возможности точно определить истинные значения параметров генеральной совокупности, исходя из выборочных наблюдений. Роль исследователя должна сводиться к выбору или формированию модели явления и последующего проведения анализа на основе объективного математического рассуждения. Таким образом, выбор наилучшего из возможных методов статистической обработки производится в соответствии с некоторым критерием качества, основанным на допущениях о вероятностной природе исходных данных.

В анализе данных (АД) осуществлен другой («обратный») подход, согласно которому исследователь на основании исходных данных стремится сделать заключение о математической модели исследуемого явления. Как правило, данные прежде всего многомерны и их многомерность порождена сложностью исследуемого явления или слабой его изученностью. Обычно кажутся сложными те системы, для которых на данный момент не известны некоторые «внутренние» факторы, определяющие их поведение. Цель анализа состоит в выявлении гипотетических величин-факторов по большому числу (возможно, порядка сотен, тысяч) экспериментальных данных, а также в последующем образовании типологических группировок, т. е. в расширении теоретических представлений об изучаемом явлении на основании эмпирических

данных. Другая цель АД — создание средств, позволяющих охватывать содержание таблиц данных большого объема через представления (предпочтительнее визуальные), которые легко доступны для понимания пользователя. Эта особенность АД тесно связана с тем, что на каждом этапе метода полученные результаты подвергаются осмыслению и в итоге принимается решение продолжать реализацию наметенной программы либо вернуться назад на определенный шаг исходной программы анализа, произвести некоторую корректировку плана, возможно, изменить критерий качества обработки и с этого момента снова продолжить анализ. К примеру, собрав требуемые данные, по определенным правилам производим предварительный анализ. При этом можем сделать заключение о том, что исходную большую таблицу данных нет смысла обрабатывать целиком, поскольку из физических соображений ясно, что необходимо ее рассматривать как некоторый набор таблиц и, следовательно, последующий план обработки должен быть изменен.

На этапах компонентного анализа или автоматической классификации на основании полученного решения (а также в результате визуального анализа отображения исходной информации на факторных плоскостях) могут быть приняты решения об изменении условий и повторном проведении обработки либо об изменении содержания последующих шагов (например, использовать другую метрику, изменить число классов разбиения и т. п.). По мере продвижения вперед результаты становятся все более содержательными. На последнем этапе, когда принимается решение об окончании обработки, производится интерпретация результатов.

Развитие математического обеспечения, направленное на решение задач АД, позволяет широко использовать ЭВМ.

Хотя ЭВМ предоставляет большие возможности, для исследователя остается широкое поле деятельности как важного элемента этого процесса. Следовательно, систематизируем свою интуицию. На пути систематизации интуиции методы АД являются шагом вперед по сравнению со статистическими методами, способствуют возникновению гипотез.

Математические методы, развиваемые для реализации ориентированной на исследователя методологии, объединяются под общим названием методов АД.

§ 2. Этапы проведения АД.

Сбор и подготовка данных

Любой эксперимент включает следующие основные этапы:

- 1) сбор и подготовка данных; постановка задачи; выбор плана эксперимента; сбор и подготовка исходных данных для обработки;
- 2) обработка данных;
- 3) анализ (интерпретация) результатов и принятие решения.

Последовательность проведения эксперимента и АД полностью соответствуют этой схеме.

Представим последовательность и содержание этапов АД.

Схема этапов АД

Сбор и подготовка данных

Этап
I

1. Представление стандартизованных данных с помощью таблиц.
2. Предварительный анализ данных.
3. Выбор метрики. Выбор критериев качества обработки.
4. Расчет требуемых статистических характеристик данных (средних значений, среднеквадратических отклонений, построение гистограмм и т. п.).
5. Редактирование данных.
6. Прочие задачи, которые следует решить до «того», как применить «основной метод».

Обработка данных

Алгоритмы обработки и анализа

Визуальное представление результатов

Этап
II

А. Компонентный анализ

1. Факторный анализ соответствий (исходные данные представлены таблицей сопряженности).
2. Компонентный анализ (исходные данные представлены таблицей объект — объект: корреляционная матрица, матрица расстояний, матрица близостей).
3. Компонентный анализ (таблицы логических расстояний)
4. Факторный анализ.
5. Иллюстративное представление объектов и признаков в факторных плоскостях. Иллюстративное представление на фоне исходных данных дополнительных элементов.

Б. Автоматическая классификация

1. Иерархическая классификация с использованием: таблицы сопряженности, факторных переменных матрицы, расстояний (близостей).
2. Разбиение исходного множества на: оптимальное число классов; на заданное число классов.
3. Иллюстративное представление дополнительных элементов с нулевой массой (не влияющих на ход классификации)

Анализ результатов

Алгоритм интерпретации результатов

Этап
III

1. Иллюстративное представление результатов в виде точек в факторной плоскости, представление дендограмм, гистограмм, таблиц.
2. Построение эллипсоидов инерции выделенных классов в иерархической и простой классификации в факторной плоскости.
3. Определение взаимного вклада классов и факторов в общую инерцию. Интерпретация факторов.

Обсуждение результатов

Принятие решения о пересмотре позиций, принятых в I этапе.
Принятие решения о пересмотре позиций, принятых во II этапе.
Принятие решения об окончании анализа.

Традиционно исходная информация для проведения компонентного (факторного) анализа представляется в виде таблиц данных о связях или сходствах объектов исследуемой системы. Данные могут быть получены на основании экспертных опросов либо с помощью измерения или вычисления реальных связей (например, корреляционных) между объектами и признаками. При этом под связями подразуме-

меваются довольно широкий спектр отношений между объектами и признаками, например, данные годовых поставок из одной отрасли в другую в задачах межотраслевого баланса, привлекательность различных видов трудовой деятельности для различного контингента населения, данные о передвижении пассажиров из одного пункта в другой для определенного города и т. д.

Наиболее распространенными являются три основных вида таблиц (хотя также используются таблицы других видов) представления данных: таблицы сопряженности; таблицы объект — признак, таблицы связи между объектами (объект — объект, признак — признак).

Т а б л и ц а с о п р я ж е н н о с т и. Допустим, ставится задача исследовать два набора элементов (явлений, свойств и т. д.). Первый набор состоит из n элементов (количество строк), и каждому элементу поставлена в соответствие определенная строка таблицы. Второй набор состоит из p элементов (количество столбцов), каждому из которых поставлен в соответствие определенный столбец той же таблицы. Относительно этих наборов известно, что в результате испытания один из элементов первой группы появляется совместно с одним из элементов второй группы.

Очевидно, $\sum_i \sum_j k_{ij} = K$, т. е. мы получили таким образом таблицу сопряженности двух групп элементов. Таблица сопряженности обладает свойством двойственности относительно строк и столбцов в том смысле, что здесь строки можно считать объектами, а столбцы — признаками, но можно считать и наоборот: простое транспонирование матрицы не нарушает ее смысла.

Рассмотрим, например, две противоборствующие стороны (теория игр). Первая сторона имеет n стратегий, вторая — p . Каждый раз после выбора стратегии i ($i = 1, 2, \dots, n$) первой стороной и стратегии j ($j = 1, 2, \dots, p$) — второй в соответствующую клетку ij будем добавлять единицу. В результате получим таблицу сопряженности. В данном случае симметрия относительно строк и столбцов очевидна.

Дальнейший анализ этой таблицы (факторный анализ соответствий, автоматическая классификация) может быть направлен, например, на установление правил («разгадывание секретов»), которыми руководствуются стороны при выборе стратегий.

Если все ненулевые значения элементов ij заменим единицами, а нулевые оставим без изменения, то получим логическую (булевскую, индикаторную) таблицу сопряженности.

Т а б л и ц а о б ъ е к т — п р и з н а к. Наиболее общепотребительная интерпретация таблицы: строки соответствуют переменным (признакам). При этом каждый отдельный элемент на пересечении строки и столбца представляет собой значение признака j для соответствующего объекта i . Пусть набор объектов представляет собой набор диодов разных типов. В качестве признаков могут выступать: напряжение, ток прямой, ток обратный, мощность рассеивания, материал, из которого изготовлен диод, и т. п. В клетке ij таблицы нужно указать значение признака j (например, обратный ток) для i -го диода.

Ясно, что значения некоторых признаков могут принимать одинаковые значения для разных объектов.

Таблица объект — признак, как правило, не обладает свойством двойственности (в отличие от таблицы сопряженности). Кажется неестественным утверждение, например, что «объект — ток обладает признаком диод». Но главное не в этом. Дело в том, что значения признаков можно сравнивать по величине (или сопоставлять), тем самым сравнивая объекты по соответствующему признаку. В то же время сравнение элементов строки, т. е. значений разных признаков на одном и том же объекте, как правило, бессмысленно. Например, можно сравнивать величину обратного тока для разных диодов, а также сопоставлять содержательно материалы, из которых изготовлены диоды. Но сопоставление тока и материала, из которого изготовлен диод (если даже предположить, что это имеет какой-то смысл), совсем другое дело.

В общем случае формальная двойственность объектов и признаков требует, чтобы модели анализа признаков допускали переформулировку в терминах объектов, и, наоборот, модели анализа объектов допускали переформулировку в терминах признаков¹.

Другой пример: при изучении технико-экономических показателей деятельности различных предприятий данной отрасли строки таблицы соответствуют предприятиям, а столбцы — показателям их деятельности.

Т а б л и ц а о б ъ е к т — о б ъ е к т. Эта таблица описывает связи между объектами. Такое представление часто оказывается удобным. Анализ таблиц объект — признак приводит к рассмотрению взаимосвязи признаков. По мере углубления изучения объекта рассматривается все большее и большее число признаков. При этом закономерность расплывается зачастую на множество малозначащих связей. В этом случае можно свойства объекта выражать не в терминах признаков, а в терминах отношений между объектами по данному свойству. В частности, такого типа матрицы используются при парном сравнении (сопоставлении) группы объектов с привлечением экспертов. К рассматриваемому типу таблиц можно отнести корреляционную (ковариационную) матрицу, определенную для некоторой совокупности случайных переменных.

Рассмотрим n -мерные векторы-столбцы X_j ($j = 1, 2, \dots, p$). Каждый вектор имеет n составляющих, так что в результате имеем $n \times p$ -матрицу $X = [x_{ij}]$.

Геометрически эту матрицу можно понимать двояко. Можно определить n -мерное евклидово пространство в соответствии с размерностью X_j . Тогда выборка будет представлена p точками. В свою очередь, можно определить p -мерное пространство, а выборку представить n точками. В соответствии с этим можно образовать ковариационную матрицу размером $p \times p$ и интерпретировать ее как таблицу объект — объект или образовать матрицу с размерами $n \times n$ и интер-

¹ Для этого требуется, чтобы объекты и признаки обладали свойством однородности (сравнимости).

претировать ее как таблицу признаков — признаков. Причем ранг этих матриц удовлетворяет соотношению — $r \leq \min(n, p)$ ².

К рассмотренному типу таблиц можно свести большое разнообразие исходных данных. Существуют и другие типы «стандартных» таблиц, но на их рассмотрении останавливаться не будем. Скажем только, что алгоритмы и прикладные программы приспособлены, как правило, к работе с определенным типом таблицы.

Исторически метод АД имеет происхождение из математической психологии. В настоящее время он особенно интенсивно развивается в связи с широким применением в социально-экономических исследованиях. Не будет ошибкой утверждать, что АД также может быть широко использован при проведении технических экспериментов.

По существу необходимо различать методы проведения наблюдений и методы анализа. Под проведением наблюдений понимают вопросы планирования эксперимента, которые учитывают влияние метода проведения и отбора наблюдений на последующий анализ в целом.

При изучении принципов экспериментирования учитывается важность варьирования в процессе эксперимента предположительно причинных переменных, чтобы определить их влияние на рассматриваемую зависимую переменную. При этом различают две фазы исследования. На начальных этапах решения задач должны быть проведены функциональные эксперименты для выяснения принципиальной возможности достижения какого-либо выхода или другого желаемого свойства. После обнаружения некоторого множества рабочих или перспективных условий следует продолжить эксперимент, варьируя одновременно все причинные переменные. При этом следует иметь в виду, что существует опасность приписывания одному или нескольким экспериментальным причинным переменным влияние на зависимую переменную, которое на самом деле объясняется изменением некоторой причинной переменной, не учтенной в структуре эксперимента. Более того, может оказаться, что не все важные причинные переменные учтены структурой эксперимента, так как некоторые переменные могут быть просто не известны.

С другой стороны, известные переменные могут необоснованно рассматриваться как не имеющие существенного значения, и они отбрасываются.

Эти вопросы могут быть разрешены применением в процессе эксперимента рандомизации (предложенной Р. Фишером в 1935 г.). Напомним формулировку принципа рандомизации: «Если при проведении эксперимента проводится размещение экспериментальных объектов по комбинациям факторов, то это должно делаться случайным образом с использованием равных вероятностей».

Из этого следует, что принцип рандомизации в эксперименте приводит к тому, что данные эксперимента представляются таблицей объект — признак, где в строке дается результат отдельного эксперимен-

² Легко обнаружить связь между таблицами объект — признак, объект — объект и признак — признак. Так, например, если X — таблица объект — признак, то $A = XX^T$ — таблица объект — объект, а $B = X^T X$ — таблица признак — признак.

та. Количество строк соответствует количеству исследуемых комбинаций значений факторов.

Другой подход к организации эксперимента состоит в следующем.

С самого начала следует написать полный перечень всех переменных (признаков), которые могут оказывать влияние на желаемые свойства разрабатываемого изделия. Для этого составляется таблица всех потенциальных причинных переменных и всех представляющих интерес свойств изделия или процесса. Эту таблицу называют матрицей влияния. Формируется таблица следующим образом.

Для каждой причинной переменной, по предположению влияющей на какой-либо результат, отводится столбец, и для каждого отклика отводится строка. В каждую клетку такой таблицы экспериментатор заносит результат своих знаний или догадок. Так, например, в ячейку ij записывается единица (возможно, с учетом знака), если предполагается, что переменная j оказывает влияние на отклик (объект) i . Знак должен учитывать направление возможного влияния переменной. Можно также указать в клетке признак, свидетельствующий о том, что характер влияния переменной неизвестен, и т. п.

Таким образом, в матрице влияния отражено мнение экспериментатора на данный момент времени относительно эффекта воздействия каждой независимой управляемой переменной на каждое свойство изделия (процесса).

Рассмотрим некоторые аспекты, характеризующие технические (промышленные) эксперименты.

Подтверждающие эксперименты проводятся на действующих системах с целью изучения реакции системы (изучается одно ее свойство) на изменение вариантов рабочих условий. Результаты этих экспериментов могут быть представлены таблицей объект — признак.

Другой вариант, когда в результате эксперимента получена совокупность N наблюдений, охарактеризованная двумя системами признаков $X = (x_1, x_2, \dots, x_n)$ и $Y = (y_1, y_2, \dots, y_m)$. Первая содержит факторные, входные признаки, а вторая — результативные, выходные. В отличие от предыдущего варианта, когда эффекты воздействия некоторых факторов на какое-либо одно свойство изделия могут быть достаточно хорошо предсказаны, более реальным является случай, когда экспериментатора интересуют несколько свойств.

В этом случае анализ многомерных данных может быть проведен с помощью группировок. Сперва производятся оценка и анализ матриц коэффициентов связи между парами признаков из X и Y , затем осуществляется построение группировок по тем или иным совокупностям данных. Такой анализ может использоваться для получения теоретического знания, поскольку структура парных связей может дать некоторое представление о причинно-следственных отношениях.

Следующий аспект основан на различии между условиями проведения эксперимента, в которых данные собираются последовательно и параллельно. Данные эксперимента представляют собой идентичные наборы измерений, произведенные в различные моменты времени. Эти данные можно также представить таблицей типа объект — признак (хотя здесь не очень ясно, что объект, а что признак), разместив

в строках траектории во времени отдельных параметров; столбцы будут соответствовать набору измеренных параметров, отнесенных к определенному моменту времени.

В связи с этой ситуацией отметим, что возможен вариант, когда в определенный момент времени измеряется не вектор, а матрица. В этом случае в результате эксперимента приходим к необходимости анализировать не отдельную матрицу, а совокупность таблиц (блок). Естественно, что эта задача намного сложнее анализа двумерных таблиц. Тем не менее в настоящее время разработаны программы для ЭВМ, которые могут работать с блоками.

Наконец, отметим эксперименты, различающиеся по чувствительности. В некоторых случаях интересующий нас эффект значительно превосходит величину среднеквадратической ошибки. Тогда не требуется большого числа опытов. Это характерно для физических измерений (например, для траекторных). Иногда оказывается, что полезный эффект значительно меньше среднеквадратической ошибки измерений. Это характерно для социально-экономических процессов, клинических опытов, для крупных экспериментальных работ, где всегда присутствуют неуправляемые факторы и незначительные достижения имеют важное значение. Здесь можно облегчить положение за счет значительного увеличения объема данных. Эта ситуация как раз характерна для методов АД.

§ 3. Обработка данных, анализ результатов и принятие решений

В настоящее время существует значительное количество алгоритмов, которые можно было бы отнести к алгоритмам обработки данных. Но основой всех этих алгоритмов являются алгоритмы компонентного анализа (корреспондент — анализ), анализ главных компонент (факторный анализ), факторный анализ матрицы расстояний (корреляционной матрицы и др.) и автоматической классификации (алгоритмы разбиения и алгоритмы иерархической классификации, а также алгоритмы дискриминантного анализа). Эти алгоритмы могут использовать различную информацию: таблицу сопряженности, таблицу объект — признак, таблицу объект — объект, а также факторные координаты, полученные в результате работы одного из алгоритмов компонентного анализа.

Основной задачей компонентного анализа является уменьшение объема исходной таблицы за счет сокращения исходного числа признаков, что, в свою очередь, позволяет представить объекты в виде точек некоторого координатного пространства невысокой размерности. При этом должна быть сохранена исходная структура: связи или сходства объектов передаются расстоянием между точками. Тогда оси нового пространства могут интерпретироваться как новые признаки, лежащие в основе связи между объектами. Каждый объект в той или иной степени характеризуется этими признаками. Новый выбор признаков строится таким образом, что с помощью первого выделенного призна-

ка можно «максимально» воссоздать исходную таблицу и никакой другой признак не может сделать это лучше.

С помощью второго признака максимально воссоздается остаток, т. е. информация, не учтенная первым признаком, причем ось первого признака ортогональна оси второго признака. Третий признак воссоздает остаток, не учтенный первыми двумя, и его ось ортогональна первым двум и т. д.

Такая конструкция системы признаков, называемая факторами, позволяет существенно сократить исходное число признаков, сохранив приемлемую точность исходной таблицы. Следует заметить, что уже на этом шаге может решаться задача интерпретации выделенных признаков. В основном это задача факторного анализа. Но если не удастся достаточно хорошо интерпретировать систему признаков, с ее помощью можно решить ряд задач сопоставительного анализа и автоматической классификации. Если ограничиться двумя-тремя признаками, то исходные объекты (и исходные признаки) могут быть изображены в виде точки на плоскости, координатами которых являются факторы. Такие плоскости называют факторными с указанием номеров факторов, которые их определяют. Напомним, что с увеличением номера фактора его вес убывает.

По построению факторы являются линейной функцией исходных признаков. Поэтому они сохраняют информацию о взаимном расположении исходных объектов, что позволяет проводить их визуальное сопоставление. Степень близости объектов в факторной плоскости свидетельствует о схожести объектов, следовательно, появляются широкие возможности по обеспечению исходной информацией процесса проведения кластерного анализа, т. е. проведения автоматической классификации. В § 2 рассматривалась двойственность таблицы объект — признак. Если двойственность имеет место, то можно говорить о сопоставлении признаков, т. е. при отображении элементов (объектов и (или) признаков) на факторную плоскость схожие элементы будут располагаться достаточно близко друг от друга и, наоборот, близко расположенные отображения соответствуют сходным элементам. Попутно отметим, что если на факторную плоскость отображены одновременно и объекты и признаки, то в окрестности объекта будут группироваться наиболее характерные для него признаки. Причину этого будет показано ниже, хотя данный факт не очевиден, так как пространство объектов и пространство признаков — это разные пространства.

Автоматическая классификация (АК) объединяет весь набор методов и алгоритмов, предназначенных для разбиения совокупности объектов, определенных набором признаков на какое-то число однородных в определенном смысле групп (классов), причем эти группы могут быть в той или иной степени связаны между собой.

Задача о разбиении множества объектов на относительно далекие группы тесно связанных между собой объектов широко рассматривается в литературе. При этом часто используются термины: автоматическая классификация, кластерный анализ, таксономия, распознавание без учителя, диагонализация матриц связи и т. п. Среди них наиболее употребительными, по-видимому, являются термины автома-

тической классификации и кластерный анализ. В дальнейшем будем придерживаться этой терминологии.

Кластерным называют анализ неклассифицированных объектов, проводимый с целью выделения структур, классов, множеств подобных объектов. До сих пор нет точного его определения. Дело в том, что кластерный анализ используется и совершенствуется в довольно различных областях (социология, психология, биология, экономика, медицина, геология, астрономия и т. д.), а в различных прикладных задачах цели анализа могут быть различными. В сущности, этот анализ тесно связан с исследованиями процесса порождения данных.

Для разбиения на классы широко используется критерий качества, использующий следующие неформальные требования к классификации: внутри групп объекты должны быть тесно связаны между собой; необходимо, чтобы элементы разных групп располагались далеко друг от друга; при прочих равных условиях распределение объектов по группам должно быть равномерным.

Результаты АК представляют собой либо разбиение данных, либо иерархию разбиений. В ряде случаев разбиения могут быть представлены «деревьями» в соответствии с теорией графов, узлы которого являются объектами, подлежащими классификации. Более того, результаты могут быть представлены в виде частично перекрывающихся кластеров или компактных (плотных) множеств, внутри которых элементы неклассифицированы.

Различают несколько типов алгоритмов классификации:

- 1) агломеративные алгоритмы, в которых кластеры строятся путем попарного объединения объектов и которые обеспечивают иерархическое разбиение объектов (восходящая иерархия);
- 2) разделительные алгоритмы, которые осуществляют дихотомизацию всего множества объектов и которые могут обеспечить иерархическое разбиение объектов (нисходящая иерархия);
- 3) алгоритмы, которые непосредственно формируют разбиение, также, как методы кластеризации вокруг перемещающихся центров или алгоритмы минимальных отклонений.

Совершенно новые эффекты обнаружились при применении к данным методов формирования факторов и методов АК с тем, чтобы с каждым фактором была связана своя типология. Часто оказывается возможным интерпретировать полученные типологические группировки, и их можно понимать как систему терминов, позволяющих достаточно коротко описывать отдельные явления.

В любом случае использование компонентного анализа в сочетании с задачей классификации позволяет идентифицировать группы и определить их относительное положение. Часто разбиения или иерархические структуры используются для более качественного анализа результатов, полученных при предварительном анализе методом главных компонент.

Компонентный анализ. Пусть X — матрица, состоящая из n строк и p столбцов. Каждую полную матрицу можно интерпретировать двояко: либо как набор из n точек (или векторов) в p -мерном пространстве (R^p), либо как набор из p точек в n -мерном пространстве (R^n).

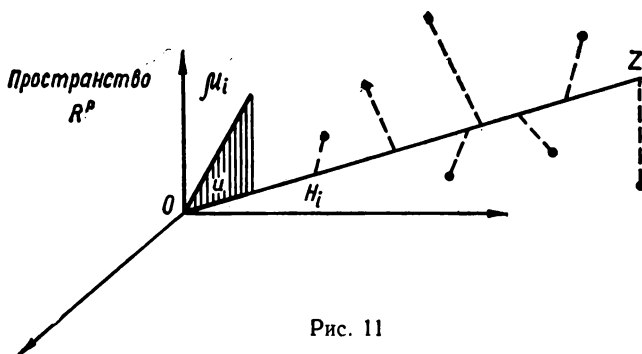


Рис. 11

Наша задача состоит в том, чтобы представить набор точек (из R^p или R^n) в пространстве меньшей размерности, стараясь при этом потерять как можно меньше информации о взаимном расположении точек в исходном множестве, т. е. мы хотим определить q -мерное подпространство R^q так, чтобы q оказалось существенно меньше p и таким, чтобы исходные n точек были достаточно представлены этим подпространством.

Начнем с того, что определим одномерное подпространство в виде «прямой», проходящей через начало координат рассматриваемой системы точек так, чтобы эта прямая наилучшим образом (в смысле некоторого критерия) приближала исходное облако точек. Это показано на рис. 11.

Пусть u — единичный вектор, тогда $u^T u = I$.

Проекция OH_i вектора OM_i на одномерное подпространство определяется как скалярное произведение OM_i и u (т. е. определяется сумма попарных произведений составляющих).

Каждая из n строк матрицы X есть вектор в R^p . Поэтому произведение матрицы X и вектора u есть вектор-столбец с n элементами, каждый из которых является скалярным произведением строки матрицы X и вектора u .

Таким образом, n компонент X являются проекциями n исходных точек на направление, определяемое вектором u .

Будем искать такое множество L , для которого сумма квадратов расстояний

$$\sum_{i=1}^n (M_i H_i)^2 = \sum_{i=1}^n (OM_i)^2 - \sum_{i=1}^n (OH_i)^2 \quad (3.1)$$

была бы минимальной.

Так как сумма $\sum_{i=1}^n (OM_i)^2$, характеризующая распределение исходного множества n точек в пространстве R^p , не зависит от вектора L , то наша задача может быть сведена к максимизации выражения

$$\sum_{i=1}^n (OH_i)^2 = (Xu)^T Xu = u^T X^T Xu \quad (3.2)$$

за счет выбора вектора u при ограничениях $u^T u = I$.

Пусть u_1 — вектор, при котором максимум (3.2) достигается. Тогда ставим задачу найти вектор u_2 , который был бы ортогонален вектору u_1 и максимизировал бы выражение $u_2^T X^T X u_2$. Таким образом, можно построить q -мерное подпространство ($q \leq p$). При этом векторы u_1, u_2, \dots, u_q являются ортогональными собственными векторами матрицы $X^T X$, соответствующие q наибольшим собственным значениям

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_q.$$

Этот результат показывает, что вектор u_1 характеризует одномерное подпространство, наилучшим образом (в смысле наименьших квадратов) приближающее n точек в R^p , и ему соответствует наибольшее собственное значение λ_1 матрицы $X^T X$. И, вообще говоря, q -мерное подпространство наилучшим образом приближает n точек R^p , определяемое первыми q собственными векторами, соответствующими q наибольшим собственным значениям симметричной матрицы $X^T X$.

Рассмотрим пространство R^n ; причем теперь матрица X будет рассматриваться как набор p точек, n координат которых представлены столбцами матрицы X .

Будем искать единичный вектор V (и соответственно q -мерное подпространство) такой, чтобы максимизировалась сумма квадратов p проекций исходных точек на вектор V , т. е. вектор V , максимизирующий выражение

$$(X^T V)^T X^T V = V^T X X^T V$$

при ограничении $V^T V = I$.

Пусть V_α — собственный вектор матрицы $X X^T$, соответствующий собственному значению μ_α .

В пространстве R^p имеем уравнение

$$X^T X u_1 = \lambda_1 u_1, \quad (3.3)$$

а в пространстве R^n

$$X X^T V_1 = \mu_1 V_1. \quad (3.4)$$

Умножив (3.3) на X , получим

$$(X X^T) X u_1 = \lambda_1 (X u_1). \quad (3.5)$$

Это уравнение показывает, что для любого собственного вектора u_1 матрицы $X^T X$ относительно ненулевого собственного значения λ_1 существует собственный вектор $X u_1$ матрицы $X X^T$, соответствующий тому же собственному значению λ_1 .

В свою очередь, μ_1 — наибольшее собственное значение матрицы $X X^T$, поэтому с учетом предыдущего предложения заключаем, что $\lambda_1 = \mu_1$.

С другой стороны, перемножим обе части (3.4) на X^T . Видим, что $X^T V_1$ — собственный вектор, соответствующий собственному значению μ_1 . Повторяя предыдущие рассуждения, заключаем, что $\mu_1 \leq \lambda_1$, из чего следует $\lambda_1 = \mu_1$.

Таким образом, приходим к выводу, что матрицы XX^T и X^TX имеют одинаковую последовательность ненулевых собственных значений.

Заметим, что вектор Xu_α имеет квадрат нормы λ_α (так как $u_\alpha^T X^T X u_\alpha = \lambda_\alpha$) и, кроме того, единичный вектор V_α соответствует тому же собственному значению λ_α . Если $\lambda_\alpha \neq 0$, то

$$V_\alpha = Xu_\alpha / \sqrt{\lambda_\alpha} \quad (3.6)$$

и, в свою очередь,

$$u_\alpha = X^T V_\alpha / \sqrt{\lambda_\alpha}; \quad (3.7)$$

u_α называются α -главными осями в пространстве R^p , а V_α — α -главными осями в R^n .

В подпространстве R^p , образованном осями u_α , координатами точек являются компоненты Xu_α ; координаты точек могут быть выражены компонентами $V_\alpha / \sqrt{\lambda_\alpha}$.

Координаты точек на соответствующих (α) главных осях в R^p пропорциональны компонентам на соответствующих (α) главных осях в R^n , и наоборот. Этот вывод имеет важное значение при визуализации данных.

Восстановление исходных данных. Перепишем (3.6) в виде

$$Xu_\alpha = \sqrt{\lambda_\alpha} V_\alpha. \quad (3.8)$$

Умножим (3.8) справа на u_α^T и просуммируем по α . При этом $\alpha = 1, 2, \dots, p$ (так как рассматривается пространство R^p); часть собственных значений λ_α могут принимать нулевые значения.

Имеем

$$X \sum_{\alpha=1}^p u_\alpha u_\alpha^T = \sum_{\alpha=1}^p \sqrt{\lambda_\alpha} V_\alpha u_\alpha^T. \quad (3.9)$$

Пусть V — матрица порядка $p \times p$, у которой столбцами являются собственные векторы u_α матрицы $X^T X$. Так как собственные векторы ортогональны и имеют норму (3.1), имеем $V^T V = I$ и $V V^T = I$.

Следовательно,

$$\sum_{\alpha=1}^p u_\alpha u_\alpha^T = V V^T,$$

откуда из (3.9) получим

$$X = \sum_{\alpha=1}^p \sqrt{\lambda_\alpha} V_\alpha u_\alpha^T. \quad (3.10)$$

Эта формула может рассматриваться как формула для восстановления матрицы исходных данных на основании базиса, построенного с помощью собственных значений λ_α .

Если $q < p$ и последующие $q + 1, q + 2, \dots$ собственные значения малы, то исходная матрица может быть представлена достаточно хорошо первыми q слагаемыми

$$X \approx X^* = \sum_{\alpha=1}^q \sqrt{\lambda_\alpha} V_\alpha u_\alpha^T. \quad (3.11)$$

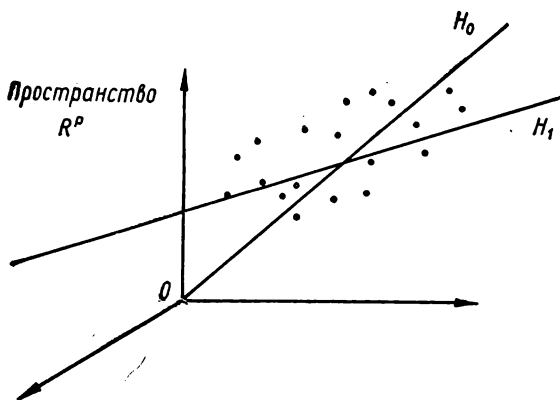


Рис. 12

Отметим, что вектор V_α имеет n компонент, а u_α — p компонент. Следовательно, np элементов матрицы X аппроксимируется с помощью $q(n+p)$ элементов. Например, если $n = 300$, $p = 500$ и $q = 10$, то 150 000 элементов исходной матрицы приближаются 8000 элементами. Как правило, на практике приближение оказывается довольно хорошим.

Качество аппроксимации с помощью q компонент оценивается величиной

$$\tau_q = \frac{\sum_{i,j} x_{ij}^{*2}}{\sum_{i,j} x_{ij}^2} \quad (3.12)$$

или

$$\tau_q = \frac{\text{sp}(X^* X^*)}{\text{sp}(X^T X)},$$

где sp означает след матрицы.

Подставляя вместо X и X^* их значения из (3.10) и (3.11), получаем

$$\tau_q = \sum_{\alpha \leq q} \lambda_\alpha / \sum_{\alpha=1}^p \lambda_\alpha.$$

Видим, что $\tau_q \leq 1$. Величину τ_q называют долей общей дисперсии, воспроизводимой q первыми факторами.

Общий анализ. Выше был рассмотрен анализ главных компонент с точки зрения сокращения исходного объема данных. Здесь подойдем к вопросу анализа с другой стороны, а именно с точки зрения изучения взаимного расположения исходных данных в R^p или R^n . Такой подход позволяет сделать заключение о сходстве (или различии) объектов (признаков).

Проведем рассуждения относительно пространства R^p . В этом пространстве с помощью матрицы R заданы n объектов. Рассмотрим всевозможные пары объектов. Ставим задачу определить линию H_1 , вдоль которой все пары растянуты максимально так, что сумма квадратов всех расстояний между объектами максимальна, в то время как сумма квадратов расстояний точек от этой линии была бы минимальна. На рис. 12 видно, что эта линия может не проходить через начало координат.

Пусть h_i и $h_{i'}$ — проекции точек i и i' на H_1 . Можно записать следующее выражение:

$$\begin{aligned} \sum_{i,i'}^n (h_i - h_{i'})^2 &= n \sum_{i=1}^n h_i^2 + n \sum_{i'=1}^n h_{i'}^2 - 2 \left(\sum_{i=1}^n h_i \right) \left(\sum_{i'=1}^n h_{i'} \right) = \\ &= 2n^2 \left(\frac{1}{n} \sum_{i=1}^n h_i^2 - h_0^2 \right) = 2n \sum_{i=1}^n (h_i - h_0)^2, \end{aligned}$$

где $h_0 = \frac{1}{n} \sum_{i=1}^n h_i$ и, следовательно, является проекцией центра масс G облака на линию H_1 .

Замечаем, что j -й координатой G в R^p является $r_j = \frac{1}{n} \sum_{i=1}^n r_{ij}$, r_{ij} — элемент матрицы R .

Таким образом, если начало координат системы (см. рис. 12) поместить в точку G , то анализ сводится к проблеме, рассмотренной выше. Этого можно достигнуть, заменив матрицу R матрицей X так, что $x_{ij} = r_{ij} - r_{0j}$.

Квадрат расстояния между двумя объектами K и K' можно выразить в виде

$$d^2(K, K') = \sum_{j=1}^p (r_{kj} - r_{k'j})^2.$$

Если каждое из слагаемых необходимо рассматривать как равноправное, то следует произвести нормировку, тогда

$$d^2(K, K') = \sum_{j=1}^p \left(\frac{r_{kj} - r_{k'j}}{S_j \sqrt{n}} \right)^2,$$

где S_j рассматривается как стандартное отклонение j -переменной:

$$S_j^2 = \frac{1}{n} \sum_{i=1}^n (r_{ij} - r_{0j})^2. \quad (3.13)$$

В этом случае при анализе матрица X имеет общий элемент

$$x_{ij} = \frac{r_{ij} - r_{0j}}{S_j \sqrt{n}}. \quad (3.14)$$

При определении главных осей используется матрица $C = X^T X$ с общим элементом

$$C_{ij} = \sum_i x_{ij} x_{ij'} = \frac{\frac{1}{n} \sum_i (r_{ij} - r_{0j})(r_{ij'} - r_{0j'})}{S_j S_{j'}}. \quad (3.15)$$

Таким образом, $C_{ij'}$ является коэффициентом корреляции между переменными j и j' .

Выражение (3.13) должно центрировать и нормировать элементы матрицы R . Это имеет значительные последствия, т. е. теперь роль строк и столбцов матрицы фиксирована.

Выразим квадрат расстояния между признаками j и j' в виде

$$d(j, j') = \sum_{i=1}^n (x_{ij} - x_{ij'})^2 = \sum_{i=1}^n x_{ij}^2 + \sum_{i=1}^n x_{ij'}^2 - 2 \sum_{i=1}^n x_{ij} x_{ij'}.$$

Учитывая (3.13), находим, что

$$\sum_{i=1}^n x_{ij}^2 = \sum_{i=1}^n x_{ij'}^2 = 1,$$

т. е. точки-признаки располагаются на сфере единичного радиуса с центром в начале осей.

Учитывая (3.15), получаем

$$d^2(j, j') = 2(1 - C_{jj'}).$$

Заметим, что расстояние такого типа обладает следующими свойствами:

1) признаки с коэффициентом корреляции $C_{jj'} \approx 1$ практически совпадают;

2) при $C_{jj'} \approx -1,0$ признаки максимально удалены;

3) два некоррелированных признака ($C_{jj'} = 0$) находятся на измененном расстоянии друг от друга.

Дополнительные переменные. По разным причинам бывает полезным дополнить матрицу R . Это можно сделать двояко: увеличивая количество либо строк, либо столбцов. При этом нас может интересовать вопрос, как эта дополнительная информация соотносится с уже имеющейся.

Пусть R^+ — дополнительные p_s столбцов и R_+ — дополнительные n_s строк к матрице R (рис. 13). Матрицы R^+ и R_+ преобразовываются соответственно в X^+ и X_+ . В R^n имеем p_s дополнительных признаков и, следовательно,

$$x_{ij}^+ = \frac{r_{ij}^+ - r_{0j}^+}{S_j \sqrt{n}}.$$

С учетом (3.6) и (3.7) абсциссы p_s дополнительных переменных могут быть выражены p_s -компонентами вектора $(X^+)^T V_\alpha$. В R^p имеем

$$x_{+ij} = \frac{r_{+ij} - r_{+0j}}{S_j \sqrt{n}}.$$

Абсциссы n_s дополнительных объектов могут быть выражены n_s -компонентами вектора $X_+ u_\alpha$.

При наличии дополнительных элементов элементы матрицы X называют активными, а дополнительные элементы (элементы матрицы X^+ и X_+) — иллюстративными.

Ч и с л е н н ы й п р и м е р. Для того чтобы показать на примере, каким образом изложенные идеи могут быть использованы практически, рассмотрим следующую задачу (упрощенную по сравнению с реальными задачами, но вместе с тем обладающую характерными особенностями таких задач).

Для оптимального протекания некоторого технологического процесса его необходимо обеспечить соответствующими ресурсами. В табл. 1 представлены данные, характеризующие поступление таких ресурсов: столбцы 77—88 обозначают время, к которому относятся объемы. Данная таблица получена в результате деления каждой из строк реальной таблицы на значение, соответствующее времени,— столбец 82. Таким образом, значения табл. 1

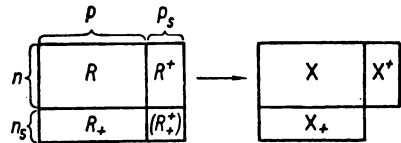


Рис. 13

Т а б л и ц а 1

| № п/п | 77 | 78 | 79 | 80 | 81 | 82 | 83 | 84 | 85 | 86 | 87 | 88 |
|-------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1 | 1701 | 1774 | 1415 | 1169 | 1130 | 1000 | 1021 | 1038 | 1549 | 1374 | 1284 | 1345 |
| 2 | 879 | 982 | 1215 | 1014 | 1064 | 1000 | 1040 | 1005 | 1407 | 1432 | 1381 | 1451 |
| 3 | 1500 | 1524 | 1373 | 1154 | 1065 | 1000 | 1012 | 1000 | 1205 | 1229 | 1089 | 1152 |
| 4 | 963 | 1262 | 1192 | 1141 | 1269 | 1000 | 1182 | 1186 | 1542 | 1460 | 1305 | 1500 |
| 5 | 651 | 819 | 848 | 849 | 892 | 1000 | 832 | 955 | 991 | 1122 | 877 | 930 |
| 6 | 517 | 556 | 866 | 1056 | 1207 | 1000 | 1202 | 1314 | 1351 | 1714 | 1421 | 1699 |
| 7 | 586 | 733 | 799 | 843 | 877 | 1000 | 842 | 988 | 1007 | 1133 | 864 | 957 |
| 8 | 489 | 489 | 800 | 1046 | 1207 | 1000 | 1225 | 1345 | 1451 | 1835 | 1502 | 1787 |
| 9 | 203 | 377 | 582 | 564 | 768 | 1000 | 997 | 1210 | 1398 | 2278 | 2010 | 2447 |
| 10 | 378 | 405 | 882 | 990 | 1345 | 1000 | 1055 | 1203 | 1292 | 2848 | 2953 | 3837 |
| 11 | 99 | 311 | 395 | 495 | 755 | 1000 | 1043 | 1205 | 1568 | 2008 | 2056 | 2130 |
| 12 | 271 | 359 | 686 | 900 | 1306 | 1000 | 1130 | 1374 | 1404 | 2378 | 2581 | 2298 |
| 13 | 1054 | 1101 | 1160 | 1086 | 1136 | 1000 | 1113 | 1136 | 1152 | 1475 | 1311 | 1479 |
| 14 | 925 | 922 | 1191 | 1051 | 1320 | 1000 | 1341 | 1461 | 1460 | 1964 | 1913 | 2152 |
| 15 | 851 | 902 | 1094 | 1114 | 1094 | 1000 | 976 | 902 | 1012 | 1078 | 871 | 980 |
| 16 | 828 | 827 | 1333 | 1176 | 1309 | 1000 | 1081 | 1103 | 1263 | 1323 | 1306 | 1374 |
| 17 | 794 | 1002 | 1296 | 1178 | 1161 | 1000 | 1022 | 1134 | 1144 | 1110 | 1092 | 1035 |
| 18 | 465 | 995 | 1139 | 965 | 1022 | 1000 | 1155 | 1315 | 1077 | 1277 | 1130 | 943 |
| 19 | 655 | 1040 | 1190 | 1220 | 1225 | 1000 | 1035 | 1038 | 970 | 1204 | 1210 | 1126 |
| 20 | 411 | 945 | 890 | 998 | 1107 | 1000 | 989 | 993 | 1035 | 1460 | 1335 | 1274 |
| 21 | 825 | 922 | 1059 | 1037 | 988 | 1000 | 1030 | 993 | 1100 | 1241 | 1125 | 1195 |
| 22 | 686 | 772 | 979 | 965 | 920 | 1000 | 1151 | 969 | 1611 | 1900 | 1651 | 1982 |
| 23 | 879 | 886 | 928 | 1031 | 1046 | 1000 | 1064 | 1085 | 1150 | 1303 | 1286 | 1302 |
| 24 | 926 | 973 | 1029 | 1128 | 1179 | 1000 | 1269 | 1219 | 1678 | 2310 | 2277 | 2363 |
| 25 | 1249 | 1442 | 1587 | 1277 | 1205 | 1000 | 1080 | 1299 | 1196 | 1284 | 1373 | 1298 |
| 26 | 791 | 828 | 1056 | 1241 | 1104 | 1000 | 1091 | 1542 | 2274 | 1803 | 1853 | 1962 |
| 27 | 1237 | 867 | 1154 | 1173 | 1185 | 1000 | 1029 | 1212 | 1187 | 1325 | 1327 | 1441 |
| 28 | 1081 | 853 | 1118 | 1297 | 1386 | 1000 | 1389 | 1667 | 2622 | 2291 | 2076 | 2709 |
| 29 | 1040 | 1083 | 1029 | 946 | 850 | 1000 | 1422 | 1543 | 1210 | 1378 | 1266 | 1143 |
| 30 | 1332 | 1431 | 1376 | 1080 | 1162 | 1000 | 1067 | 1623 | 1860 | 1939 | 1737 | 1865 |
| 31 | 1613 | 1694 | 1496 | 1597 | 1343 | 1000 | 2319 | 2730 | 3020 | 2532 | 2073 | 2298 |
| 32 | 957 | 953 | 1016 | 969 | 1108 | 1000 | 1113 | 1646 | 2180 | 2108 | 1918 | 2418 |
| 33 | 955 | 1120 | 1231 | 1095 | 1078 | 1000 | 999 | 1110 | 1170 | 1231 | 1163 | 1167 |
| 34 | 853 | 910 | 1164 | 1039 | 1182 | 1000 | 1184 | 1250 | 1468 | 1741 | 1677 | 1897 |
| 35 | 886 | 954 | 1067 | 1080 | 1089 | 1000 | 1008 | 1077 | 1098 | 1256 | 1181 | 1215 |
| 36 | 852 | 965 | 1158 | 1148 | 1283 | 1000 | 1143 | 1180 | 1447 | 1505 | 1432 | 1559 |

приобрели однородный характер и таблица данных может условно интерпретироваться как «таблица сопряженности». Не будем утверждать, что этот путь «стандартизации» данных окажется подходящим всегда. Этот вопрос особый, и здесь он не рассматривается.

| Номер собствен- ных зна- чений | Процент от общей дистер- сии | Суммар- ный про- цент | Гистограмма распределения собственных значений |
|---|---------------------------------------|-----------------------------|--|
| 1 | 67,44 | 67,44 | +++++ |
| 2 | 13,55 | 80,99 | +++++ |
| 3 | 7,96 | 88,95 | +++++ |
| 4 | 3,75 | 92,70 | ++++ |
| 5 | 2,73 | 95,43 | +++ |
| 6 | 1,90 | 97,33 | ++ |
| 7 | 1,13 | 98,46 | ++ |
| 8 | 0,69 | 99,15 | + |
| 9 | 0,52 | 99,66 | + |
| 10 | 0,34 | 100,00 | + |
| 11 | 0,00 | 100,00 | + |
| 12 | 0,00 | 100,00 | + |
| 13 | 0,00 | 100,00 | + |

Добавим, что столбец 88 получен в результате «прогноза таблицы». Какую роль сыграли при этом методы анализа данных, будет видно из дальнейшего.

Т а б л и ц а 3

| № п/п | F1 | F2 | F3 | № п/п | F1 | F2 | F3 |
|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 | —0,28 | 0,08 | 0,13 | 19 | —0,11 | —0,13 | 0,03 |
| 2 | —0,06 | —0,01 | —0,02 | 20 | 0,03 | —0,11 | 0,03 |
| 3 | —0,29 | 0,02 | —0,10 | 21 | —0,10 | —0,05 | 0,01 |
| 4 | —0,11 | 0,01 | 0,01 | 22 | 0,10 | 0,02 | —0,03 |
| 5 | —0,09 | —0,05 | 0,05 | 23 | —0,06 | —0,03 | 0,00 |
| 6 | 0,12 | —0,05 | 0,09 | 24 | 0,12 | 0,02 | —0,09 |
| 7 | —0,05 | —0,05 | 0,08 | 25 | —0,22 | —0,03 | —0,05 |
| 8 | 0,16 | —0,03 | 0,09 | 26 | 0,09 | 0,10 | 0,04 |
| 9 | 0,39 | 0,01 | —0,02 | 27 | —0,12 | —0,01 | —0,03 |
| 10 | 0,38 | —0,11 | —0,14 | 28 | 0,10 | 0,13 | 0,02 |
| 11 | 0,44 | 0,04 | 0,03 | 29 | —0,09 | 0,06 | 0,06 |
| 12 | 0,37 | —0,07 | —0,04 | 30 | —0,07 | 0,10 | —0,06 |
| 13 | —0,11 | —0,03 | —0,03 | 31 | —0,02 | 0,19 | 0,08 |
| 14 | 0,05 | —0,01 | —0,02 | 32 | 0,10 | 0,13 | —0,01 |
| 15 | —0,17 | —0,09 | 0,04 | 33 | —0,15 | —0,03 | 0,00 |
| 16 | —0,07 | —0,07 | 0,03 | 34 | 0,02 | —0,01 | —0,02 |
| 17 | —0,14 | —0,08 | 0,05 | 35 | —0,11 | —0,05 | 0,01 |
| 18 | —0,04 | —0,07 | 0,11 | 36 | —0,04 | —0,03 | 0,01 |

Анализ собственных значений матрицы табл. 1 представлен в табл. 2. При анализе использовались столбцы с идентификаторами 77—87; столбец 88 будет использован как иллюстративный. В табл. 2 в первой колонке даны номера собственных значений; во второй — процент от общей дисперсии, объясняемый фактором, соответствующим данному собственному значению; третья колонка содержит суммарный (накопленный) процент, объясняемый совместно предыдущими факторами (например, первый — 67,44 %, первый и второй — 67,44 % + 13,55 % = 80,99 и т. д.). Далее представлена гистограмма распределения собственных значений.

Табл. 3 соответствует табл. 1, «сжатой» по столбцам (было 11, не считая столбца 88, а стало три), а табл. 4 соответствует табл. 1, «сжатой» по строкам (было 36 строк, а осталось три; таблица представлена в транспонированном виде).

Т а б л и ц а 4

| № п/п | F1 | F2 | F3 |
|-------|-------|-------|-------|
| 77 | —0,32 | 0,11 | —0,10 |
| 78 | —0,30 | 0,02 | —0,04 |
| 79 | —0,18 | —0,06 | —0,01 |
| 80 | —0,11 | —0,07 | 0,03 |
| 81 | —0,02 | —0,11 | 0,01 |
| 82 | —0,01 | —0,11 | 0,04 |
| 84 | 0,01 | 0,00 | 0,08 |
| 84 | 0,04 | 0,06 | 0,09 |
| 85 | 0,08 | 0,14 | 0,04 |
| 86 | 0,21 | 0,00 | —0,04 |
| 87 | 0,24 | —0,02 | —0,09 |
| 88 | 0,28 | 0,01 | —0,11 |

С помощью обеих таблиц можно восстановить табл. 1, разумеется, с некоторой потерей точности, определяемой скоростью сходимости суммарного накопленного процента (см. табл. 2) к 100 %.

Этот момент был использован для получения прогнозных данных табл. 1 на время 88. Собственно прогноз можно получить различными путями: запрогнозировать каждую строку в отдельности, получить искомый набор с помощью множественного регрессивного анализа (с тем, чтобы учесть взаимную зависимость строк) и т. д.

В данном примере прогноз был получен из иных соображений. Были отобраны несколько факторов, соответствующих наибольшим собственным значениям. Последующие факторы вносят все меньше информации и, как правило (хотя и не всегда), несут «шумовой» эффект. Поэтому ими можно пренебречь. Мы ограничились тремя первыми факторами. Как следует из ранее изложенного, факторы ортогональны между собой. Следовательно, процедура прогнозирования каждого фактора в отдельности вполне законна. Последующий переход к исходной табл. 1 с учетом прогнозных значений факторов дает нам набор 88 ожидаемого распределения ресурсов, на который можно ориентироваться при оценке изучаемого процесса. Не вдаваясь в подробности, заметим, что такой подход к прогнозу больших массивов информации, заданной таблично, имеет заметное преимущество перед другими методами (в частности, методом множественной регрессии).

Для визуального анализа данные табл. 3 и 4 были использованы для отображения на факторную плоскость. На рис. 14 горизонтальная ось соответствует первому фактору, вертикальная — второму. Оче-

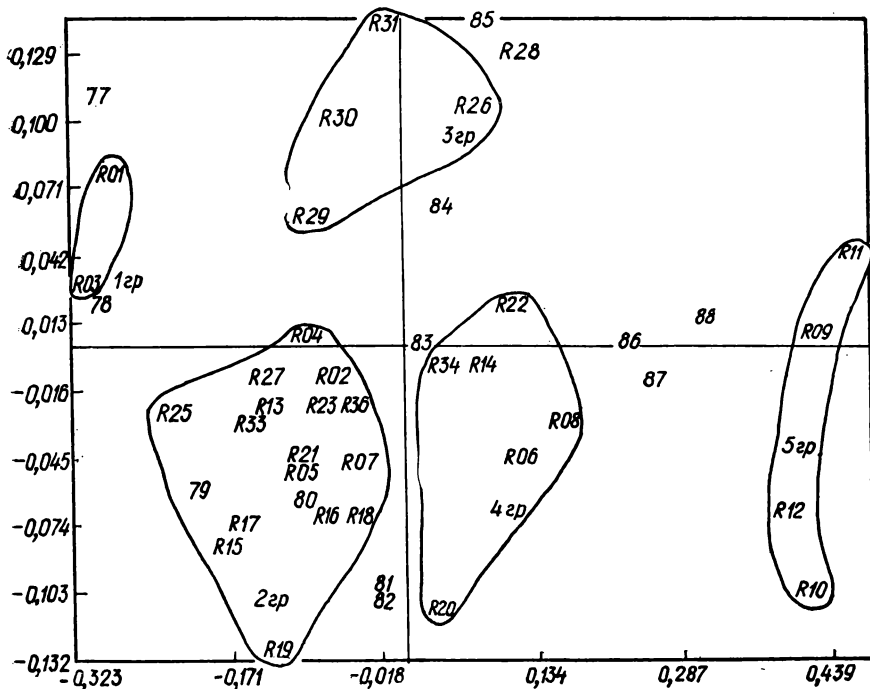


Рис. 14

видно, что аналогичное отображение можно построить для второго и третьего факторов и т. д. На рис. 14 точка «88» (полученная в результате прогноза и, следовательно, не учтенная при определении факторов) является дополнительной (иллюстративной). Об этом говорилось выше. Видим, что «88» ведет себя «схоже» с «86» и «87», но совсем «непохоже» на «85». Обращает на себя внимание, что кривые R_{02} , R_{04} , R_{13} , R_{23} , R_{27} , R_{36} и т. д. ведут себя «сходным» образом. Схожие группы обозначены на графике, но это удалось сделать в результате проведения следующего этапа АД — классификации.

После проведения этого мероприятия мы, естественно, приходим к мысли, что прогнозировать надо бы не сразу всю таблицу, а каждую сходную группу в отдельности. И действительно, в результате полной процедуры прогноз удалось еще несколько улучшить. Таким образом, сочетание компонентного анализа и автоматической классификации является мощным инструментом методов АД. Поэтому перейдем к рассмотрению основных методов автоматической классификации и затем снова вернемся к данному примеру.

Автоматическая классификация. Представление классов центрами масс довольно распространено. Широкое практическое применение получила классификация относительно подвижного центра масс. Хотя этот метод базируется на относительно слабой теоретической основе, а его эффективность больше подтверждается в ходе практических работ, он, по всей видимости, является одним из лучших методов разбиения больших массивов данных. Его чаще используют в сочетании с другими методами (например, для предварительной кластеризации перед применением методов главных компонент). Этот метод относится к классу методов, известных как алгоритм k -средних.

Алгоритм работает следующим образом. Пусть $i \in I$, $i = 1, 2, \dots, n$, I — множество объектов, которые описаны некоторым набором признаков. Предположим, что в пространстве R^p , содержащем n точек, определено евклидово расстояние d . Максимальное число кластеров обозначим K_0 .

Далее рассмотрим последовательность шагов.

Шаг 0. Из подлежащей кластеризации совокупности объектов извлекаем случайным образом K_0 элементов, которые используем в качестве начальных приближений для центров K_0 искомых кластеров. Эти K_0 центров

$$C_1^0, C_2^0, \dots, C_k^0, \dots, C_{k_0}^0$$

разбивают множество I на k_0 кластеров

$$I_1^0, I_2^0, \dots, I_k^0, \dots, I_{k_0}^0$$

в соответствии с правилом: точка i включается в класс I_k^0 , если она ближе к C_k^0 , чем по всем остальным центрам (рис. 15).

Шаг 1. Определим K_0 центров новых кластеров

$$C_1^1, C_2^1, \dots, C_{k_0}^1,$$

используя центры тяжести кластеров

$$I_1^0, I_2^0, \dots, I_{k_0}^0.$$

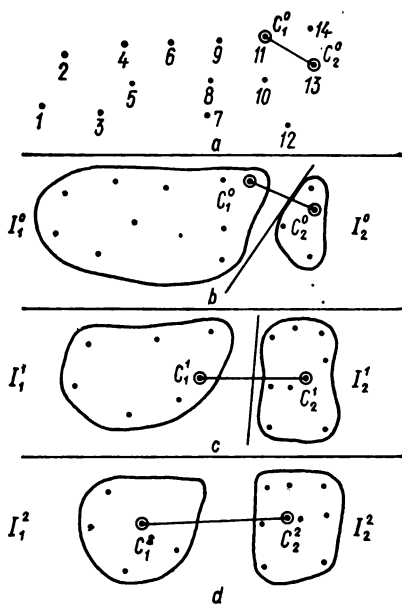


Рис. 15

Эти новые центры образуют новые разбиения в соответствии с теми же правилами, что и на предыдущем шаге. Получим разбиение

$$I_1^1, I_2^1, \dots, I_{k_0}^1.$$

Шаг m . Определим K_0 центров новых кластеров

$$C_1^m, C_2^m, \dots, C_{K_0}^m,$$

используя центры масс кластеров

$$I_1^{m-1}, I_2^{m-1}, \dots, I_{K_0}^{m-1}.$$

Новые центры образуют разбиения исходного множества I на кластеры

$$I_1^m, I_2^m, \dots, I_{K_0}^m.$$

Алгоритм заканчивается либо когда две последовательные итерации приводят к одному и тому же разбиению, либо когда выбранный критерий (например, значение дисперсии внутри кластеров) перестает существенно

уменьшаться, либо когда достигнуто предварительно установленное максимальное число итераций.

В каждом случае результирующее разбиение зависит от выбора исходных положений центров масс кластеров.

Покажем, что внутригрупповая дисперсия от шага к шагу убывает. Более того, за счет выбора соответствующих правил назначения можно добиться строгого убывания дисперсии (лишь бы исходное множество I было зафиксировано).

Предположим, что каждому из объектов множества I присвоены веса, причем

$$\sum_{i=1}^n p_i = 1,$$

и пусть $d^2(i, C_k^m)$ — квадрат расстояния между i -м объектом и центром k -го кластера на шаге m .

Введем количественный критерий

$$V(m) = \sum_{k=1}^{K_0} \left\{ \sum_{i \in I_k^m} p_i d^2(i, C_k^m) \right\}. \quad (3.16)$$

Вспомним, что на шаге m группа I_k^m сформирована из объектов, которые ближе к C_k^m , чем к другим центрам (эти центры являются центрами масс групп I_k^{m-1} предыдущего шага).

Внутригрупповая дисперсия на m -м шаге имеет вид

$$V(m) = \sum_{k=1}^{k_0} \left\{ \sum_{i \in I_k^m} p_i d^2(i, C_k^{m+1}) \right\}, \quad (3.17)$$

где C_k^{m+1} — центр масс группы I_k^m . Для $m + 1$ -го шага имеем

$$V(m+1) = \sum_{k=1}^{k_0} \left\{ \sum_{i \in I_k^{m+1}} p_i d^2(i, C_k^{m+1}) \right\}. \quad (3.18)$$

Покажем, что

$$V(m-1) \geq V(m) \geq V(m+1). \quad (3.19)$$

Выражение (3.17) отличается от (3.16) тем, что в фигурных скобках указана инерция класса относительно центра масс, в то время как в (3.16) точка, относительно которой определяется инерция, может не совпадать с центром масс. В таком случае относительно этой точки инерция класса будет больше. Из этого следует справедливость левой части неравенства (3.19). Справедливость правой его части вытекает из того, что при переходе от (3.17) к (3.18) происходит только переназначение объектов к тем центрам, к которым они в данный момент ближе. Такое переназначение может привести только к уменьшению (3.18) по сравнению с (3.17).

Существует целый ряд алгоритмов, которые в принципе подобны изложенному. Однако они имеют ряд особенностей.

Метод динамических сгущений. Нетрудно показать, что большая группа задач АК может быть выражена в форме оптимизации критерия, хорошо определенного с математической точки зрения. Метод динамических сгущений оптимизирует критерий, который оценивает согласованность между покрытием множества объектов, подлежащих классификации, и способом представительства классов этого покрытия. Напомним, что сгущением называется такое множество объектов S , в котором любая его внутренняя связь x_{ij} ($i, j \in S$) превышает внешние x_{ki} ($k \in S, l \notin S$). Выше мы рассмотрели случай, когда в качестве покрытия использовались разбиения, а в качестве представительства — центры масс классов. В методах динамических сгущений задача оптимизации формулируется как задача одновременного отыскания покрытия и представительства его классов среди множества возможных покрытий и представительств, которые оптимизируют заданный критерий. При этом представитель класса имеет самый широкий смысл; центр тяжести, точка класса, группа точек класса, ось, случайная переменная. В качестве покрытий (кроме разбиений) рассматриваются иерархии и покрытия.

Метод k средних. Начинают с k кластеров (например, берут первые k точек). В качестве критерия используется минимальное расстояние внутри кластера относительно среднего.

Как только элемент включается в кластер, среднее пересматривается. Это продолжается до тех пор, пока не размещены все точки. Такая процедура обеспечивает хорошее разбиение в единичной итерации.

Но она зависит от порядка записи в файле, в то время, как описанная выше процедура таким недостатком не обладает.

ISODATA. Начинают с k кластеров и относят все элементы в соответствии с внутрикластерным критерием минимальности. После того, как все элементы отнесены, средние пересматриваются и все продлевается снова до тех пор, пока не перестанем получать улучшение внутрикластерного критерия минимальности.

Устойчивые кластеры. Алгоритмы классификации относительно перемещающегося центра сходятся к локальному минимуму. Проблема нахождения оптимального разбиения на K_0 кластеров с использованием критерия минимума дисперсии внутри группы до настоящего времени не решена ввиду отсутствия удовлетворительного алгоритма. Разбиения в общем зависят от выбора начальных центров.

Процедура нахождения устойчивых кластеров, предложенная Дидэ, является по крайней мере частичным решением проблемы.

Метод состоит в выполнении нескольких разбиений при различных наборах центров и выделении устойчивых кластеров, т. е. в определении объектов, принадлежащих к одному и тому же кластеру в каждом разбиении.

Пусть p_1, p_2, \dots, p_s — S результатов разбиения множества I на k_0 кластеров. Имеем $(k_0)^s$ кластеров. Устойчивые группы образуют те объекты, которые в разбиениях p_1, p_2, \dots, p_s чаще оказываются в одном кластере. На практике число устойчивых кластеров с прочными базами (число повторяющихся объектов) значительно меньше $(K_0)^s$. Так, например, для четырех разбиений на пять кластеров для 1000 объектов получено 20 устойчивых кластеров с базами, которые больше 10.

Иерархическая классификация. Пусть между парой объектов установлена мера близости. Далее, существуют правила расчета мер близости между расчлененными кластерами объектов. Например, если x, y и z — три объекта и если x и y объединяются в кластер h , можно определить меру близости между этим кластером и z как минимальное расстояние между элементами h и z , т. е.

$$d(h, z) = \min \{d(x, z), d(y, z)\}.$$

Это расстояние определяет односвязный метод (ближайшего соседа).

Другое правило, которое часто используется для определения расстояния между h и z , выражается в виде

$$d(h, z) = \frac{d(x, z) + d(y, z)}{2}.$$

В общем случае, когда x и y состоят соответственно из n_x и n_y элементов, а h — из $n_x + n_y$ элементов, имеем

$$d(h, z) = \frac{n_x d(x, z) + n_y d(y, z)}{n_x + n_y}.$$

Алгоритм классификации. Алгоритм восходящей иерархической классификации состоит в следующем. Будем называть точками объек-

ты, подлежащие классификации, или кластеры объектов, сформированные алгоритмом.

1. Имеем n классифицируемых точек (n объектов).

2. Находим две ближайшие друг к другу точки и объединяем их.

3. Рассчитываем меру близости между новой точкой и оставшимися точками. Возвращаемся к шагу 1, имея $n - 1$ точек.

4. Снова определяем две ближайшие точки, которые снова объединяем. Пересчитываем меры близости, и процедуру повторяем до тех пор, пока не останется одна точка.

Сделаем некоторые замечания. В случае односвязного метода алгоритм использует меры близости в виде неравенств. При этом «дерево» оказывается вытянутым с равностоящими уровнями.

Семейство кластеров, формируемых данным алгоритмом, образует иерархию. Этому семейству свойственно то, что оно содержит как весь массив объектов, так и отдельные объекты.

Каждый раз новый кластер формируется из расчлененных точек; новый кластер становится, в свою очередь, новой точкой и, следовательно, обязательно включается в последующий кластер (рис. 16, а).

Иерархия является индексной, если численные значения $V(h) \geq 0$ связаны с каждой частью h иерархии таким образом, что $V(h) \leq V(h')$, если $h \subset h'$. Иерархия, представленная на рис. 15, проиндексирована по значениям меры близости, соответствующим каждому агрегированию. Эти меры близости представляются ординатой.

Проведем горизонтальную линию L (рис. 16, б). Образованный этой линией срез дерева образует разбиение исходного множества на две группы $\{4, 5\}$ и $\{1, 2, 3\}$. По мере опускания вниз число групп в срезе будет возрастать. Если линия будет опущена ниже уровня 1, то срез дерева образует разбиение, при котором в каждой группе содержится только один объект. Таким образом, иерархия позволяет получить набор разбиений от n групп до одной. Рассмотрим классификацию по односвязному методу (ближайшего соседа) и «минимальное дерево».

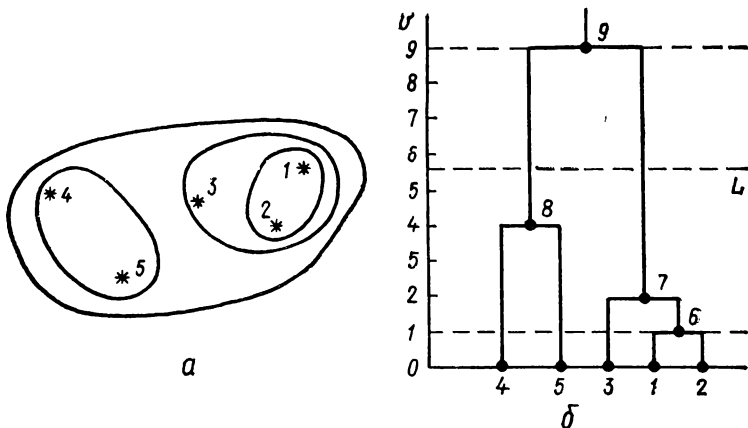


Рис. 16

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | — | 1 | 2 | 9 | 9 |
| 2 | 1 | — | 2 | 9 | 9 |
| 3 | 2 | 2 | — | 9 | 9 |
| 4 | 9 | 9 | 9 | — | 4 |
| 5 | 9 | 9 | 9 | 4 | — |

Рис. 17

Ультраметрика. Со всякой иерархической группировкой объектов можно связать следующую упрощенную меру связи состояний между объектами: $d(x, y)$ равно номеру того уровня иерархии, на котором n и $n - 1$ впервые попадают в один класс. Так, для иерархии, описанной выше, матрица связей имеет вид, представленный на рис. 17.

Мера расстояния $d(x, y)$ обладает следующим свойством: для любых x, y, z выполнено неравенство

$$d(x, y) \leq \max \{d(x, z), d(y, z)\}. \quad (3.20)$$

Это неравенство выражает очевидный факт: если x находится в одном классе с z , причем y также находится в одном классе с z , то x и y принадлежит одному классу.

Вспомним теперь, как определяется обычная метрика:

- 1) $d(x, y) = 0$, если $x = y$;
- 2) $d(x, y) = d(y, x)$ (свойство симметрии);
- 3) $d(x, y) \leq d(x, z) + d(y, z)$ (неравенство треугольника).

Мера расстояния называется ультраметрической, если имеем

$$d(x, y) \leq \max \{d(x, z), d(y, z)\}.$$

Ультраметрика остается ультраметрикой при любом ее монотонном преобразовании, так как при этом меняется упорядочение значений, а значит, выполняется (3.20).

Всякая ультраметрика устроена довольно просто. В силу (3.20) очевидно, что для любых x, y, z все три значения $d(x, y)$, $d(x, z)$, $d(y, z)$ не могут быть различными: максимальное из них все-таки не больше какого-то из оставшихся, так что два обязательно совпадают, а третье не больше их общего значения.

Выше мы с помощью таблицы перешли от индексной иерархии к ультраметрике. С другой стороны, каждой ультраметрической мере близости d соответствует индексная иерархия, у которой d является соответствующим индексом.

Минимальное дерево. Пусть R^p — пространство, содержащее n точек, описанных p признаками. Для каждой пары точек установим меру близости. Соединим каждую пару точек линией, длина которой соответствует индексу близости. Таким образом, набор из n объектов оказывается представленным однозначным графом. Однако если имеем дело с большим количеством объектов, такой способ представления оказывается неприемлемым.

Поэтому следует выделить из этого графа подграфы, обладающие некоторыми полезными свойствами, а именно подграфы, обладающие древовидной структурой. Для нас представляет интерес граф без циклов, имеющий n вершин и $n - 1$ ребро. Длиной дерева является сумма значений индексов.

Наибольший интерес представляет «дерево минимальной длины» ввиду его замечательных описательных свойств, обусловленных связью

с иерархической классификацией. Например, с помощью минимального дерева можно вручную произвести анализ структуры корреляционной матрицы, размеры которой составляют несколько десятков.

В дальнейшем для простоты будем считать, что ребра единственны, а это существенно упрощает описание алгоритмов.

Укажем несколько известных алгоритмов построения минимального дерева.

А л г о р и т м К р у с к а л а (1956). $n(n - 1)/2$ ребер располагают в порядке возрастания значений индекса. Начиная с первых двух ребер отбираются ребра, не образующие цикл с уже отобранными ребрами. Процедура заканчивается, когда окажутся отобранными $n - 1$ ребер. В этом случае получим минимальное дерево.

А л г о р и т м П р и м а (1957). Начиная с любого узла графа. На первом шаге выбирается ближайший объект и присоединяется к исходному. На k -м шаге к сформированным последовательностям ребер V_{k-1} присоединяется самое короткое ребро V_k , которое может быть соединено с одним из узлов V_{k-1} и не образуют цикл. Образованное дерево имеет минимальную длину, так как V_k в любой момент является деревом минимальной длины с k узлами.

А л г о р и т м Ф л о р е к а и д р. (1951). Первоначально строятся отдельные ветви дерева. Для этого начиная с произвольного (ранее не рассматривавшегося) объекта i определяют i_1 , на котором достигается $\min_i d(x_i, x_j)$, затем — i_2 , на котором достигается $\min_i d(x_i, x_j)$, и так до тех пор, пока очередной i_k не совпадет с одним из ранее полученных. Последовательность i, i_1, \dots, i_{k-1} образует искомую ветвь. После получения всех ветвей происходит их «сшивание». Для этого каждая ветвь присоединяется к своему ближайшему соседу; процедура заканчивается, когда будет получен связанный граф. Этот алгоритм считается одним из самых «быстрых» при ручной обработке матриц достаточно больших размерностей.

Минимальное дерево используется не только для непосредственного анализа, но и как основа для дальнейшего применения формальных методов. Так, разрезанные ребра характеризуют структуру связей между классами разбиения.

Связь между минимальным деревом и односвязным методом. Пусть V — минимальное дерево, построенное на основании матрицы расстояний между n объектами. Если V не имеет циклов и является связанным, существует единственный путь, соединяющий две вершины x и y . Обозначим $dv(x, y)$ длину самого длинного ребра на этом пути. Покажем, что $dv(x, y)$ является ультраметрической мерой близости.

Пусть v — самое длинное ребро. Устранение v приводит к разделению V на два отдельных связанных дерева. Если существует путь, связывающий x с y (и не обязательно проходящий через ребра V), у которого наибольшее ребро короче, чем у V , то должно существовать ребро u , отличное от v и короче v , которое соединит образованные два дерева. Это позволит получить дерево меньшей длины, чем V , что противоречит исходной предпосылке. Таким образом, $dv(x, y)$ действительно является наименьшим максимальным скачком.

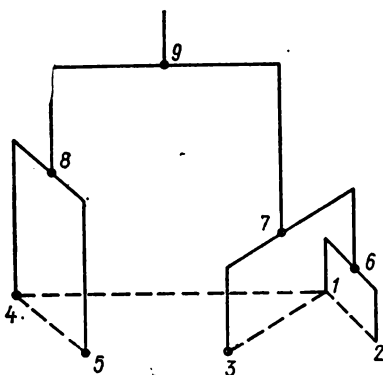


Рис. 18

Этот вывод обеспечивает построение односвязным методом иерархии, начинающейся с минимального дерева.

Для построения иерархии размыкаем самое большое ребро; полученные две группы являются самыми дальними, и индекс, соответствующий их объединению, является длиной этого ребра. Затем размыкаем ребра в порядке уменьшения их длин, спускаясь вдоль иерархии до тех пор, пока не достигнем отдельных элементов. Последнее ребро соответствует первому объединению двух элементов

восходящим алгоритмом. Теперь можно одновременно представить иерархию и минимальное дерево в перспективе (рис. 18).

Таким образом, к классической программе добавляется вспомогательная информация, которая позволит лучше представить относительное положение точек. Бывает особенно полезно при использовании метода главных компонент отразить минимальное дерево на факторную плоскость, чтобы скомпенсировать искажения, связанные с проекцией.

Продолжение численного примера. В результате проведенного выше анализа табл. 1 получены наборы данных (см. табл. 3, 4), которые были использованы для построения индексной иерархической классификации (рис. 19). Верхняя горизонтальная линия определяет уровень объединения в классы. На рис. 19 выделены пять групп, хотя некоторые из них можно объединить (агрегировать), например 1 и 2, 3 и 4. Это как раз и есть один из тех моментов, о которых говорилось выше: за счет визуального анализа принимается решение. Было условлено обрабатывать не всю исходную табл. 1 совместно, а разбив ее на три группы, и обрабатывать каждую группу в отдельности. При этом, как отмечалось ранее, произошло улучшение прогноза.

Табл. 4 была подвергнута двоякой обработке. Первый раз (верхняя часть рис. 20) при классификации учитывались данные столбцов 77—87. Явное разделение на две группы 77—82 и 83—87 объединяется организационными мерами, направленными на улучшение хода процесса.

Второй раз (нижняя часть рис. 20) в процессе классификации в качестве иллюстративной точки включена 88. Она не влияла на ход классификации, но ее «траектория» определялась с учетом значений ее признаков. Как видно, ожидаемое распределение ресурсов на время «88» сходно с распределениями на 86 и 87.

Заметим, что в настоящее время имеется довольно большое число программ, позволяющих выполнить различные операции по АД. Это прежде всего программы, разрабатываемые в лаборатории прикладных методов математической статистики и теории вероятностей ЦЭМИ АН СССР.

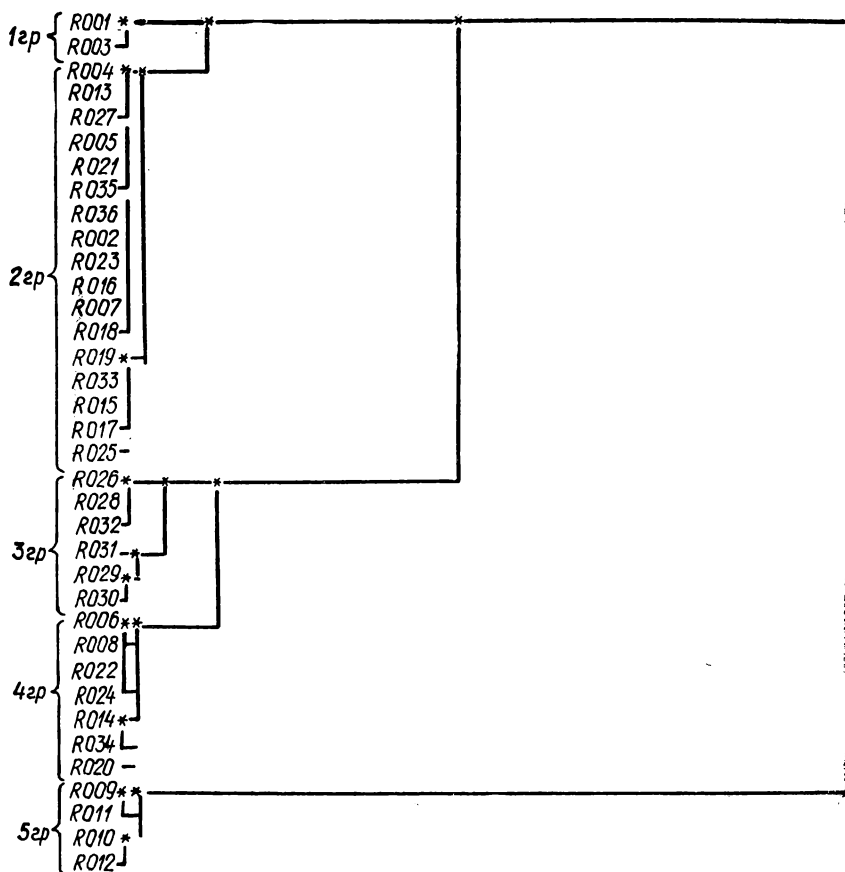


Рис. 19

Методы дискриминантного анализа. Дискриминантный анализ базируется на методах теории решений. Основной задачей этой теории является построение решающей функции (решающего правила) δ , отображающей выборочное пространство X в множество решений D . Для выбора s вводится функция потерь $L(\Theta, d) \geq 0$, которая интерпретируется как убыток при наблюдении величины ξ от принятия решения d при условии, что ξ имеет распределение $F(x, \Theta)$ из параметрического класса распределений $F = \{F(x, \Theta), \Theta \in \Omega\}$. Тогда для любого правила δ можно определить функцию риска $R(\Theta, \delta) = ML(\Theta, \delta(\xi))$. Очевидно, что правило δ_1 лучше правила s_0 , если $R(\Theta, \delta_1) \leq R(\Theta, \delta_0)$ для $\forall \Theta \in \Omega$ и $\exists \Theta^*$ такое, что $R(\Theta^*, \delta_1) < R(\Theta^*, \delta_0)$.

Однако не все правила δ могут быть сравнимы по данному критерию. Поэтому для дальнейшего уточнения принято выполнять два подхода.

При байесовском подходе предполагается, что параметр Θ — случайная величина с плотностью распределения $\Pi(\Theta)$. Тогда наилучшее пра-

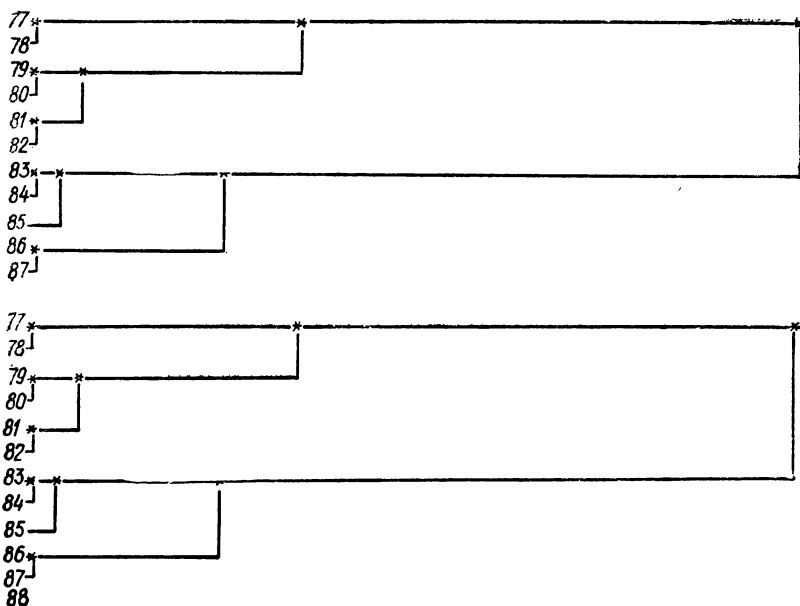


Рис 20

вило s^* выбирается из условия минимизации функционала $\sigma(\delta) = \int R(\Theta, \delta) \pi(\Theta) d\Theta$, называемого байесовским риском.

При минимаксном подходе минимизируется функционал $m(\delta) = \sup_{\Theta \in \Omega} R(\Theta, \delta)$, причем информация о распределении параметра Θ отсутствует.

Дискриминантный анализ как частный случай теории решений изучает конечное множество Ω , т. е. имеется всего L распределений $F_i(x, \Theta) = F(x, \Theta_i)$, $i = \overline{1, L}$. Типичной задачей здесь является задача разбиения выборочного пространства X на L классов $W_i : W_i \cap \cap W_j = \emptyset, i \neq j; \bigcup_{i=1}^L W_i = X$. Можно записать $W_i = \{x : \delta(x) = d_i\}$. Введя функцию потерь $L(\Theta_i d_j) = l(j/i)$, т. е. убыток от отнесения к j -му классу объекта, в действительности принадлежащего i -му классу, запишем функцию риска $R(\delta) = (R_1(\delta), \dots, R_L(\delta))$ и L -мерный вектор, где

$$R_i(\delta) = R(\Theta_i, \delta) = \sum_{j=1}^L l(j/i) p(j/i);$$

здесь $p(j/i)$ — вероятность отнести объект i -го класса к j -му. Байесовским решением в данном случае будет функция δ , минимизирующая функционал $r(\delta) = \sum_{i=1}^L R_i(\delta) \pi_i = \sum_{i=1}^L \pi_i \sum_{j=1}^L l(j/i) p(j/i)$. Минимакс-

ным решением будет функция, минимизирующая функционал

$$m(\delta) = \max_{i=\overline{1,L}} R_i(\delta).$$

Практическое решение задач АД состоит в построении дискриминантной функции $D(y)$, $y \in Y$ — пространства признаков объектов.

В общем случае y является p -мерным вектором. Пусть для всех классов W_i , $i = \overline{1,L}$, нам известны $f_i(y)$ — плотности распределения этих классов в пространстве признаков. Тогда объект x , обладающий признаками y , относят к l -му классу, если

$$D(y) = \frac{\pi_l f_l(y)}{\pi_{l'} f_{l'}(y)} > 1 \quad \forall l' \neq l.$$

В случае неизвестных плотностей $f_i(y)$ применяется другой метод. Дискриминантной функцией называется функция

$$D(y) = a + b_1 y_1 + b_2 y_2 + \dots + b_p y_p, \quad y = (y_1, \dots, y_p) \in Y \subset R^p,$$

причем на разных классах значения $D(y)$ должны различаться как можно больше. Это требование эквивалентно максимизации отношений $T(D)/W(D)$, где T — общая (по всему выборочному пространству X) сумма квадратов, а W — внутригрупповая. Для решения такой задачи можно использовать, например, метод пошагового ДА, в котором из пространства всех признаков выделяется подпространство тех признаков, на которых функция будет удовлетворять упомянутому выше условию.

ЭРГОНОМИЧЕСКИЙ АНАЛИЗ ДЕЯТЕЛЬНОСТИ ОПЕРАТОРА АВТОМАТИЗИРОВАННЫХ СИСТЕМ

§ 1. Задачи, решаемые автоматизированной системой сбора и обработки радиолокационной информации. Роль и место человека-оператора в системе

Автоматизированная система сбора и обработки радиолокационной информации (АССОРЛИ) обеспечивает получение информации о состоянии среды.

Источником информации является ИВК, а носителем информации — выходной сигнал приемника радиолокационной станции (РЛС). Наряду с сигналами, отраженными от воздушного объекта, выходной сигнал содержит шумы приемника, а также помеху. Поэтому АССОРЛИ должна решать следующие задачи:

- 1) выделение отраженного сигнала из выходного сигнала приемника, представляющего собой смесь сигнала и шума;
- 2) определение координат цели по отраженному сигналу;
- 3) распознавание отраженного сигнала (свой, чужой объект) и решение целей (групповая или одиночная).

Эти задачи составляют содержание первичной обработки информации.

Для определения параметров движения цели и повышения достоверности ее обнаружения необходимо обработать информацию. Эта обработка сводится к обнаружению траектории цели (см. [1]).

В условиях маневра цели необходимо непрерывно или с допустимой дискретностью измерять параметры движения, т. е. осуществлять слежение за целями.

Обнаружение траекторий и слежение за ними составляют содержание вторичной обработки информации.

Существующие ИВК обладают значительно меньшей дальностью действия, чем это требуется. Это вынуждает использовать в качестве источников информации систему территориально разнесенных ИВК. Объединение ИВК в систему создает возможность существенно улучшить характеристики радиолокационной информации (дальность обнаружения, точность определения координат и параметров движения и т. д.). Для реализации этих возможностей АССОРЛИ решает задачи сбора и объединения информации, т. е. осуществляет третичную обработку информации [22].

Третичная обработка информации включает: приведение координат цели и параметров её движения к единой системе координат и единому времени; отождествление отметок цели, поступающих от различных ИВК, с одной из трасс; вычисление обобщенных координат цели и параметров её движения.

АССОРЛИ представляют собой иерархические многоуровневые системы, функционирующие в сложных тактических условиях.

В иерархической структуре АССОРЛИ можно выделить три уровня: отдельного ИВК, в частности отдельной РЛС; отдельного узла ИВК, например радиолокационного узла (РЛУ); группировки узлов ИВК, в частности РЛУ.

Для каждого из указанных уровней характерны свои задачи управления, определяемые функциональным назначением уровней АССОРЛИ.

Так, на уровне РЛС решаются следующие задачи управления: оптимальный выбор режима работы РЛС (частота посылки импульсов, период обзора РЛС и т. д.); выбор режима защиты от помех (автоматическая регулировка усиления, включение СДЦ и т. д.); отстройка по частоте от помех; управление энергетическим потенциалом РЛС.

На втором уровне управления АССОРЛИ решаются следующие основные задачи: рациональное использование пропускной способности РЛС узла (зоно- и целераспределение); оптимальное распределение энергии разнотипных радиолокационных средств с целью устранения провалов в зоне выдачи информации РЛУ; выбор оптимального расписания работы радиолокационных средств при централизованном управлении с РЛУ.

Наконец, на третьем уровне управления решаются задачи управления отдельными РЛУ (зоно-, целераспределение, разгрузка узлов и т. д.).

АССОРЛИ являются системами, имеющими сложную структуру с большим количеством комплектующих элементов, связанных между собой. При исследовании эффективности АССОРЛИ необходимо всегда учитывать надежность их функционирования. Известно, что надежность системы поддерживается правильной организацией технического обслуживания, обеспечиваемого обслуживающим персоналом.

По степени участия человека-оператора в сборе и обработке радиолокационной информации системы можно условно разделить на неавтоматизированные, автоматизированные и автоматические.

В неавтоматизированных системах все задачи, о которых говорилось выше, решаются операторами визуально с использованием индикаторов и планшетов.

В автоматических системах задачи решаются с помощью автоматических устройств сбора и обработки информации. Функции человека при этом сводятся к контролю за работой системы и её техническому обслуживанию.

Однако наибольшее распространение в настоящее время получили не автоматические, а автоматизированные системы сбора и обработки информации, в которых человек является неотъемлемой частью системы. В таких системах наиболее рационально используются преимущества человека и ЭВМ при выполнении отдельных операций сбора и обработки. Человек-оператор способен качественно оценить ситуацию и определить возможные варианты решения. ЭВМ компенсирует ограниченные возможности человека быстро количественно оценить тот или иной вариант решения.

Такой подход к проектированию позволяет получить систему, обладающую достаточно хорошими характеристиками радиолокационной информации при приемлемой стоимости системы.

Предметом изучения данной и следующей глав будет деятельность человека-оператора в автоматизированных системах сбора и обработки информации [5, 81].

Исходя из приведенного выше описания задач, решаемых любой системой сбора и обработки информации, можно выделить три вида деятельности оператора в таких системах:

- 1) участие в процессе сбора и обработки радиолокационной информации;
- 2) участие в управлении системой;
- 3) техническое обслуживание системы.

Рассмотрим более подробно деятельность человека-оператора в автоматизированной системе сбора и обработки радиолокационной информации [43]. С этой целью остановимся на первичной, вторичной и третичной обработках информации и на функции человека-оператора при этом.

Задачи первичной обработки информации решаются устройствами первичной обработки (УПО).

В автоматизированных УПО (наиболее распространенных) обычно автоматизируется решение задачи определения координат цели и их кодирования. Задача обнаружения отметки цели решается визуально с помощью устройства съема и кодирования координат. Как правило, съём координат с экрана индикатора при автоматизированной обработке производит оператор.

Автоматизированная обработка повышает темп и точность съёма координат целей по сравнению с неавтоматизированной обработкой.

При автоматической обработке все задачи решают автоматические УПО. Функции человека при этом сводятся к контролю за работой системы.

Поскольку преимущества человека по сравнению с машиной при обнаружении сигнала на фоне шумов и помех неоспоримы, в современных и перспективных радиолокационных системах обнаружения целей и съём их координат осуществляет оператор [15].

Все операции процесса вторичной обработки принципиально могут быть формализованы и выполнены с помощью ЭВМ. В этом случае устройство обработки автоматическое.

При автоматизированной вторичной обработке операции обнаружения траекторий и сличения координат экстраполированной точки с координатами вновь поступивших отметок и выбор одной из отметок для продолжения траектории выполняются, как правило, оператором.

Автоматизированная система сопровождения целей изображена на рис. 21.

Автоматизированная система сопровождения включает устройство отображения совмещенной (первичной и вторичной) информации (УО), оператора (ОП), устройство съема координат (УСК) и полуавтоматическое устройство вторичной обработки (УВО).

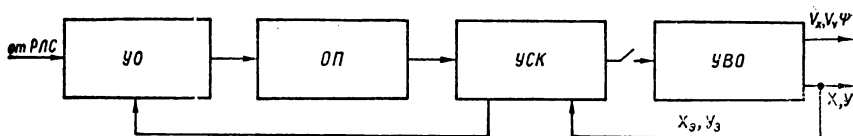


Рис. 21

В полуавтоматическом УВО оператор после обнаружения отдельных отметок на экране УО в процессе первичной обработки самостоятельно или по данным целеуказания выбирает те из них, которые принадлежат одной цели. После этого оператор с помощью УСК вводит в экстраполятор цели в двух периодах обзора пространства и посылает в УВО команду, по которой начинается дальнейшая обработка этой группы отметок, т. е. слежение за выбранной траекторией. Координаты цели при втором вводе могут быть сняты оператором либо в следующем после первого ввода обзоре, либо через несколько обзоров, что определяется такими факторами, как устойчивость обнаружения цели РЛС от обзора к обзору, загрузка оператора функциями обработки информации и т. д.

После обнаружения оператором трассы цели экстраполятор рассчитывает параметры движения цели и её упрежденные координаты.

Параметры движения и экстраполированные координаты являются выходными данными УВО. Кроме того, экстраполированные координаты используются для перемещения маркера (экстраполированной точки) на экране индикатора. Траектория движения маркера должна отображать с некоторой точностью траекторию движения цели. При слежении за целью в момент получения новой отметки, принадлежащей траектории данной цели, оператор оценивает разность между наблюдаемой отметкой и положением маркера, сопровождающего эту цель. Если рассогласование превышает допустимую величину, оператор вводит корректуру по координатам, совмещая маркер с отметкой от цели.

В полуавтоматических УВО оператор выполняет роль дискриминатора ошибок. Характеристика дискриминатора системы сопровождения нелинейна, так как существует зона нечувствительности, в пределах которой корректура не вводится. Размеры этой зоны определяются разрешающей способностью УО и индивидуальными особенностями оператора [15].

Для улучшения условий сопровождения целей и качества сопровождения УВО регулярно, через несколько обзоров, РЛС может выдавать сигнал на УО, свидетельствующий о необходимости ввода корректуры оператором.

Точность полуавтоматического сопровождения целей определяется многими факторами, основными из которых являются психофизиологические характеристики оператора, уровень его подготовки, количество одновременно сопровождаемых целей (загрузка оператора), скорость движения, интенсивность маневра, устойчивость обнаружения цели РЛС от обзора к обзору и т. п.

Автоматические УВО реализуются на базе ЭВМ и обеспечивают автоматическое сопровождение целей. Как правило, в существующих и перспективных АСУ автоматический способ обработки используется в комбинации с полуавтоматическим. Автоматическое сопровождение цели начинается только после обнаружения трассы цели оператором и наблюдения за трассой в течение некоторого времени. Для перехода в режим автоматического сопровождения должны соблюдаться условия устойчивой выдачи информации от обзора к обзору, отсутствие резких маневров цели, помех, отражений от местных предметов и т. д. Такой комбинированный метод вторичной обработки информации позволяет, во-первых, рационально использовать преимущества оператора при обнаружении сигналов цели на фоне шумов и помех, во-вторых, более рационально использовать ресурсы ЭВМ.

Назначение третичной обработки — собрать информацию от нескольких источников и объединить ее в общую картину.

Здесь также имеется принципиальная возможность формализации задач третичной обработки. В этом случае процесс сбора и объединения информации автоматический и реализуется с помощью ЭВМ. Автоматический метод решения рассматриваемой задачи обладает такими качествами, как большая точность решения задачи и высокая вероятность правильного отождествления сообщений.

В современных системах также распространен автоматизированный способ решения задачи сбора и объединения информации с участием оператора [15].

При автоматизированном способе вычислительные и логические операции (сбор сообщений, их приведение к единой системе координат и единому времени, сглаживание координат) выполняются, как правило, ЭВМ. Оператор реализует функции отождествления сообщений и группирования целей.

Основными операциями, выполняемыми человеком при отождествлении, является классификация признаков одиночных сообщений и распознавание принадлежности сообщений одной цели по совокупности признаков. Классификационными признаками сообщений могут быть координаты целей и их признаки (количественный состав, тип цели и т. д.). Иногда на устройствах отображения имеется информация о курсе скорости движения целей. Использование этой информации как классификационных признаков существенно улучшает качество решения задачи отождествления.

После классификации признаков одиночных сообщений оператор осуществляет распознавание принадлежности сообщений одной цели (отождествление). Общая схема решения задачи распознавания оператором при этом сводится к следующему.

Допустим, что при распознавании можно использовать N классификационных признаков сообщений. Тогда каждое сообщение можно представить в N -мерном пространстве определенной точкой. В таком представлении совокупности сообщений, относящихся к одной и той же цели, будет соответствовать ансамбль точек, рассеянных внутри некоторой области пространства. Очевидно, что точки, формирующие изображения одного класса, будут группироваться внутри некоторой

ограниченной области N -мерного пространства. Если расстояние между точками внутри данного класса меньше, чем расстояние между любой парой точек, принадлежащих разным классам, то задача распознавания может быть решена практически безошибочно.

В реальных условиях возможны ошибки восприятия и классификации признаков (например, оператор определяет с ошибкой расстояние между сообщениями). Поэтому вероятность правильного решения задачи отождествления оператором имеет некоторую величину, меньшую единицы. Например, экспериментально установлено, что при использовании в качестве классификационных признаков, кроме координат объекта и его характеристик курса, скорости цели вероятность правильного отождествления составляет величину, равную 0,85—0,9.

Рассмотрим особенности выполнения операции группирования целей человеком. Операция выполняется в тех радиолокационных системах, где не требуется информации по каждой отдельной цели или же плотность потока целей оказывается выше пропускной способности системы.

В принципе группирование осуществляется теми же способами, что и отождествление сообщений, и ведется по признаку близости классификационных признаков целей. Отличия здесь заключаются в том, что, во-первых, степень близости устанавливается самим оператором и определяет уровень группирования; во-вторых, очень важным классификационным признаком при группировании становится такой тактический признак, как принадлежность целей одной группе, выполняющей определенную тактическую задачу. Эти особенности группирования позволяют использовать преимущества человека по сравнению с ЭВМ при распознавании тактических ситуаций [23].

§ 2. Основные задачи и принципы эргономического анализа

Объектом изучения в эргономике является единая система среда — человек — машина. Конечная цель эргономики — создание высокоэффективных систем среда — человек — машина на основе рационального использования возможностей человека и техники в процессе их проектирования, разработки и эксплуатации (рис. 22).

Эргономика призвана решать две группы взаимосвязанных задач, которые в общем виде могут быть определены следующим образом.

Как техническая наука эргономика изучает посты и пульты управления, процессы и алгоритмы функционирования технических устройств, чтобы выяснить, какие требования они предъявляют к психофизиологическим особенностям человека-оператора, т. е. решает задачу приспособления человека к технике и условиям труда.

Как психологическая наука эргономика изучает [87] психические процессы и свойства человека, выясняя, какие требования к техническим устройствам вытекают из особенностей человеческого организма, т. е. решает задачу приспособления техники и условий труда к человеку.

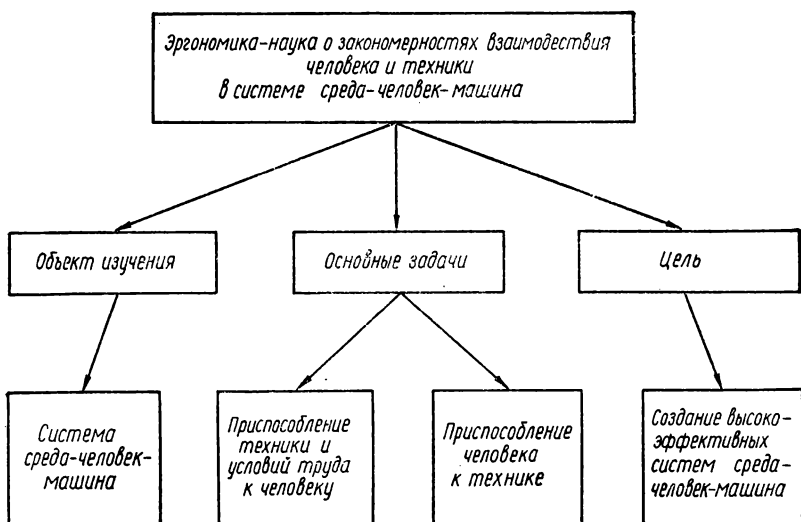


Рис. 22

Деятельность специалистов по эргономике простирается от ранних стадий разработки системы человек — машина до ее анализа на испытаниях и в процессе эксплуатации серийных образцов. Однако, несмотря на некоторую общность задач и методов эргономического исследования на этапах разработки системы и ее эксплуатации, эти этапы существенно отличаются друг от друга.

На этапе эргономического проектирования основное внимание уделяется решению задачи приспособления техники к человеку. Задача приспособления человека к технике решается лишь ориентировочно (рис. 23, штрихом обозначены задачи проектирования, решаемые ориентировочно). Последнее объясняется тем, что на этапе проектирования системы и ее предварительной эргономической оценки конструктор может исследовать ее поведение только в лабораторных условиях, далёких, как правило, от условий её применения. При этом многие особенности функционирования системы остаются неясными, в результате чего трудно определить полный перечень задач, решаемых оператором, их характеристики, достоверно оценить характеристики деятельности оператора и эффективность системы человек — машина.

На этапе эргономической оценки системы в процессе ее испытаний и эксплуатации центр тяжести переносится на решение задачи приспособления человека к технике. Возможности внесения изменений в конструкцию и алгоритмы функционирования системы человек — машина ограничены, так как эти изменения обычно связаны со значительными затратами времени и средств. Таким образом, повышение эффективности системы может быть достигнуто лишь в результате тщательного изучения деятельности оператора и определения путей оптимизации этой деятельности с учётом особенностей системы.

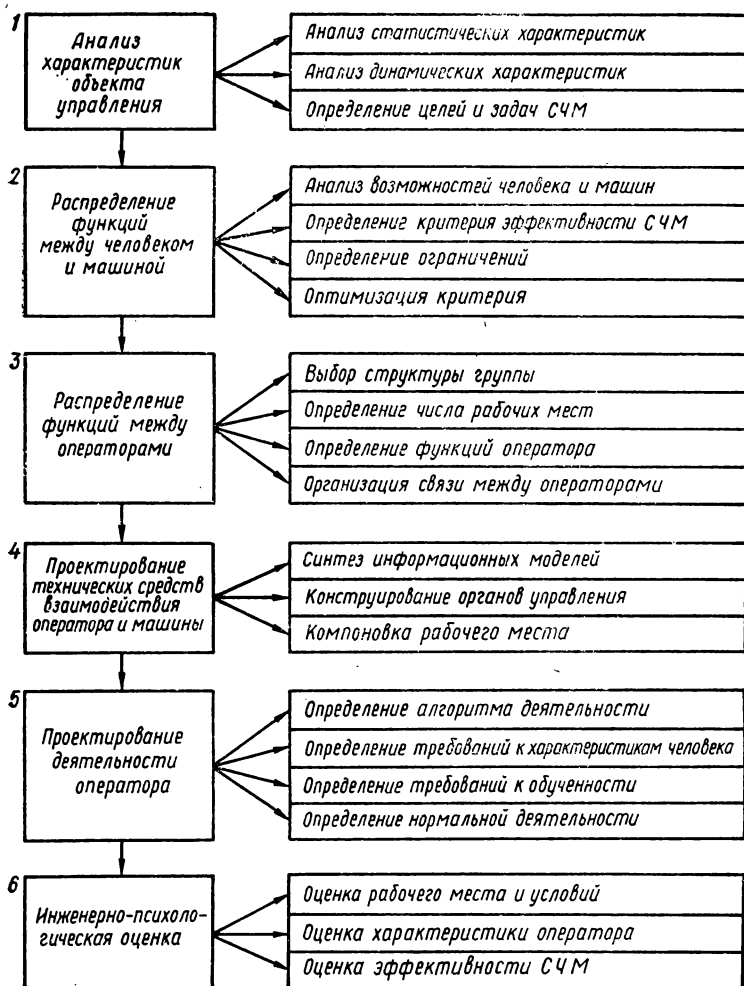


Рис. 23

Эргономическая оценка эксплуатируемых систем типа человек — машина производится в соответствии со схемой, изображенной на рис. 24. Кратко охарактеризуем основные этапы оценки.

1. Знакомство с системой. Устанавливается назначение системы. По техническим описаниям изучаются принципы ее функционирования, по инструкциям по эксплуатации — особенности обслуживания и эксплуатации.

2. Изучение структурной схемы системы. Разрабатывается структурная схема системы, отражающая связи отдельных подсистем, потоки информации и протекание процесса управления. Изучаются принципы распределения функций между машиной

и оператором. Указываются прямые и обратные связи между операторами и машиной. Определяются интенсивность связей (число связей в единицу времени) и их относительная важность.

3. Исследование влияния среды. Изучается действие природных факторов, а также различного рода «помех», оказывающих влияние на функционирование системы и работу операторов. Описываются возможные изменения внешней обстановки, приводящие к развитию экстремальных ситуаций.

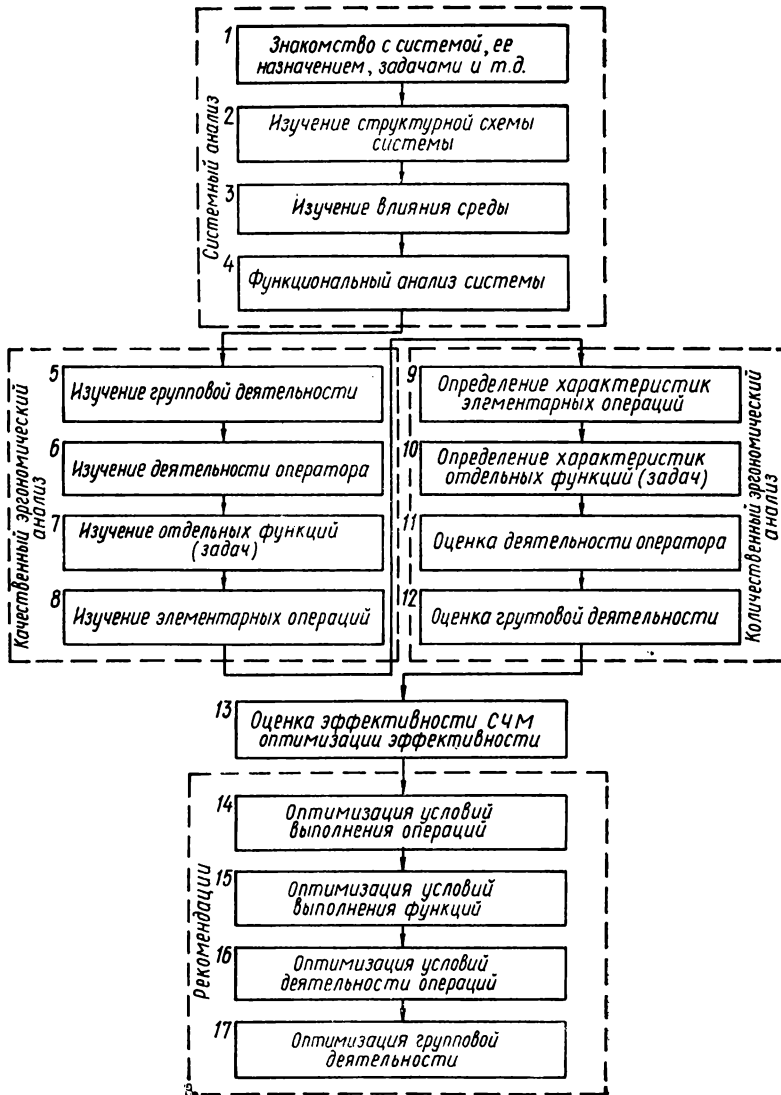


Рис. 24

4. **Функциональный анализ системы.** Изучаются функции системы и её подсистем при всех режимах работы (режимы: применения, подготовки к применению, восстановления).

5. **Исследование групповой деятельности операторов.** Изучаются структура группы, распределение функций между операторами, характер связей между ними, интенсивность связей и их важность.

6. **Изучение деятельности оператора.** Определяются функции оператора как для нормальных режимов работы, так и для различных экстремальных (конфликтных) ситуаций, строится профессиограмма деятельности оператора.

7. **Изучение отдельных функций оператора.** Исследуется каждая функция (задача), выполняемая оператором. Составляется алгоритм ее выполнения в виде последовательности операций.

8. **Исследование элементарных операций.** Изучается каждая элементарная операция, входящая в алгоритм выполнения функции. Определяется ее психологическое содержание.

Этапы 5—8 являются этапами качественного исследования деятельности операторов. В результате их выполнения подготавливаются необходимые сведения для количественной эргономической оценки системы человек — техника.

9. **Определение количественных характеристик элементарных операций.** Экспериментальным путём определяются количественные характеристики элементарных операций: время выполнения, точность, надёжность.

10. **Определение количественных характеристик задач.** Экспериментальным (или расчетным) путем определяются количественные характеристики задач (важность, интенсивность, время существования) и характеристики их решения оператором (время решения, точность, надежность).

11. **Оценка деятельности оператора.** Экспериментальным (или расчётным) путем определяются характеристики деятельности оператора (качество решения всех задач, производительность), оценивается режим работы (напряженность).

12. **Оценка групповой деятельности операторов.** Экспериментальным (или расчётным) путем определяются характеристики групповой деятельности операторов (качество решения задач с учётом взаимодействия операторов, производительность группы операторов).

13. **Оценка эффективности системы человек — машина.** Оценивается эффективность системы с учетом деятельности группы операторов, определяется степень влияния их деятельности на эффективность системы.

Этапы 9—13 количественной эргономической оценки являются основой для выработки рекомендаций, обеспечивающих повышение эффективности системы человек — машина, улучшение условий деятельности операторов.

14. **Оптимизация условий выполнения операций.** Анализируются количественные характеристики выполнения операций. Определяются операции, имеющие неудовлетворительные характеристики, изыскиваются возможности по улучшению характеристик, даются рекомендации по улучшению общей компоновки рабочего места, совершенствованию информационной модели, органов управления. Определяются требования к характеристикам человека.

15. **Оптимизация условий выполнения функций.** Анализируются количественные характеристики выполнения функций (задач). Определяются задачи, имеющие неудовлетворительные характеристики. Намечаются пути улучшения характеристик задач, во-первых, за счёт приспособления человека к технике (оптимизация алгоритмов решения задач), во-вторых, за счёт приспособления техники к человеку (компоновка рабочего места, улучшение информационной модели, органов управления). Определяются требования к обученности оператора при решении задач (нормативы).

16. **Оптимизация условий деятельности оператора.** Анализируются количественные характеристики работы оператора. Определяются возможности их улучшения за счёт приспособления человека к технике (дисциплина обслуживания задач и т. д.), а также даются рекомендации по совершенствованию рабочих мест.

17. **Оптимизация условий групповой деятельности операторов.** Анализируются количественные характеристики групповой деятельности. Решается задача оптимизации групповой деятельности за счёт выбора рациональной структуры группы, распределения функций между операторами и определения их количества. Разрабатываются рекомендации по улучшению условий групповой деятельности: уточняется необходимое количество автоматизированных рабочих мест, даются предложения по улучшению средств отображения коллективного и группового пользования, организации связи между операторами и т. д.

18. **Оптимизация критерия эффективности системы.** Анализируется достигнутый уровень эффективности системы, устанавливается возможность его повышения за счёт уточнения варианта распределения функций между операторами и машиной. Определяются задачи, которые желательно автоматизировать; даются рекомендации по улучшению условий взаимодействия операторов и машины.

Таково содержание эргономической оценки системы человек — машина в процессе ее испытаний и эксплуатации. Общая схема эргономической оценки позволяет наметить основные методические вопросы, которые необходимо решить для ее проведения. В основу такой оценки должен быть положен системный подход, который предполагает наличие следующих основных этапов исследования: определение цели исследования; выбор показателя эффективности; выбор количественных методов исследования; определение факторов, влияющих на эффективность и их количественное описание; установление коли-

чественных отношений между показателем эффективности и влияющими на него факторами; оптимизация показателя эффективности.

Таким образом, реализация системного подхода при проведении эргономической оценки требует рассмотрения показателей качества деятельности оператора в их связи с показателями эффективности АССОРЛИ, а также разработки методов количественной и качественной оценок деятельности оператора.

Наиболее характерной чертой деятельности оператора АССОРЛИ является то, что он лишен возможности непосредственно наблюдать за управляемыми объектами и вынужден пользоваться информацией, которая поступает к нему по каналам связи. Деятельность человека, совершаемая не с реальными объектами, а с их заменителями или имитирующими их образцами, называют деятельностью с информационными моделями реальных объектов.

Информационная модель — совокупность информации о состоянии и функционировании объекта управления и внешней среды. Объём информации, включенной в модель, и правила её организации должны соответствовать задачам и способам управления. Физически информационная модель реализуется с помощью устройств отображения информации.

Рассмотрим структуру деятельности оператора при его работе с информационной моделью в процессе решения некоторой произвольной задачи.

Первый этап — восприятие информации — процесс, включающий обнаружение объекта восприятия, выделение в объекте отдельных признаков, отвечающих стоящей перед оператором задаче, ознакомление с выделенными признаками и опознавание объекта восприятия.

В процессе ознакомления с выделенными признаками оператор устанавливает связи между ними, формирует системы эталонов, на основании которых он может впоследствии опознать объект или ситуацию. Процессам ознакомления и опознавания сопутствуют обычно укрупнение признаков, объединение их в структуры, которые затем выступают как единые оперативные единицы восприятия. Формирование этих единиц обеспечивает целостность и предметность восприятия, а также возможность в дальнейшем мысленного реконструирования особенностей объекта, не нашедших отражения в информационной модели.

Второй этап — оценка информации, её анализ и обобщение на основе заранее заданных или сформулированных критериев оценки. Оценка производится на основе сопоставления воспринятой информационной модели со сложившейся у оператора внутренней моделью обстановки, которая представляет собой продукт осмысливания оператором сложившейся ситуации с учётом стоящих перед ним задач.

В содержание внутренней модели оператора входят образцы и модели реальной и прогнозируемой обстановки, совокупность возможных управляющих воздействий и исполнительных реакций системы. В модель могут входить также такие субъективные элементы, как представление о целях и критериях функционирования системы, мотивы деятельности и т. д.

Завершающей операцией переработки информации на втором этапе является синтез, т. е. организация полученных после декодирования сведений в целостную систему взаимосвязанных характеристик.

Третий этап — принятие решения о действиях — выполняется на основе проведенного анализа информационной и внутренней модели обстановки.

В ряде случаев определяется заранее заданный и известный оператору алгоритм решения. Тогда основой взаимодействия оператора с информационной моделью является, по существу, выбор наилучшего из имеющихся в его распоряжении средств.

Деятельность оператора с информационной моделью усложняется, если ситуация не предусмотрена заданным алгоритмом решения. В этом случае это взаимодействие преследует уже две цели — постановку самой задачи и нахождение способа ее решения, что достигается посредством оперативного мышления.

Четвертый этап — приведение принятого решения в исполнение посредством определенного действия (системы действий) или отдачи соответствующих распоряжений.

Пятый этап — контроль за результативностью исполнения принятого решения.

После завершения этого этапа оператор приступает к решению другой возникшей задачи. Если оператор решает одновременно несколько задач, то этапы могут быть сдвинуты относительно друг друга.

Первые два этапа называют информационным поиском, последние три — обслуживанием.

В реальной работе оператора перечисленные этапы не обязательно представлены полностью и в указанной последовательности. Нередко они настолько переплетены, что трудно выделить каждый из них в отдельности. В некоторых видах деятельности оператора тот или иной этап оказывается сокращенным, в то время как другой сильно разрастается и начинает составлять главное содержание действия.

Деятельность оператора имеет иерархический характер. Иерархия заключается в последовательном выполнении человеком ряда действий разных порядков (уровней). Действием является поведенческий акт, имеющий осознанную цель, а совокупность действий, объединенных единым мотивом, представляет собой деятельность. Относительно стабильные компоненты действия называют операциями. У операции всегда имеются четко выраженные критерии выполнения, начало и конец. Очень часто к действиям оператора применяют термин «задача».

Таким образом, для количественного исследования деятельности оператора необходимо установить критерии выполнения операций, задач и, наконец, всей деятельности оператора. Остановимся на этих критериях [43].

Критериями выполнения операций является время выполнения, надёжность, точность.

Время выполнения операции $\tau_{оп}$ является случайной величиной и характеризуется законом распределения $\varphi(\tau_{оп})$ или его числовыми характеристиками: математическим ожиданием $\bar{\tau}_{оп}$ и среднеквадратическим отклонением $\sigma_{оп}$.

Надёжность выполнения операций $P_{\text{оп}}$ определяется вероятностью их правильного выполнения. Для оценки этой вероятности используется величина

$$P_{\text{оп}} = m/n, \quad (2.1)$$

где m — число правильно выполненных операций; n — общее число выполненных операций.

Ошибку оператора можно определить как действие (или бездействие) оператора, которое приводит к нарушению нормального функционирования системы или снижает эффективность её применения.

Ошибки оператора делятся на следующие: сенсорные ошибки, заключающиеся в неправильном восприятии информации; ошибки переработки информации и оценки ситуаций; ошибки в принятии решения и его реализации.

Точность работы оператора характеризуется законом распределения $\eta(x)$ и его числовыми характеристиками: математическим ожиданием m_x и среднеквадратическим отклонением σ_x . Величина m_x определяет систематическую ошибку оператора, а величина σ_x — случайную.

Результаты экспериментальных исследований показывают, что, как правило, распределение ошибок можно аппроксимировать нормальным законом.

При изучении характеристик задач выделяют две группы характеристик: собственно характеристики задач и характеристики качества их решения оператором [43].

К характеристикам i -й задачи относят важность задачи C_i , интенсивность её появления λ_i , время существования задачи $t_i^{\text{доп}}$.

Важность i -й задачи C_i определяется степенью её влияния на эффективность системы. Практически возможным методом определения величины C_i является метод экспертных оценок. Экспертами могут быть операторы, работающие на системе и имеющие достаточно большой опыт.

Под интенсивностью появления i -й задачи λ_i понимается количество задач, возникающих в единицу времени. Величина определяется, с одной стороны, условиями функционирования системы, а с другой, — особенностями самой системы (уровнем автоматизации, методами решения задач и т. д.).

Величина, обратная интенсивности λ_i , определяет среднее время между моментами возникновения задачи Δt_i . Исчерпывающей характеристикой величины Δt_i является закон распределения $\Phi_i(\Delta t)$, который часто принимают экспоненциальным, т. е.

$$\Phi_i(\Delta t) = 1 - \exp(-\lambda_i \Delta t). \quad (2.2)$$

Под временем существования i -й задачи $t_i^{\text{доп}}$ понимается время от момента реального возникновения задачи (момента возникновения ситуации, требующей вмешательства оператора в процессе управления) до момента ее исчезновения (момента, когда решение задачи становится практически нецелесообразным). Величина $t_i^{\text{доп}}$ определяется промежутком времени, в течение которого задержка в решении задачи не приведет к каким-либо последствиям, нежелательным для системы

(аварии, существенному снижению эффективности применения системы и т. д.). Это время иначе называют допустимым временем решения задачи оператором. Оно характеризуется функцией распределения $G_i(t)$ и ее числовыми характеристиками $\bar{t}_i^{\text{доп}}$ и $\sigma_i^{\text{доп}}$. Очень часто принимают закон распределения равномерным, т. е.

$$G_i(t) = (t_i^{\text{доп}} - t_{i\min}^{\text{доп}}) / (t_{i\max}^{\text{доп}} - t_{i\min}^{\text{доп}}). \quad (2.3)$$

К характеристикам качества решения i -й задачи оператором относятся время ее решения τ_i и надежность решения P_i .

Время решения τ_i характеризуется законом распределения $B_i(\tau)$ и его числовыми характеристиками $\bar{\tau}_i$ и σ_i .

При расчетах довольно часто принимают закон распределения Эрланга k -го порядка, который описывается выражением

$$dB_i(\tau) = \frac{v_i (v_i \tau)^{k-1}}{k!} \exp(-v_i \tau) (\tau \geq 0, k \geq 1), \quad (2.4)$$

где k — порядок закона Эрланга; $v_i = 1/\bar{\tau}_i$ — интенсивность решения i -й задачи.

Надежность решения i -й задачи P_i определяется как вероятность безошибочного решения и рассчитывается аналогично (2.1).

Наконец, деятельность оператора в целом характеризуется такими показателями, как производительность Π и напряженность ρ .

Под производительностью оператора Π обычно понимают его возможности по управлению объектами, обслуживанию аппаратуры, сопровождению целей и т. д. при надёжности работы P (вероятности безошибочного решения задач) не ниже заданной.

Функциональное напряжение организма при трудовой деятельности определяется двумя сторонами деятельности: энергетической и информационной [26]. Обе они имеют место при всех видах трудовой деятельности, но в зависимости от того, какая сторона преобладает, труд относят к физическому или умственному. Операторская деятельность является, как правило, умственной и оценивается напряженностью этой деятельности.

По своему характеру напряженность деятельности оператора может быть двух видов: операционной и эмоциональной.

Операционная напряженность возникает в результате сложности выполняемой работы (большая информационная нагрузка, трудности различения сигналов и т. д.), эмоциональная при воздействии на человека раздражителей. Последний вид напряженности развивается в результате появления у человека отрицательных эмоций. Предметом нашего изучения в дальнейшем будет операционная напряженность.

Необходимо отметить, что рассмотренные характеристики деятельности оператора в значительной степени зависят от способа его деятельности, т. е. от его навыков и умения, а также от мотивов его поведения и тех условий, в которых он работает.

Очень важным условием возможности изучения и оптимизации деятельности оператора является наличие методов, позволяющих определить характеристики более высокого уровня по известным характеристикам более низкого уровня.

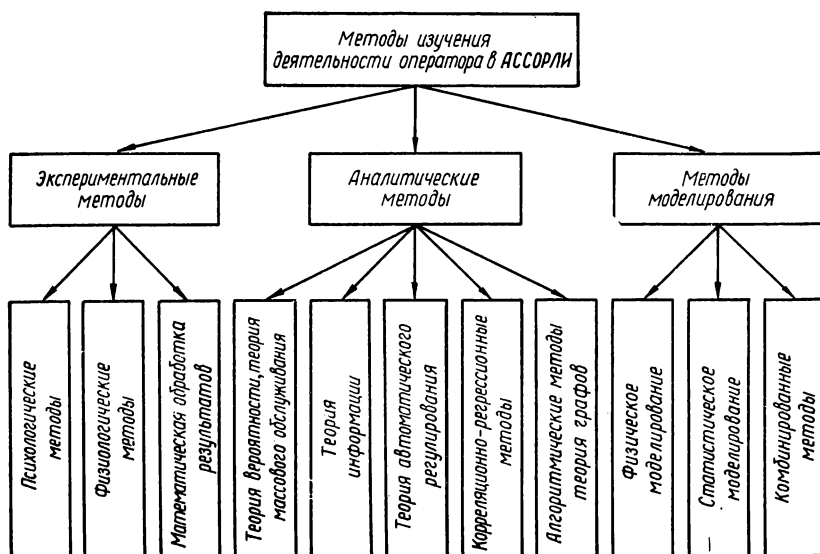


Рис. 25

Сделаем ряд замечаний относительно характеристик методов эргономического исследования.

По своему характеру и особенностям методы изучения деятельности оператора, применяемые в эргономике, делятся на экспериментальные, аналитические и моделирования. Классификация этих методов представлена на рис. 25.

Специфика проведения эргономической оценки системы в процессе её испытаний или эксплуатации определяет особую важность экспериментального метода как наиболее достоверного метода исследования. Общая схема проведения такого исследования соответствует схеме, приведенной на рис. 24. Основой экспериментального эргономического исследования АССОРЛИ являются статистические данные, получаемые в процессе проведения натурных (летных) экспериментов. Экспериментальный метод включает в качестве составляющих методы психологические, физиологические и обработки экспериментальных данных.

С помощью психологических методов осуществляется психологический анализ деятельности оператора в реальных условиях, оценки влияния на его деятельность различного рода факторов.

Физиологическими методами изучается физиологическое состояние оператора в ходе деятельности, определяется реакция различных систем организма на выполнение данной работы (см. гл. IV).

Основные характеристики деятельности оператора (время решения задач, число ошибок, физиологические показатели и др.) являются случайными величинами, изменяющимися от опыта к опыту. Случайность этих характеристик обусловлена влиянием на них огромного числа факторов объективного характера. Поэтому обработка результа-

тов эксперимента ведется методами математической статистики. Задачами статистической обработки результатов эксперимента являются оценка параметров искомых характеристик, сравнение этих параметров, построение эмпирических зависимостей.

Аналитические методы и методы моделирования составляют группу математических методов. Математические методы находят широкое применение на стадии проектирования системы, а также дополняют экспериментальный метод при проведении испытаний и в период эксплуатации системы (особенно для тех условий, где возможности проведения эксперимента ограничены).

Особенностью математических методов является то, что в них исследуется не сама деятельность оператора, а некоторые искусственно созданные объекты (модели), соотносимые с этой деятельностью [43]. Моделирование деятельности оператора заключается в исследовании этой деятельности на моделях и переносе результатов на реальную деятельность.

При математическом моделировании деятельность оператора исследуется с помощью математических моделей. Исследование заключается в выводе формул, решении уравнений и т. д. В основу построения математических моделей могут быть положены различные математические методы. Наибольшее распространение в эргономике находят методы теорий: массового обслуживания, вероятностей, автоматического регулирования, информации и т. д.

При физическом моделировании исследуется деятельность оператора как бы в лабораторных условиях с помощью специального, имитирующего реальную систему, оборудования (тренажеры, макеты, моделирующие стенды и установки и т. д.). Смысл моделирования заключается в воспроизведении психологической структуры и особенностей реальной деятельности, а также в математическом планировании и обработке результатов эксперимента.

Под статистическим моделированием понимается имитация деятельности оператора при помощи ЭВМ с учётом воздействия случайных факторов, обусловленных влиянием условий деятельности, процесса управления и самого оператора. Метод позволяет получить удовлетворительную точность определения характеристик деятельности оператора и может быть применён на любой стадии проектирования системы.

Методы математического моделирования являются новыми для исследования деятельности оператора, и их применение обусловлено потребностями эргономического проектирования и анализа в количественных характеристиках деятельности оператора.

§ 3. Аналитические методы исследования деятельности оператора

Методы теории автоматического регулирования. Количественные методы исследования деятельности человека-оператора, использующие аппарат теории автоматического регулирования, занимают ведущее место среди других методов [49, 51]. Это объясняется, во-первых, широким использованием человека-оператора в следящих

системах, во-вторых, достаточно хорошей разработкой методов исследования сложных динамических систем.

По аналогии с теорией автоматического регулирования оператора представляют динамическим звеном, имеющим сенсорный вход и моторный выход. Важнейшей характеристикой динамического звена является передаточная функция $G(t)$, связывающая выходную $U_{\text{вых}}(t)$ и входную $U_{\text{вх}}(t)$ величины:

$$G(t) = U_{\text{вых}}(t)/U_{\text{вх}}(t). \quad (3.1)$$

Получение передаточной функции производится в предположении, что человек — это одноканальная замкнутая система с обратными связями, в которой различаются следующие временные интервалы: T_1 — время анализа (составляет примерно 0,3 с); T_2 — время формирования выходной реакции (около 0,2 с); T_3 — время восприятия входных возмущений.

В качестве обобщенного варианта линейной непрерывной модели деятельности оператора многие авторы считают возможным использовать передаточные функции вида

$$W(p) = K_{\text{оп}} M(p)/N(p) \exp(\tau_p), \quad (3.2)$$

где $K_{\text{оп}}$ — коэффициент усиления оператора; τ_p — время запаздывания реакции оператора; $M(p)$, $N(p)$ — полиномы, порядок и коэффициенты которых зависят от типа и конкретных условий решения задачи и могут изменяться в широких пределах.

Полином $M(p)$ в общем случае имеет порядок от нулевого до второго и позволяет отразить способности оператора к учету первой и второй производных входного сигнала.

Полином $N(p)$ может иметь порядок от нулевого до четвертого. В его структуре должны быть отражены характеристики нервно-мышечной системы человека, свойства фильтрации высокочастотных составляющих входного сигнала, параметры целевой функции управления, а также другие факторы, характеризующие специфику решаемой задачи.

Однако в силу ряда экспериментально установленных факторов модель (3.2) не адекватна деятельности оператора в реальных условиях. Прежде всего следует указать на наличие некоторой зоны нечувствительности в сенсорной сфере оператора. Помимо того, спектральный анализ выходного сигнала оператора обнаруживает частоты, которые не содержались во входном сигнале, что свидетельствует о нелинейности преобразования сигнала оператором. Далее, линейная непрерывная модель не может объяснить экспериментально наблюдаемых явлений интерполяции и экстраполяции реакции оператора при временном исчезновении входного сигнала. Постоянство параметров модели не позволяет ставить вопросы об учете таких важнейших характеристик человека, как утомляемость, адаптация к объекту и внешним условиям, степень обученности и тренированности и т. д.

Для устранения указанных недостатков математических моделей оператора при описании его деятельности применяются дискретные импульсные модели.

Структурная схема дискретной модели оператора в режиме слежения приведена на рис. 26: 1 — индикатор; 2 — импульсный элемент; 3 — экстраполятор; 4 — непрерывная часть передаточной функции оператора; 5 — управляемый элемент.

Общий вид математической модели деятельности оператора $W(p)$ в данном случае представляется произведением двух передаточных функций

$$W(p) = H(p) G(p), \quad (3.3)$$

где $G(p)$ — непрерывная часть передаточной функции оператора, описываемая соотношением типа (3.2) с учётом запаздывания в формирующей цепи; $H(p)$ — передаточная функция экстраполятора.

Например, для экстраполятора первого порядка имеем

$$H_1(p) = \left(\frac{1+pt}{T} \right) \left(\frac{1-e^{-pt}}{p} \right)^2, \quad (3.4)$$

где T — период дискретности импульсного элемента.

Экспериментальные исследования дискретной импульсной модели показали, что она дает результаты, в большей степени соответствующие реальному поведению оператора, чем линейная непрерывная модель [64].

Основным вопросом, возникающим при использовании методов теории автоматического регулирования для описания деятельности человека-оператора, является определение структуры и числовых значений параметров передаточной функции, соответствующей тем или иным видам деятельности оператора, с учетом динамических свойств объекта управления, критериев качества его поведения, характеристик окружающей среды, конструкции органов управления, средств индикации, а также индивидуальных способностей, обученности и тренированности оператора.

Для определения передаточной функции оператора, соответствующей конкретным условиям его деятельности, производится математическая обработка экспериментальных данных [6].

Наиболее просто эта задача может быть решена в предположении детерминированного характера деятельности оператора при детерминированных входных сигналах. В этом случае управляющее воздействие оператора также является детерминированной функцией времени. Как известно, изображения по Лапласу входного $X(p)$ и выходного $Y(p)$ сигналов линейной стационарной системы при нулевых начальных условиях связаны между собой соотношением

$$Y(p) = W(p) X(p). \quad (3.5)$$

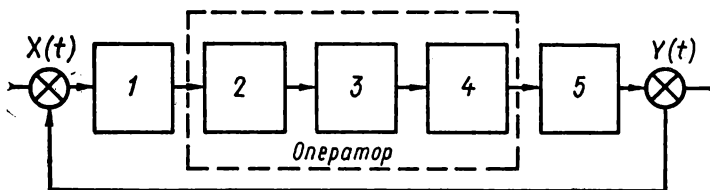


Рис. 26

Здесь $W(p)$ соответствует искомой передаточной функции оператора, а $Y(p)$ и $X(p)$ находятся по формулам

$$\begin{aligned} Y(p) &= \int_0^{\infty} e^{-pt} Y(t) dt, \\ X(p) &= \int_0^{\infty} e^{-pt} X(t) dt, \end{aligned} \quad (3.6)$$

где $X(t)$ — входной сигнал оператора; $Y(t)$ — соответствующая ему реакция, определяемая по данным эксперимента.

Вычисление передаточной функции оператора существенно усложняется при анализе его реакции на случайный входной сигнал. В этом случае даже для линейной непрерывной модели с постоянными коэффициентами приходится учитывать стохастическую природу ответной реакции оператора и использовать статистические методы обработки сигналов [61]. Для стационарных случайных процессов решение задачи дается известными в статистической динамике линейных систем соотношениями

$$S_{yx}(\omega) = W(j\omega) S_x(\omega), \quad (3.7)$$

$$S_y(\omega) = |W(j\omega)|^2 S_x(\omega), \quad (3.8)$$

где $S_x(\omega)$, $S_y(\omega)$ — спектральные плотности входного и выходного сигналов оператора; $S_{yx}(\omega)$ — взаимная спектральная плотность входного и выходного сигналов; $W(j\omega)$ — передаточная функция оператора (комплексный коэффициент усиления).

Спектральные плотности сигналов выражаются через соответствующие корреляционные функции по формулам

$$\begin{aligned} S_x(\omega) &= \int_{-\infty}^{+\infty} R_x(\tau) e^{-j\omega\tau} d\tau, \\ S_y(\omega) &= \int_{-\infty}^{+\infty} R_y(\tau) e^{-j\omega\tau} d\tau, \\ S_{yx}(\omega) &= \int_{-\infty}^{+\infty} R_{yx}(\tau) e^{-j\omega\tau} d\tau. \end{aligned} \quad (3.9)$$

При отыскании адекватного выражения передаточной функции оператора в классе дискретных импульсных моделей может быть использован аппарат дискретного преобразования Лапласа или z -преобразования. Эти преобразования связаны между собой простой заменой переменной $z = e^T$ и приводят к тождественным результатам.

Если необходимо определить значения входного и выходного сигналов оператора лишь в некоторые дискретные моменты времени $t = nT$ ($n = 0, 1, \dots$), то, пользуясь z -преобразованием, их можно

представить в виде

$$X(z) = \sum_{n=0}^{\infty} X(nT) z^{-n}, \quad (3.10)$$

$$Y(z) = \sum_{n=0}^{\infty} Y(nT) z^{-n}.$$

Следовательно, дискретная передаточная функция оператора $W(z)$ по аналогии с непрерывными системами может быть определена из соотношения

$$W(z) = Y(z)/X(z). \quad (3.11)$$

При необходимости в структуре передаточной функции может учитываться наличие того или иного типа экстраполятора входного сигнала оператора.

Таким образом, можно сделать следующие выводы о возможностях применения теории автоматического регулирования для исследования деятельности оператора [43].

1. Математические модели деятельности человека-оператора, описываемые средствами теории автоматического регулирования, отражают динамический характер воздействия человека на машины и поэтому имеют преимущества перед другими моделями во всех тех случаях, когда важен не только результат того или иного управляющего воздействия на объект, но и динамические характеристики управления во времени.

2. Преимущественные области применения методов теории автоматического регулирования в исследованиях систем человек — машина охватывают пока сравнительно частный класс однокоординатного компенсаторного и преследующего слежения.

3. Во всех случаях, когда передаточная функция человека-оператора известна, методы теории автоматического регулирования являются весьма эффективным средством исследования систем человек — машина и позволяют получить исчерпывающие данные об их свойствах и характеристиках.

4. Основная часть математических моделей деятельности оператора, разработанных на базе методов теории автоматического регулирования, относится к классу стационарных моделей с фиксированными параметрами. Это не позволяет в принципе ставить вопросы об учете изменения во времени таких важнейших характеристик человека, как утомляемость, адаптация к внешним условиям, степень обученности и тренированности. Поэтому здесь следует переходить к нестационарным нелинейным моделям, например, в рамках аналитической теории систем [5].

Информационные методы исследования. Информационные методы исследования систем человек — машина базируются на отождествлении человека-оператора, передающего информацию со средств отображения на органы управления, с каналом связи. Теоретической базой для исследования каналов связи является теория информации.

Теория информации изучает количественные закономерности, связанные с получением, обработкой, хранением и передачей информа-

ции. Под информацией в эргономике принято понимать любые изменения в управляемом процессе, отображаемые средствами представления информации или непосредственно воспринимаемые оператором, а также команды, указания о необходимости тех или иных воздействий на процесс.

Любое соотношение представляет собой совокупность сведений о некоторой физической системе. Применительно к деятельности оператора сообщение — это совокупность зрительных, акустических и иных сигналов, воспринятых в данный момент, а также «сигналов», хранимых в памяти оператора. Сообщение содержит определенное количество информации тогда, когда состояние системы заранее не известно, случайно, т. е. системе заранее присуща какая-то степень неопределенности.

В качестве меры неопределенности физической системы X , принимающей состояния X_i ($i = 1, 2, \dots, n$), используется понятие энтропии

$$H(x) = - \sum_{i=1}^n p_i \log p_i, \quad (3.12)$$

где p_i — вероятность i -го состояния системы.

При равновероятном появлении различных состояний ($p = 1/n$) энтропия максимальна:

$$H(x) = \log n. \quad (3.13)$$

Неопределенность системы уменьшается при получении каких-либо сведений об этой системе. Чем больше объем полученных сведений, чем они более содержательны, тем большей информацией о системе можно располагать. Поэтому естественно количество информации измерять уменьшением энтропии той системы, для уточнения состояния которой предназначены эти сведения.

Если априорная (исходная) энтропия системы была $H(x)$, а после получения сведения об этой системе она стала $H_0(x)$, то количество полученной при этом информации выразится формулой

$$I_x = H(x) - H_0(x). \quad (3.14)$$

Энтропия и количество информации измеряются в двоичных единицах информации или битах.

В общем случае структурная схема гиповой технической системы связи соответствует рис. 27, где 1 — источник информации; 2 — кодировщик; 3 — кодировщик канала; 4 — канал; 5 — декодировщик канала; 6 — декодировщик информации; 7 — приемник информации (адресат).

Под источником информации здесь понимается любое устройство или человек, генерирующий информацию, подлежащую передаче.

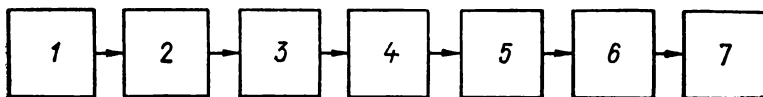


Рис. 27

Термином «канал» определяют физическую среду, используемую при передаче информации, а адресатом называют человека или устройство, которое получает информацию. Эти три блока считаются фиксированными и обязательно присутствуют в любой системе связи.

Блоки 2 и 3 могут структурно изменяться и соответствуют устройствам, преобразующим сообщение на выходе источника в сигнал, пригодный для передачи по данному каналу. Блоки 5 и 6 также не имеют фиксированной структуры и соответствуют устройствам, воспроизводящим полезную информацию из сигналов, поступивших на выход канала.

Основной принцип функционирования канала может быть выражен преобразованием последовательности входных событий x , принадлежащих некоторому пространству X , в выходные события y , являющиеся точками пространства Y .

Канал называется дискретным, если его пространства входных и выходных событий дискретны. При этом преобразование точки входного пространства в точку y выходного пространства в общем случае определяется условным распределением вероятностей $p(y/x)$, заданным на пространстве Y для любого $x \in X$ и описывающим случайные шумы в канале. Если пространства событий непрерывны, то при описании функционирования канала под $p(y/x)$ следует понимать соответствующую плотность условного распределения вероятностей.

Канал, в котором условное распределение вероятностей одно и то же для всех последовательных входных и выходных событий, называют стационарным каналом без памяти. Если это условное распределение вероятностей является функцией предыдущих событий на входе или выходе, то говорят, что канал обладает памятью. Память называют конечной, если эта зависимость распространяется только на конечное число предшествующих событий.

С точки зрения функционирования каналов во времени различают каналы с дискретным и непрерывным временем. В первом случае источник информации порождает входные события канала в строго определенные моменты времени, во втором — каждое очередное событие в произвольные моменты времени.

Для оценки степени адекватности модели канала специфике деятельности человека в системах человек — техника [70] структурная схема типовой системы человек — техника представлена на рис. 28, где 1 — объект управления (источник информации); 2 — индикатор (кодер источника); 3 — рецептор анализатора (кодер канала); 4 — нервные пути (канал); 5 — мозговые центры анализаторов; 6 — центры обработки информации и принятия решения; 7 — центры управления речевым и двигательным аппаратами человека; 8 — исполнительные элементы речевого и двигательного аппаратов; 9 — центры оперативной памяти; 10 — центры долговременной памяти (блоки 5—10 — декодер канала); 11 — органы управления объекта (декодер приемника); 12 — средства преобразования информации (декодер адресата); 13 — адресат.

Рассмотрим функционирование собственно канала связи, выделенного на рис. 28 штрихпунктирным прямоугольником.

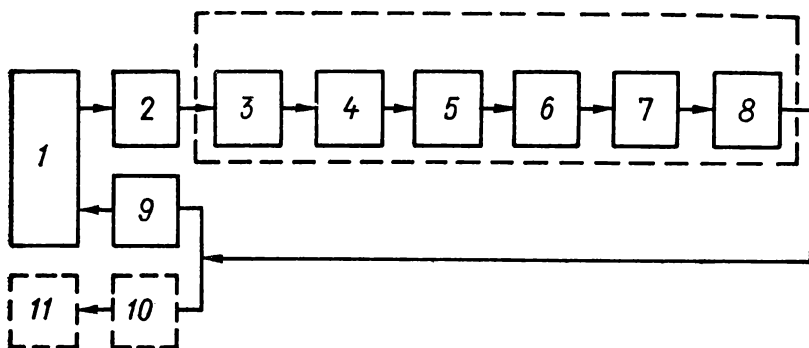


Рис. 28

Входные события канала формируются индикационными устройствами объекта, которые являются аналогами кодера источника и преобразуют первичные контролируемые процессы в специальные сигналы — носители информации, характеризующиеся их физической природой и интенсивностью и приспособленные для восприятия человеком.

Роль кодера канала связи играют рецепторы соответствующих анализаторов человека. Основной функцией рецептора, как известно, является превращение энергии действующего раздражителя в нервный процесс. Каждый рецептор приспособлен к приему сигналов-раздражителей вполне определенной модельности, в то время как выходные сигналы всех рецепторов по своей природе едины и приспособлены для передачи по нервным путям. Именно это свойство рецепторов и позволяет отождествлять их функции с функциями устройств кодирования информации перед передачей ее по каналам связи.

Функции собственно канала как физической среды для распространения энергии сигналов — носителей информации возлагаются на нервные ткани или нервные пути, связывающие рецепторы с мозговыми центрами анализаторов и мозговые центры с исполнительными органами моторной сферы человека.

Выходные события канала, формируемые его декодером, согласно рис. 20 преобразуются далее декодером приемника информации. Роль декодера приемника в системах человек — техника отводится непосредственно органам управления объекта. Именно органы управления преобразуют управляющие реакции человека-оператора в соответствующие управляющие воздействия на объект [23]. В этом смысле можно говорить, что органы управления декодируют вводимую оператором командную информацию в форму или язык, приемлемый и понятный для объекта.

Из сравнения схем на рис. 27 и 28 можно установить, что наряду с имеющейся аналогией между системой человек — машина и системой связи существуют значительные расхождения, которые заключаются в следующем.

1. В отличие от системы связи система человек — машина всегда является структурно-замкнутой, т. е. источник информации одновре-

менно является адресатом, а условия формирования входных событий канала зависят от событий на его выходе.

2. Объекты управления в силу динамического характера управляемых процессов относятся к категории нестационарных источников. Поэтому вероятностные характеристики элементов пространства входных событий канала зависят от времени. Оператор сталкивается с дополнительной неопределенностью в моментах времени формирования входных событий канала.

3. При выработке входных управляющих воздействий учитывается информация не только о текущем состоянии процесса, но и его предистория и прогноз будущего состояния. Поэтому декодер канала содержит центры оперативной памяти, логической обработки информации и принятия решений на управление.

4. При принятии решения должна учитываться целевая функция системы в целом. Поэтому декодер канала содержит центры долговременной памяти, где сконцентрирована информация о целевой функции памяти.

5. Оператору присущи свойства запаздывания реакции на сигнал, а также конечная разрешающая способность по различению стимулов раздражителей и ответных реакций. Поэтому оператор может быть представлен только каналом с дискретным временем.

Таким образом, в качестве наиболее адекватной модели деятельности человека в эргатических системах при использовании информационных методов в общем случае следует рассматривать модель нестационарного дискретного канала с дискретным временем, конечной оперативной памятью и внутренним источником информации о целевой функции системы.

Рассмотрим информационную модель деятельности оператора, разработанную с учетом приведенных выше соображений для простейшей структуры деятельности, когда известна реакция на каждый из сигналов множества [5].

Под пространством входных событий канала L_x будем понимать множество различных состояний индикационных устройств, размещенных на информационной панели. Пусть n — общее число индикаторов панели и q_i ($i = 1, 2, \dots, n$) — число различных для оператора состояний индикатора. Тогда общее число различных состояний дается формулой

$$N = \prod_{i=1}^n q_i. \quad (3.15)$$

Теперь x_j , $j = 1, 2, \dots, N$, будет обозначать j -й элемент пространства L_x каналов связи.

Аналогично под пространством выходных событий канала L_y будем понимать множество различных состояний совокупности органов управления, предусмотренных конструкцией пульта управления.

Пространства L_x и L_y эквивалентны по числу элементов, т. е. стимулу x_j соответствует реакция оператора y_j .

Для характеристики функционирования канала по времени на отрезке $0 \leq t \leq T$ введем упорядоченную систему моментов времени

$$0 = t_0 < t_1 < t_2 \dots < t_k = T$$

такую, что протяженность ν -го отсчета оператора во времени составляет величину

$$\delta_\nu = t_\nu - t_{\nu-1}, \quad (3.16)$$

где $\nu = 1, 2, \dots, k$.

Исчерпывающее статистическое описание пространства входных событий канала может быть получено, если для любого $j = 1, 2, \dots, N$ задать безусловные вероятности $p(x_j, \nu)$ реализации элемента x_j на отрезке времени, соответствующему ν -му отсчету оператора.

По условию нормировки

$$\sum_{j=1}^N p(x_j, \nu) = 1. \quad (3.17)$$

Будем трактовать вероятности $p(x_j, \nu)$ как компоненты N -мерного вероятностного вектора состояний информационной панели, т. е.

$$P(X, \nu) = \{p(x_1, \nu), p(x_2, \nu), \dots, p(x_N, \nu)\}. \quad (3.18)$$

Аналогично для y_j -го управляющего воздействия имеем

$$P(Y, \nu) = \{p(y_1, \nu), p(y_2, \nu), \dots, p(y_N, \nu)\}. \quad (3.19)$$

С учетом последних выражений математическую модель деятельности оператора можно выразить формулой

$$P(Y, \nu) = \Psi[P(X, \nu), \nu], \quad (3.20)$$

где Ψ — символ в общем случае нестационарного преобразования дискретной последовательности различных состояний информационной панели в дискретную последовательность различных управляющих воздействий оператора.

При отсутствии ошибок в формировании управляющего воздействия выражение (3.20) приобретает вид

$$P(Y, \nu) = P(X, \nu) E, \quad (3.21)$$

где E — единичная $N \times N$ -матрица.

В случае ошибок оператора при любом входном событии сохраняется ненулевая вероятность реализации любого из управляющих воздействий. Поэтому, применив формулу полной вероятности, можно записать

$$\begin{aligned} P(Y, \nu) &= p(x_1, \nu) p_\nu(y_1/x_1) + \dots + p(x_i, \nu) p_\nu(y_i/x_i) + \\ &+ \dots + p(x_N, \nu) p_\nu(y_N/x_N). \end{aligned} \quad (3.22)$$

Используя последнее соотношение, в общем случае получаем

$$P(y, x) = P(x, \nu) M(\nu), \quad (3.23)$$

где $M(\nu)$ — $N \times N$ -матрица условных вероятностей, определяемая соотношением

$$M(\nu) = \begin{bmatrix} p_\nu(y_1/x_1) & p_\nu(y_2/x_1) & \dots & p_\nu(y_N/x_1) \\ p_\nu(y_1/x_2) & p_\nu(y_2/x_2) & \dots & p_\nu(y_N/x_2) \\ \dots & \dots & \dots & \dots \\ p_\nu(y_1/x_N) & p_\nu(y_2/x_N) & \dots & p_\nu(y_N/x_N) \end{bmatrix}. \quad (3.24)$$

Математические модели вида (3.21) и (3.23) могут быть непосредственно использованы для оценки фактической информационной напряженности оператора.

Рассмотрим произвольный ν -й отсчет. Пусть x_μ — состояние информационной модели, отвечающее этому отсчету. Тогда условная энтропия управляющих реакций оператора на входной сигнал x_μ выражается формулой

$$H_\nu(Y/x_\mu) = - \sum_{j=1}^N p_\nu(y_j/x_\mu) \log_2 p_\nu(y_j/x_\mu). \quad (3.25)$$

При усреднении по всем возможным состояниям информационной модели получим общую условную энтропию

$$H_\nu(Y/X) = \sum_{\mu=1}^N \sum_{j=1}^N p(x_\mu, \nu) p_\nu(y_j/x_\mu) \log_2 p_\nu(y_j/x_\mu). \quad (3.26)$$

Количество информации, которое оператор перерабатывает при ν -м отсчете, можно представить в виде

$$I_\nu(Y, X) = H_\nu(Y) - H_\nu(Y/X), \quad (3.27)$$

где энтропия $H_\nu(Y)$ определяется соотношением

$$H_\nu(Y) = - \sum_{j=1}^N p(y_j, \nu) \log_2 p(y_j, \nu). \quad (3.28)$$

В окончательном виде количество информации, перерабатываемое оператором в течение ν -го отсчета, выразится соотношением

$$\begin{aligned} I_\nu(Y, X) = & - \sum_{j=1}^N p(y_j, \nu) \log_2 p(y_j, \nu) + \\ & + \sum_{\mu=1}^N \sum_{j=1}^N p(x_\mu, \nu) p_\nu(y_j, x_\mu) \log_2 p_\nu(y_j/x_\mu). \end{aligned} \quad (3.29)$$

Из последнего выражения легко видеть, что в случае безошибочного функционирования оператора на ν -м интервале количество обрабатываемой информации представится в виде

$$I_\nu = - \sum_{i=1}^N p(x_i, \nu) \log_2 p(x_i, \nu). \quad (3.30)$$

Для того чтобы определить количество информации, перерабатываемой оператором в течение некоторого времени, приведенные выше соотношения должны быть дополнены соотношениями, характеризующими деятельность оператора во временной области.

Основные трудности, возникающие при этом, обусловлены главным образом стохастическим характером процесса деятельности операторов. Это приводит к необходимости считать временные затраты оператора на реализацию ν -го отсчета случайными величинами, подчиненными определенным законам распределения вероятностей. Число различных по содержанию отсчетов, с которыми оператору приходится иметь дело, совпадает с числом элементов пространства входных событий L_x . Поэтому для описания деятельности оператора во времени достаточно задать N функций распределения вероятностей временных затрат оператора на реализацию рассматриваемой совокупности отсчетов.

Для учета изменения характеристик работоспособности оператора на отрезке времени функционирования в зависимости от обучаемости воздействия стресса или усталости параметры указанных распределений должны отличаться для каждого ν -го отсчета.

С учетом сказанного всю необходимую исходную информацию для описания деятельности оператора во времени целесообразно представить в виде матричной функции распределения вероятностей

$$F(N, K, \Theta) = \begin{bmatrix} F_{11}(\Theta) & F_{12}(\Theta) & \dots & F_{1k}(\Theta) \\ F_{21}(\Theta) & F_{22}(\Theta) & \dots & F_{2k}(\Theta) \\ & & \dots & \\ F_{N1}(\Theta) & F_{N2}(\Theta) & \dots & F_{Nk}(\Theta) \end{bmatrix}. \quad (3.31)$$

Матрица $F(N, K, \Theta)$ имеет размерность $N \times K$, причем нумерация строк матрицы совпадает с нумерацией элементов пространства входных данных событий, а нумерация столбцов соответствует порядковым номерам отсчетов, реализуемых оператором с момента $t = 0$ [16].

Произвольный элемент $F_{ij}(\Theta)$ матрицы по физическому смыслу представляет собой интегральный закон распределения временных затрат δ_{ij} на реализацию j -го отсчета, соответствующего элементу x_i пространства входных событий, т. е.

$$F_{ij}(\Theta) \equiv p(\delta_{ij} < \Theta). \quad (3.32)$$

Следует отметить, что степень изученности информационных процессов в центральной нервной системе человека не позволяет аналитически рассчитать элементы матрицы (3.31).

Единственно возможный путь их практического определения в настоящее время состоит в сборе, систематизации и обобщении статистических данных, полученных при экспериментальном исследовании деятельности оператора с использованием многоэлементных информационных панелей.

Информационные методы исследования деятельности оператора применяются при решении ряда практических задач [29]. Во-первых, количество перерабатываемой человеком информации является мерой сложности его работы. Следовательно, таким способом можно сравнивать между собой различные виды деятельности. Во-вторых, зная количество информации, можно оценить время, которое оператор затра-

чивает на ее переработку, поскольку между ними, как правило, существует линейная зависимость. В-третьих, знание количества информации позволяет согласовать темп ее выдачи (производительность источника сообщений) с психофизиологическими возможностями человека по ее приему и обработке, иными словами, с пропускной способностью оператора.

Рассмотрим последнюю задачу, получившую наибольшее распространение в инженерно-психологических исследованиях.

Прежде всего рассчитывается количество информации, которое оператор должен переработать за заданный промежуток времени. Для подсчета ее количества при выполнении определенных видов деятельности можно воспользоваться следующими соотношениями.

Количество информации, получаемой оператором от k независимых индикаторов, определяется по формуле

$$H_{\text{общ}} = \sum_{j=1}^k H_j = - \sum_{j=1}^k \sum_{i=1}^n p_{ij} \log_2 p_{ij}, \quad (3.33)$$

где H_j — энтропия j -го индикатора ($j = 1, 2, \dots, k$).

Количество информации при считывании показаний прибора определяется выражением

$$H_{\text{пр}} = \log_2 (x_{\text{max}} - x_{\text{min}})/2\delta, \quad (3.34)$$

где x_{max} , x_{min} — максимальное и минимальное значения шкалы прибора; δ — абсолютная погрешность считывания показаний с прибора.

Информация, подлежащая запоминанию, может быть трех видов. Если запоминанию подлежит m символов (знаков, сигналов) из общего числа k символов, то количество запоминаемой информации вычисляется по формуле

$$H_{\text{зап}} = \log_2 k^m = m \log_2 k. \quad (3.35)$$

Если от оператора не требуется запоминать порядок поступления сигналов, а необходимо запомнить только сами сигналы, то количество информации будет меньше:

$$H_{\text{зап}} = \log_2 m = \log_2 k. \quad (3.36)$$

В тех случаях, когда совокупность сигналов заранее известна, но неизвестен порядок их следования, количество информации во всей последовательности будет таким:

$$H_{\text{зап}} = \log_2 m = \log_2 k. \quad (3.37)$$

Количество информации, используемое при осуществлении в уме простейших вычислительных операций (сложение, деление и т. д.), определяется выражением

$$H_{\text{выч}} = \sum_{i=1}^k \log_2 N_i + \log_2 R, \quad (3.38)$$

где N_i — максимальные значения используемых при вычислении чисел; R — максимально возможное значение результатов вычисления; k — количество чисел, используемых для получения результата R .

При проверке логических условий количество информации имеет вид

$$H_{\text{лог}} = \sum_{i=1}^l \log_2 n_i, \quad (3.39)$$

где l — число проверяемых логических условий; n_i — число возможных альтернатив (исходов), возникающих при проверке i -го условия.

Извлечение из памяти осуществляется всякий раз, когда нужно принять решение по полученной информации. Оператор должен вспомнить, какое действие необходимо осуществить в ответ на появление данного сигнала. Количество информации при этом определяется формулой

$$H_{\text{пом}} = -Q \sum_{i=1}^r p_i \log_2 p_i, \quad (3.40)$$

где p_i — вероятность осуществления i -го действия; r — общее число различных действий, которые может выполнять оператор; Q — число одновременно выполняемых действий из r возможных ($Q \leq r$).

Наконец, количество информации при выполнении двигательной задачи вычисляется по формуле

$$H_{\text{дв}} = \log_2 (2A/W), \quad (3.41)$$

где A — амплитуда движения руки, т. е. расстояние, на которое перемещается рука; W — ширина цели, т. е. ширина органа управления, на который должна попасть рука оператора.

Среднее время переработки всей информации при решении задачи оператором выразится равенством

$$T = \sum_{i=1}^n (H_i/V_i), \quad (3.42)$$

где H_i — количество информации i -го вида; V_i — скорость переработки информации i -го вида.

Ориентировочные значения скорости переработки некоторых видов информации приводятся в табл. 5.

Полученное среднее время переработки информации при решении задачи T сравнивается с допустимым временем решения задачи T_0 . Считается, что при выполнении условия

$$T \leq T_0 \quad (3.43)$$

оператор может решить задачу, т. е. его пропускная способность не меньше, чем скорость поступления информации (производительность источника сообщений).

Т а б л и ц а 5

| Вид информации | v , дв·ед/с |
|--|---------------|
| Восприятие поступающих сигналов | 18,2—24,4 |
| Решение логических задач | 8,2—15,1 |
| Арифметические операции | 6,0—12,0 |
| Информация для долговременного запоминания | 16,4—24,5 |
| Нахождение органа управления | 18,2—24,4 |
| Двигательные задачи | 8,5—11,5 |
| Выдача речевых команд | 17,8—26,3 |
| Снятие показаний со стрелочных приборов | 4,2—8,4 |

В заключение необходимо подчеркнуть, что применение методов теории информации в практике инженерно-психологических исследований связано с рядом трудностей [54].

1. До конца не раскрыты способы формирования оператором алфавита сигналов. Принято считать, что этот алфавит совпадает с физическим алфавитом сигналов, что, однако, не всегда имеет место. В процессе переработки информации человек может укрупнять информацию, производить перекодирование, использовать прошлый опыт и т. д., т. е. пользоваться алфавитом, отличающимся от входного алфавита сигналов.

2. Субъективные вероятности появления того или иного состояния системы могут отличаться от объективных. Например, для необученного оператора все состояния равновероятны; для полностью обученного оператора, хорошо усвоившего статистическую природу сигналов, субъективные вероятности близки к объективным.

3. Теория информации не учитывает смысловую сторону информации. На деятельность человека оказывает влияние не только неопределенность поступления сигнала, но и его смысл.

4. В теории информации рассматриваются только стационарные процессы, статистические характеристики которых не меняются с течением времени. Свойства человека не остаются неизменными: они меняются по мере обучаемости, утомления, воздействия помех и т. д.

Наличие этих ограничений приводит на практике к тому, что рассчитанное количество информации не всегда совпадает с количеством действительно переработанной человеком информации. Поэтому информационные методы могут быть использованы лишь для ориентировочных оценок. Тем не менее эти методы находят широкое применение в инженерно-психологических исследованиях.

Корреляционно-регрессионные методы исследования. Для изучения характера статистических зависимостей между количественными характеристиками деятельности оператора и влияющими на них факторами (условиями деятельности, характеристиками алгоритма, свойствами оператора и т. д.) по результатам экспериментальных данных в инженерной психологии применяются корреляционно-регрессионные методы исследования.

Проведение исследования корреляционно-регрессионными методами включает следующие основные этапы, выполняемые, как правило, последовательно.

1. Выбор основных факторов, влияющих на характеристики деятельности.

2. Сбор экспериментальных данных, проверка выполнения условий (ограничений) применимости корреляционно-регрессионного анализа.

3. Нанесение данных на график, предварительный анализ корреляционного поля, измерение тесноты связи, вычисление выборочного коэффициента корреляции.

4. Исследование природы влияния факторов на характеристики деятельности для установления вида зависимости между ними. Установление общего вида зависимости с максимально возможным использованием априорной информации о процессе.

5. Построение эмпирической линии регрессии методом наименьших квадратов.

6. Исследование статистических свойств регрессионной зависимости, проверка гипотез о параметрах, об адекватности приближений. Останемся на содержании указанных этапов исследования.

1. Выбор факторов, влияющих на количественные характеристики деятельности оператора. Этот выбор может быть осуществлен либо путем изучения деятельности оператора, либо с помощью опроса специалистов (экспертная оценка). Возможно также комбинированное использование обоих методов.

Т а б л и ц а 6

| Фактор | Ранжировочный ряд каждого эксперта | | | | x_i | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$ |
|--------|------------------------------------|---|-----|-----|-------|-----------------|---------------------|
| | 1 | 2 | ... | m | | | |
| 1 | | | | | | | |
| 2 | | | | | | | |
| 3 | | | | | | | |
| ⋮ | | | | | | | |
| ⋮ | | | | | | | |
| n | | | | | | | |

Для изучения деятельности оператора можно применить алгоритмический анализ [42]. После составления алгоритма деятельности оператора определяются основные факторы, влияющие на выполнение операций алгоритма. Затем эти факторы группируются и производится их окончательный отбор.

Выбор факторов методом экспертной оценки производится группой экспертов в составе восьми — десяти специалистов, хорошо знакомых с характером изучаемой деятельности. Каждому эксперту предлагается произвести отбор и ранжировку факторов, влияющих на исследуемую характеристику деятельности.

Ранжирование заключается в построении предпочтительного ряда факторов, т. е. на первое место становится фактор, который, по мнению эксперта, более всего влияет на исследуемую характеристику, на второе место — фактор, менее влияющий, и т. д. Результаты ранжирования всех экспертов по каждому фактору обычно сводятся в таблицы в виде табл. 6, элементы которой называются ранжировочными баллами.

Наиболее значимым будет фактор, сумма ранжировочных баллов для которого оказалась минимальной, следующим — фактор, у которого минимальна сумма баллов из оставшихся, и т. д.

Степень согласованности мнений экспертом оценивается с помощью коэффициента согласованности мнений (коэффициента конкордации), который рассчитывается по формуле

$$W = \frac{12S}{m^2(n^3 - n)}, \quad (3.44)$$

где x_i — суммарная оценка по i -му фактору — сумме рангов ($i = 1, 2, \dots, n$); $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ — средняя сумма рангов; S — сумма квадратов отклонений, т. е. сумма чисел, стоящих в последнем столбце табл. 6; m — число экспертов; n — число факторов.

Коэффициент конкордации — это величина, изменяющаяся в пределах от 0 до 1. Значимость коэффициента конкордации при выбранном уровне значимости p (%) проверяется по специальным таблицам. Если коэффициент конкордации значим, то существует неслучайная согласованность во мнениях экспертов (чем ближе коэффициент конкордации к единице, тем больше согласованность мнений).

При отсутствии согласованности мнений проводится совещание экспертов с целью рассмотрения причин разногласий и выработки согласованной оценки.

Таким образом, в результате экспертной оценки оказывается отобранное множество факторов, подлежащих изучению и оценке их влияния на характеристики деятельности оператора.

2. Оценка применимости корреляционно-регрессионного анализа. Основными условиями применимости корреляционно-регрессионного анализа являются:

а) случайность выборки (стохастическая независимость наблюдений), несвязанность наблюдения с предыдущими и отсутствие влияния на последующие;

б) однородность дисперсий: дисперсии S_y^2 оцениваемой характеристики y одинаковы для разных фиксированных наборов факторов $\{x_1, x_2, \dots, x_n\}$;

в) результаты измерения оцениваемой характеристики $y = \{y_1, y_2, \dots, y_k\}$ (k — число измерений) для разных фиксированных наборов факторов $\{x_1, x_2, \dots, x_n\}$ являются независимыми нормально распределенными случайными величинами.

Однако невыполнение перечисленных условий не всегда означает необходимость немедленного отказа от применения метода корреляционно-регрессионного анализа. При больших объемах выборок при-

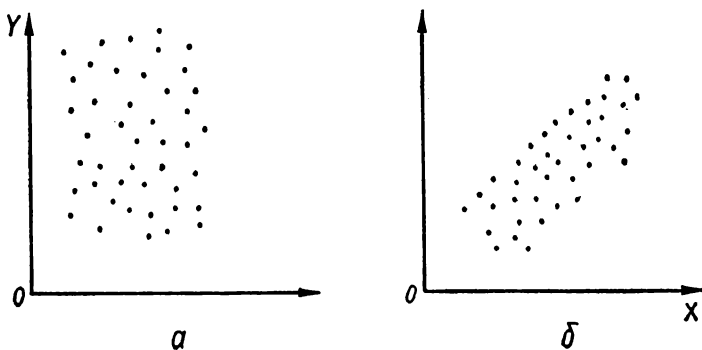


Рис. 29

менение метода дает хорошие результаты даже при невыполнении ограничений.

3. Анализ корреляционного поля. Вычисление коэффициента корреляции. Экспериментальные данные, подготовленные для корреляционно-регрессионного анализа, обычно изображаются на так называемом корреляционном поле [18]. Корреляционное поле — график в прямоугольных координатах, каждая точка которого соответствует одному наблюдению и имеет координатами наблюденные значения коррелируемых признаков (рис. 29).

Если точки корреляционного поля тесно группируются вокруг некоторой прямой, есть основания предполагать, что между y и фактором x_i существует корреляционная связь (рис. 29, б). Если же облако «рыхлое» (рис. 29, а), точки располагаются часто одинаково во всех направлениях и вряд ли стоит ожидать наличия корреляционной связи переменных.

Для окончательного принятия решения о корреляции между y и фактором x_i рассчитывается выборочный коэффициент корреляции

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}. \quad (3.45)$$

Здесь x_1, x_2, \dots, x_n и y_1, y_2, \dots, y_k — две выборки из нормальных совокупностей: $x_{11}, x_{12}, \dots, x_{1n}; \dots; x_{l1}, x_{l2}, \dots, x_{ln}; x_{k1}, x_{k2}, \dots, x_{kn}; y_1, y_2, \dots, y_k$.

Статистически значимое отличие r от нуля можно понимать как наличие связи случайных величин x и y . Проверка наличия связи осуществляется по критерию Стьюдента с $k - 2$ степенями свободы (k — число наблюдений) [12].

Гипотеза о наличии связи принимается, если выполняется условие

$$r > \sqrt{\frac{t_{rp}}{n - 2 + t_{rp}}}, \quad (3.46)$$

где t_{rp} — граничное значение распределения Стьюдента с $k - 2$ степенями свободы.

4. Установление вида зависимости. Пусть после анализа корреляционного поля оказалось, что исследуемая характеристика оператора y связана с факторами x_1, x_2, \dots, x_n некоторым соотношением. Тогда задача отыскания этого соотношения распадается на два этапа: установление общего вида зависимости и вычисление оценок ее параметров.

Стандартные методы выбора общего вида кривой не известны в статистике: необходимо сочетать визуальный анализ корреляционного поля с качественным анализом природы переменных, причем последний должен иметь немалый удельный вес. Зачастую оказывается, что точки ложатся в корреляционном поле так, что известные статистические критерии не отвергают согласия с экспериментальными данными нескольких различных по типу кривых, и только априорные знания материала позволяют решить, какая из них более адекватна. Очень

часто для описания кривой применяются полиномиальные многочлены.

5. Построение линии регрессии (нахождение параметров зависимости). Общим методом построения оценки параметров кривой

$$y = f(x_1, x_2, \dots, x_n; a_0, a_1, \dots, a_n) \quad (3.47)$$

является метод наименьших квадратов.

Суть этого метода состоит в том [6], что параметры зависимости a_0, a_1, \dots, a_n подбираются так, чтобы величина

$$\Delta = \sum_{j=1}^k (y_j - \hat{y}_j)^2 \quad (3.48)$$

принимала наименьшее значение; здесь \hat{y}_j — оценивающая зависимость; k — количество наблюдений.

Рассмотрим достаточно простой и часто используемый на практике случай, когда зависимость величины y от факторов x_1, x_2, \dots, x_n линейна, т. е.

$$y = a_0 + a_1 x_1 + \dots + a_n x_n. \quad (3.49)$$

На основе исходных данных составляется система условных уравнений:

$$\begin{aligned} y_1 &= a_0 + a_1 x_{11} + \dots + a_i x_{1i} + \dots + a_n x_{1n}, \\ &\dots \\ y_j &= a_0 + a_1 x_{j1} + \dots + a_i x_{ji} + \dots + a_n x_{jn}, \\ &\dots \\ x_k &= a_0 + a_1 x_{k1} + \dots + a_i x_{ki} + \dots + a_n x_{kn}. \end{aligned} \quad (3.50)$$

В соответствии с методом наименьших квадратов записывается уравнение

$$\sum_{j=1}^k (y_j - \hat{y}_j)^2 = \sum_{j=1}^k [y_j - (a_0 + a_1 x_{j1} + \dots + a_n x_{jn})]^2. \quad (3.51)$$

После дифференцирования по переменным $a_0, a_1, a_2, \dots, a_n$ и приравнивания к нулю полученных уравнений получаем систему нормальных уравнений.

Решение системы удобно производить в матричной форме

$$Y = XA, \quad (3.52)$$

где Y — вектор-столбец зависимости переменной; X — матрица значений x_{ij} ; A — вектор-столбец неизвестных коэффициентов.

После умножения обеих частей уравнения на транспонированное значение матрицы X^T и на обратную матрицу $(X^T X)^{-1}$ имеем

$$(X^T X)^{-1} (X^T Y) = (X^T X) (X^T X)^{-1} A. \quad (3.53)$$

Поскольку произведение $(X^T X) (X^T X)^{-1} = I$, получаем

$$A = (X^T X)^{-1} (X^T Y), \quad (3.54)$$

откуда для произвольного коэффициента a_l находим

$$a_l = \sum_{i=1}^n c_{il} \sum_{j=1}^k y_j x_{jl}, \quad (3.55)$$

где c_{il} — элемент матрицы $(X^T X)^{-1}$, лежащий на пересечении j -й строки и l -го столбца.

В частном случае получаем

$$y = a_0 + a_1 x, \quad (3.56)$$

$$a = \frac{\sum_{j=1}^k y_j}{k} = \bar{y}, \quad (3.57)$$

$$a_1 = \frac{\sum_{j=1}^k (x_j - \bar{x}) y_j}{\sum_{j=1}^k (x_j - \bar{x})^2},$$

иначе

$$\hat{y} = \bar{y} + r \frac{S_y}{S_x} x, \quad (3.58)$$

где r — выборочный коэффициент корреляции;

$$S_y = \sqrt{\sum_{j=1}^k (y_j - \bar{y})^2}; \quad S_x = \sqrt{\sum_{j=1}^k (x_j - \bar{x})^2}. \quad (3.59)$$

6. Статистический анализ полученного уравнения. Этот анализ заключается в определении значимости уравнения в целом и значимости каждого из коэффициентов регрессии a_i .

Для определения значимости уравнения вычисляем дисперсию зависимой переменной S_y^2 и остаточную дисперсию $S_{\text{ост}}^2$ соответственно по формулам

$$S_y^2 = \frac{\sum_{j=1}^k (y_j - \bar{y})^2}{k - 1}; \quad S_{\text{ост}}^2 = \frac{\sum_{j=1}^k (y_j - \hat{y}_j)^2}{k - n - 1}. \quad (3.60)$$

Затем вычисляем значение F -критерия Фишера

$$F = S_y / S_{\text{ост}}, \quad (3.61)$$

которое сравниваем с табличным значением F_τ для выбранного p (%) уровня значимости и для числа степеней свободы $\nu_1 = k - 1$, $\nu_2 = k - n - 2$.

Если $F > F_\tau$, то уравнение считается значимым.

В случае значимости уравнения проводим проверку значимости каждого коэффициента регрессии по t -критерию Стьюдента с числом степеней свободы, равным $\nu = k - n - 1$.

Вычисляем значение t_{ai} :

$$t_{ai} = a_i / S_{ai}, \quad (3.62)$$

где $S_{ai} = S_{\text{ост}c_{ii}}$ — дисперсия коэффициента a_i ; c_{ii} — i -й диагональный элемент матрицы, обратной к матрице системы нормальных уравнений.

Рассчитанное значение t_{ai} сравнивается с табличным t_r . Если $t_{ai} > t_r$, то коэффициент a_i признается значимым (отличным от нуля). В противном случае i -й фактор исключается из рассмотрения и производится корректировка уравнения.

Статистический анализ уравнения заканчивается оценкой коэффициента множественной корреляции R , определяющего, на сколько процентов вариация переменной y объясняется полученным уравнением

$$R = \sqrt{1 - D/D_y}, \quad (3.63)$$

где D — определитель матрицы коэффициентов парной корреляции окончательного уравнения

$$\begin{bmatrix} 1 & r_{yx_1} & r_{yx_2} & \dots & r_{yx_k} \\ r_{yx_1} & 1 & r_{x_1x_2} & & r_{x_1x_k} \\ r_{yx_2} & r_{x_2x_1} & 1 & & r_{x_2x_k} \\ & & & \dots & 1 \end{bmatrix}, \quad (3.64)$$

D_y — минор, полученный вычеркиванием первой строки и первого столбца.

Для интерпретации уравнения регрессии применяются частные коэффициенты эластичности e_i , показывающие, на сколько процентов в среднем изменяется функция с изменением аргумента на 1 % при фиксированном значении других аргументов:

$$e_i = a_i \bar{x}_i / \bar{y}. \quad (3.65)$$

Используются также β -коэффициенты, показывающие, на какую часть среднеквадратического отклонения изменяется функция с изменением соответствующего фактора на величину своего среднеквадратического отклонения:

$$\beta_i = a_i S_{x_i} / S_y. \quad (3.66)$$

Методы теории массового обслуживания. Общая структурная схема системы массового обслуживания показана на рис. 30.

В структуре системы массового обслуживания можно выделить следующие четыре элемента.

1. Объект обслуживания, описываемый математически одним или несколькими потоками случайных заявок на обслуживание. Эти потоки называют входными потоками системы массового обслуживания.

2. Очередь как совокупность заявок, ожидающих обслуживания в данный момент времени. Она обычно характеризуется некоторой «дисциплиной обслуживания»: обслуживание в порядке поступления требований, случайный выбор заявок на обслуживание, обслуживание заявок с учетом их важности и т. д.

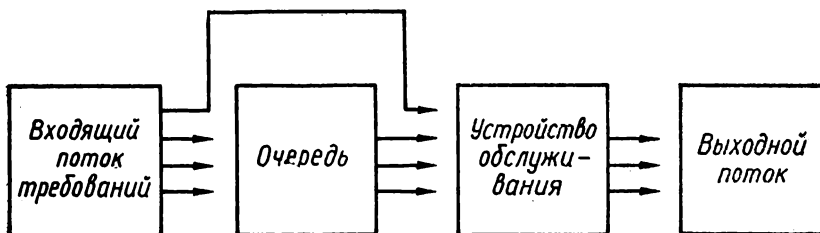


Рис. 30

3. Устройство обслуживания, которое характеризуется количеством каналов обслуживания, комбинацией каналов, законом распределения времени обслуживания и т. д.

4. Выходной поток, или поток требований, покидающий систему массового обслуживания. Он состоит из обслуженных и необслуженных требований.

Основные закономерности функционирования систем массового обслуживания [19], а также показатели качества обслуживания существенно зависят от структуры системы, свойств входного потока заявок, принятой дисциплины обслуживания и параметров процесса обслуживания.

Рассмотрим типовую постановку задачи исследования системы массового обслуживания.

Пусть задан поток однородных случайных событий $x(t)$, который является входным потоком заявок для системы массового обслуживания. Функция $x(t)$ по физическому смыслу определяет собой число требований, нуждающихся в обслуживании на отрезке времени $(0, t)$. Поскольку заявки случайны, то $x(t)$ — случайная функция времени и может быть описана n -мерным законом распределения

$$\Phi(t_1, \dots, t_n; k_1, \dots, k_n) = p\{x(t_1) = k_1, \dots, x(t_n) = k_n\}. \quad (3.67)$$

Реально существующие потоки заявок обладают рядом специфических свойств, значительно облегчающих поведение исследований. Основными из указанных свойств являются свойства стационарности, ординарности и отсутствия последействия.

Стационарными называются такие потоки, для которых количество заявок, поступающих в течение промежутка $(0, t)$, зависит лишь от длины этого промежутка и не зависит от начала отсчета времени [19].

Ординарными называются такие потоки, для которых вероятность поступления более чем одной заявки на любом промежутке (t_1, t_2) стремится к нулю при стремлении к нулю длины этого промежутка.

Поток заявок называется потоком без последействия, если число заявок на обслуживание, поступающих в систему после некоторого момента времени t , не зависит от того, какое число заявок поступило в систему до этого момента.

Потоки заявок, обладающие всеми перечисленными выше свойствами одновременно, получили наименование простейших.

Пусть теперь задано некоторое устройство обслуживания. Для него определено число каналов обслуживания, а также задана парал-

лельно-последовательная структура, в которую эти каналы объединены.

Каждый j -й канал обслуживания является заданным, если известна функция распределения времени обслуживания $B_j(y)$, где $j = 1, 2, \dots, m$, m — число каналов обслуживания.

По смыслу

$$B_j(Y) = p_j(Y \leq y) \quad (3.68)$$

выражает вероятность того, что непрерывная случайная величина Y , характеризующая время обслуживания заявки в j -м канале, не превзойдет y .

В зависимости от конкретного выражения функций распределения Φ и B , относящихся соответственно к входному потоку заявок и процессу их обслуживания, различают две основные модели массового обслуживания: марковскую и немарковскую.

Модель массового обслуживания называется марковской, если входной поток требований на обслуживание оказывается процессом Пуассона

$$P\{X(t) = k\} = p_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad (3.69)$$

а процессу обслуживания свойственно экспоненциальное распределение времени обслуживания

$$p(Y < t) = 1 - e^{-\nu t}. \quad (3.70)$$

Здесь λ — интенсивность потока требований, т. е. математическое ожидание числа требований, поступающих в систему обслуживания в единицу времени; $\nu = \frac{1}{\tau}$ — величина, обратная математическому ожиданию времени обслуживания.

Если поток заявок не пуассоновский или время обслуживания не экспоненциальное, или имеет место то и другое одновременно, то процесс обслуживания относится к более обширному и менее изученному классу немарковских моделей.

С точки зрения очереди заявок на обслуживание различают системы с неограниченной длиной очереди (системы без отказов в обслуживании) и системы с ограниченной длиной очереди (системы с потерями). Если обслуживание характеризуется временем ожидания, то различают системы с неограниченным временем ожидания, с ограниченным временем ожидания обслуживания или ограниченным временем пребывания в системе.

В системе с ограниченным временем пребывания требование теряется, если время его ожидания в очереди превышает допустимое. В этой системе каждое требование, поступающее в систему обслуживания, может находиться в ней не более некоторого допустимого времени, включающего время ожидания в очереди и время обслуживания.

С точки зрения дисциплины обслуживания различают обслуживание в порядке поступления требований и обслуживание с преимуществом. Среди систем массового обслуживания с преимуществом выделяются системы с относительным и абсолютным приоритетом. В систе-

ме с относительным приоритетом появление требования старшего ранга не прерывает начатого обслуживания. В системах с абсолютным приоритетом обслуживание прерывается немедленно при поступлении требования старшего ранга. Возможны также разнообразные комбинации дисциплины обслуживания.

Основной целью исследования системы массового обслуживания является [19], как правило, определение эффективности функционирования системы. При этом в качестве критериев эффективности наиболее часто используют следующие характеристики: вероятность обслуживания требований или вероятность отказа в обслуживании в произвольный момент времени; пропускная способность системы; степень загрузки системы; среднее время ожидания обслуживания и т. д.

После краткого ознакомления с основными понятиями теории массового обслуживания рассмотрим возможности ее применения к инженерно-психологическим исследованиям деятельности оператора.

Прежде всего очевидно, что объект контроля и управления должен быть отождествлен с объектом обслуживания, а оператор — с устройством или аппаратом обслуживания. При этом под заявками следует понимать ситуации (задачи), которые требуют каких-то действий оператора, или однозначно соответствующие им состояния индикационных устройств. Тогда моменты поступления заявок в систему обслуживания можно отождествить с моментом перехода индикатора из одного различного состояния в другое. Располагаемое время обслуживания каждой заявки может определяться в одних случаях временем экспозиций различных состояний индикатора, в других — временем существования ситуаций, в течение которого оператор должен принять решение; располагаемое время определяется характеристиками объекта управления. Отсюда следует, что фактическое время обслуживания заявки оператором не должно превосходить располагаемое, т. е. оператор должен рассматриваться как система массового обслуживания с ограниченным временем ожидания в очереди или пребывания в системе.

Далее, поскольку в любой момент времени индикатор может иметь не более одного различного состояния, то входной поток заявок обладает свойством ординарности. Строго говоря, в силу ряда причин (влияние выходного потока системы на входной поток, корреляции заявок различного типа и т. д.) свойства стационарности и ординарности потока заявок для оператора не соблюдаются. Однако с достаточной для практики точностью потоки заявок различного типа, отличающиеся содержательным аспектом обслуживания, а значит, как правило, временем обслуживания, рассматриваются как простейшие.

В отличие от классических задач теории массового обслуживания, в которых исчерпывающей характеристикой является закон распределения времени обслуживания, при изучении деятельности оператора необходимо дополнительно учитывать содержательную сторону процесса обслуживания. Эта проблема тесно связана с сознательным, творческим характером деятельности человека, а также с его способностями к адаптации и обучению. В терминах теории массового

обслуживания эта проблема непосредственно касается вопросов выбора оператором оптимальной стратегии обслуживания; установления приоритетов обслуживания для различных типов заявок; установления дисциплины очереди среди заявок одного и того же приоритета и т. д.

Таким образом, человек-оператор может быть интерпретирован как одноканальная система массового обслуживания [50] со следующими специфическими особенностями:

1) входящий поток заявок включает потоки заявок нескольких типов, каждый из которых является простейшим, имеет свою интенсивность, время обслуживания и допустимое время ожидания в очереди (пребывания в системе);

2) обслуживание заявок является приоритетным обслуживанием;

3) система массового обслуживания является системой с потерями, время ожидания в очереди (пребывания в системе) ограничено;

4) закон распределения времени обслуживания, как правило, не является экспоненциальным, наиболее часто время обслуживания распределено по нормальному закону или по закону Эрланга k -го порядка;

5) допустимое время ожидания в очереди (пребывания в системе) заявки случайно и определяется особенностями объекта управления;

6) закон распределения допустимого времени ожидания (пребывания) обычно принимается равновероятным или для упрощения задачи допустимое время ожидания (пребывания) принимается постоянным.

Учитывая эти особенности решения задач оператором, рассмотрим три наиболее типичные модели массового обслуживания с точки зрения описания деятельности оператора АСУ [29].

Простейшая модель массового обслуживания. Пусть каждый i -й поток заявок оператору описывается характеристиками: интенсивностью потока λ_i ($i = 1, 2, \dots, n$); функцией распределения длительности обслуживания $B_i(\tau)$ с числовыми характеристиками $\bar{\tau}$ и σ_i ; функцией распределения допустимой длительности ожидания начала обслуживания $F_i(t)$ со средним значением $\bar{t}_{\text{ож}}^{\text{доп}}$; важностью заявки C_i .

Пусть далее величины $\bar{\tau}_i$ и $\bar{t}_{\text{ож}}^{\text{доп}}$ слабо зависят от индекса, а важность задачи примерно одинакова. Тогда работу оператора можно рассматривать как обслуживание им суммарного потока заявок, который, будучи образован суперпозицией простейших потоков, сам является простейшим. Обслуживание заявок суммарного потока произойдет в порядке их поступления, а его характеристики определяются соотношениями

$$\Lambda = \sum_{i=1}^n \lambda_i; \quad \nu = \frac{1}{\bar{t}_{\text{ож}}^{\text{доп}}}; \quad \rho = \Lambda \bar{\tau}; \quad (3.71)$$

$$\bar{\tau} = \frac{1}{\Lambda} \sum_{i=1}^n \lambda_i \bar{\tau}_i; \quad \bar{t}_{\text{ож}} = \frac{1}{\Lambda} \sum_{i=1}^n \lambda_i \bar{t}_{\text{ож}}^{\text{доп}}.$$

Известно, что вероятность обслуживания заявки P_0 в стационарном режиме функционирования системы массового обслуживания

определяется лишь средними значениями длительности обслуживания $\bar{\tau}$ и допустимого времени ожидания $t_{\text{ож}}^{\text{доп}}$. Поэтому можно рассчитать искомую вероятность P_0 по формуле, полученной для экспоненциальных распределений величин τ и $t_{\text{ож}}^{\text{доп}}$ и имеющей вид

$$P_0 = 1 - \frac{\nu}{\Lambda} \frac{\rho \sum_{s=1}^{\infty} \frac{s \rho^s}{\prod_{m=1}^s (1 + m\nu\bar{\tau})}}{1 + \rho + \rho \sum_{s=1}^{\infty} \frac{\rho^s}{\prod_{m=1}^s (1 + m\nu\bar{\tau})}}. \quad (3.72)$$

Для практического расчета величины P_0 в соответствии с формулой (3.72) существуют таблицы определения вероятности системы массового обслуживания с ожиданием, вход в которые производится по величинам ρ и $\alpha = \nu\bar{\tau}$.

Для определения пропускной способности оператора Λ строится график зависимости $P_0(\rho)$ (рис. 31) и находится точка $\rho_0 = \Lambda_0 \bar{\tau}$, в которой выполняется условие $P_0(\rho) = P_0^{\text{доп}}$. Затем вычисляется пропускная способность Λ_0 по очевидному соотношению

$$\Lambda_0 = \frac{P_0}{\bar{\tau}}. \quad (3.73)$$

Модель массового обслуживания с относительными приоритетами. Пусть характер задач, решаемых оператором, таков, что допустимые времена ожидания решения $t_{\text{ож}}^{\text{доп}}$ для разных типов задач существенно различаются. Тогда оператор в первую очередь решает задачи с малыми значениями времени ожидания, т. е. при обслуживании устанавливаются относительные приоритеты.

Используем для оценки вероятности решения задач i -го типа P_{0i} систему массового обслуживания с относительными приоритетами. Очевидно, всегда справедливо соотношение

$$P_{0i} = \frac{\lambda_i^0}{\lambda_i} = \frac{\rho_i^0}{\rho_i}, \quad (3.74)$$

где λ_i^0 — интенсивность потока обслуженных заявок i -го типа ($i = 1, 2, \dots, n$).

Величина λ_i^0 определяется из соотношения

$$\lambda_i^0 = \lambda_i P_{0i}(\Lambda^0, t_{\text{ож}}^{\text{доп}}), \quad (3.75)$$

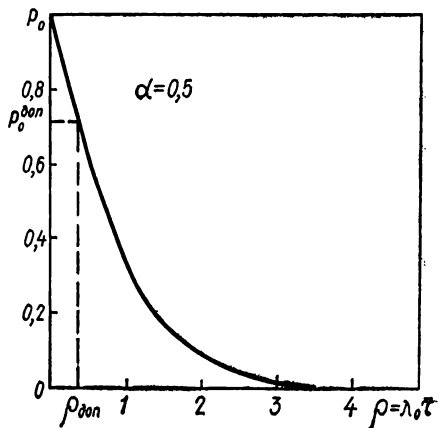


Рис. 31

где

$$P_{0i}(\Lambda^0, t_{i0ж}^{\text{доп}}) = P(t_{i0ж} \leq t_{i0ж}^{\text{доп}});$$

$t_{i_{\text{ож}}}$ — среднее время ожидания начала обслуживания заявки i -го типа; $\Lambda^0 = |\lambda_1^0, \dots, \lambda_n^0|$ — вектор-строка потоков обслуженных заявок.

Для произвольного вида функций распределения времени обслуживания заявок $B_i(\tau)$ среднее значение длительности ожидания заявки i -го потока вычисляется по формуле [24]

$$t_{\text{ioк}} = \frac{\sum_{r=1}^n \lambda_r (\bar{v}_r^2 + \sigma_r^2)}{2 \left(1 - \sum_{r=1}^{i-1} \rho_r\right) \left(1 - \sum_{r=1}^i \rho_r\right)}, \quad (3.76)$$

где $\rho_r = \lambda_r \bar{\tau}_r$.

Поскольку длительность ожидания заявок ограничена, то часть их теряется. Поток обслуженных заявок образуется прореживанием входного потока. Заявка i -го потока с вероятностью p_i обслуживается, а с вероятностью $q_i = 1 - p_i$ теряется. Прореженный поток является простейшим с параметром $\lambda_i^0 = p_i \lambda_i$. Потерянные заявки не влияют на длительность ожидания оставшихся заявок, т. е. $t_{i\text{ож}} = t_{i\text{ож}}^0$.

Задача состоит в том, чтобы от исходной системы с ограниченными временем ожидания $t_{i\text{ож}}^{\text{доп}}$ и потоком заявок λ_i перейти к эквивалентной системе с параметрами λ_i^0 и $t_{i\text{ож}}^0$. Практически эта задача сводится к совместному решению системы уравнений:

$$\begin{aligned} \lambda_1^0 &= f_1(\Lambda^0, t_{1\text{oж}}^0); \\ . &. \\ \lambda_n^0 &= f_n(\Lambda^0, t_{n\text{oж}}^0). \end{aligned} \quad (3.77)$$

Для решения данной системы можно использовать метод последовательных приближений. Условия сходимости процесса последовательных приближений хорошо известны. В случае медленной сходимости следует применять более быстрые итерационные методы типа метода Ньютона.

Предположим, что длительность обслуживания требования имеет экспоненциальное распределение, т. е.

$$B_i(\tau) = 1 - e^{-v_i \tau}. \quad (3.78)$$

Тогда длительность ожидания тех заявок, которые застают систему занятой, имеет экспоненциальное распределение с параметром $\gamma_i = \nu_i - \lambda_i$. Таким образом, имеем

$$P(t_{i_{\text{ож}}} > t) = \rho_i e^{-(\nu_i - \lambda_i)t}. \quad (3.79)$$

Учитывая, что в системе с относительными приоритетами влияние младших приоритетов на старшие значительно слабее, чем наоборот, решение задачи следует начинать с самого старшего приоритета, фиксируя параметры других потоков.

Первый цикл приближения начинается с $i = 1$. При $i = 1$ имеем

$$\begin{aligned}\lambda_1^0 &= \lambda_1 (1 - q_1); \\ \gamma &= \frac{1}{\tau_1} (1 - \bar{\tau}_1 \lambda_1); \\ q_1 &= \left(\lambda_1^0 \bar{\tau}_1 + \sum_{r=2}^n \lambda_r \bar{\tau}_r \right) e^{-\gamma, t_{1\text{ож}}^{\text{доп}}}.\end{aligned}\quad (3.80)$$

Таким образом, процесс последовательных приближений сводится к решению трансцендентного уравнения вида

$$\frac{\lambda_1^0}{\lambda_1} = 1 - \left(\lambda_1^0 \bar{\tau}_1 + \sum_{r=2}^n \lambda_r \bar{\tau}_r \right) e^{-(1 - \bar{\tau}_1 \lambda_1^0) t_{1\text{ож}}^{\text{доп}} \bar{\tau}_1}.\quad (3.81)$$

Переходя к относительным (безразмерным) величинам, получаем следующее уравнение:

$$\frac{\rho_1^0}{\rho_1} = 1 - \left(\rho_1^0 + \sum_{r=2}^n \rho_r \right) e^{-(1 - \rho_1^0) \varepsilon_{01}},\quad (3.82)$$

$$\text{где } \varepsilon_{01} = \frac{t_{1\text{ож}}^{\text{доп}}}{\tau_1}.$$

Решить уравнение (3.82) можно графическим методом. Он состоит в непосредственном определении координат точки пересечения луча

$y = \frac{\rho_1^0}{\rho}$ с кривой вида

$$y = 1 - (\rho_1^0 + \rho_\Sigma) e^{-(1 - \rho_1^0) \varepsilon_{01}},\quad (3.83)$$

$$\text{где } \rho_\Sigma = \sum_{r=2}^n \rho_r.$$

Для сокращения расчетов кривая строится в области практически ожидаемой загрузки ρ .

Для потоков других приоритетов ($i > 1$) вместо уравнения (3.82) имеем

$$\frac{\rho_i^0}{\rho_i} = 1 - \left(\sum_{r=1}^{i-1} \rho_r^0 + \rho_i^0 + \sum_{r=i+1}^n \rho_r \right) e^{-\left(1 - \sum_{r=1}^{i-1} \rho_r^0\right) \left(1 - \sum_{r=1}^{i-1} \rho_r^0 + \rho_i^0\right) \varepsilon_{0i}}.\quad (3.84)$$

Решение уравнения производится аналогично. Закончив первый цикл приближения, получаем набор величин $\rho_{i(1)}^0$ и $P_{i(1)}^0 = \frac{\rho_{i(1)}^0}{\rho_i}$.

Второй цикл приближения производится по той же методике, но с использованием величин $\rho_i = \rho_{i(1)}^0$. Решение уравнения завершается, если изменением величин $\rho_{i(k)}$ по сравнению с величинами $\rho_{i(k-1)}^0$ можно пренебречь. Во многих случаях практически достаточно провести два цикла.

Как показано в работе [2], вероятность обслуживания заявок в системах рассматриваемого класса существенно зависит лишь от сред-

них значений времени обслуживания $\bar{\tau}_i$ и допустимого времени ожидания $t_{i\text{ож}}^{\text{доп}}$ и мало зависит от вида функций распределения этих величин. Поэтому данная методика может быть использована для оценки работы операторов и при произвольном распределении времени обслуживания и допустимого времени ожидания.

Модель массового обслуживания с абсолютными приоритетами. Оператор обслуживает две группы требований, причем требования первой группы, имеющей n простейших потоков с интенсивностями ν_j ($j = 1, 2, \dots, m$), обладают при обслуживании абсолютным приоритетом по отношению к требованиям второй группы, имеющей m простейших потоков с интенсивностями ν_j ($j = 1, 2, \dots, m$).

Внутри каждой из групп установлен относительный приоритет, убывающий в порядке возрастания номера потока. Время обслуживания требований в общем случае задается произвольным законом распределения $B_i(t)$ для требований i -го потока и $B_j(t)$ — для j -го потока.

Критерием оценки [25] своевременного решения задачи любого i -го (j -го) типа является выполнение условия

$$t_{i(j)\text{ож}} + \tau_{i(j)} \leq t_{i(j)\text{доп}}, \quad (3.85)$$

где $t_{i(j)\text{доп}}$ — допустимое время пребывания задач i -го (j -го) типов в системе обслуживания.

Для первой группы, пренебрегая требованиями второй группы, не влияющими на их обслуживание, имеем

$$t_{i\text{ож}} = \frac{\sum_{r=1}^n \lambda_r (\bar{\tau}_r^2 + \sigma_r^2)}{2 \left(1 - \sum_{r=1}^{i-1} \rho_r \right) \left(1 - \sum_{r=1}^i \rho_r \right)}, \quad (3.86)$$

$$t_{i\text{пр}} = \bar{t}_{i\text{ож}} + \bar{\tau}_i, \quad \rho_i = \lambda_i \bar{\tau}_i.$$

Чтобы определить величины $t_{j\text{ож}}$ и $t_{j\text{пр}}$ для требований второй группы, необходимо привести исходную систему к некоторой системе массового обслуживания со следующими свойствами:

1) в длительность обслуживания требования j -го потока включаются все периоды обслуживания требований первой группы (прерывание обслуживания);

2) периоды занятости системы обслуживанием только требований первой группы исключаются;

3) в остальных периодах занятости, начавшихся с обслуживания требований первой группы, начальный период включается во время обслуживания требования второй группы, которое обслуживается первым в данном периоде занятости.

Очевидно, в приведенной системе обслуживания время ожидания начала обслуживания и время пребывания в системе требований второй группы могут быть определены по формулам, аналогичным (3.85)

и (3.86). Поэтому решение задачи сводится к нахождению величин τ_j и σ_j^2 в приведенной системе [19].

Случайную величину приведенного времени обслуживания требований j -го потока можно записать следующим образом:

$$t_j = t_{j0} + P_{jk} \sum_{\xi=1}^k t_{\xi \text{пр}} + r_j t_y, \quad (3.87)$$

где t_{j0} — исходное время обслуживания с распределением $B_j(t)$; P_{jk} — вероятность получения ровно k прерываний обслуживания j -го требования требованиями первой группы; $t_{\xi \text{пр}}$ — случайная величина длительности ξ -го прерывания; r_j — вероятность увеличения времени обслуживания j -го требования за счет третьей особенности приведенной системы; t_y — увеличение длительности обслуживания, имеющее, очевидно, то же распределение, что и величина $t_{\text{пр}}$.

Распределение длительности прерывания соответствует распределению периода занятости системы обслуживанием только требований первой группы.

Функция плотности распределения периода занятости обслуживанием требований первой группы $g(t)$ и основные числовые характеристики равны соответственно:

$$g(t) = \sum_{k=0}^{\infty} \frac{e^{-\Lambda t} (\Lambda t)^k}{(k+1)!} f_k(t),$$

$$\mu = \frac{\mu_1}{1 - \rho_1}; \quad \sigma_2 = \frac{\sigma_1^2 + \rho_1 \mu_1^2}{(1 - \rho_1)^2}, \quad (3.88)$$

где $f_k(t)$ — функция плотности распределения суммы k независимых длительностей обслуживания обобщенного требования первой группы с распределением

$$B(t) = \frac{1}{\Lambda} \sum_{i=1}^n \lambda_i B_i(t); \quad (3.89)$$

$\mu_1 = \frac{1}{\Lambda} \sum_{i=1}^n \lambda_i \bar{\tau}_i$ — математическое ожидание распределения $B(t)$; $\sigma_1 = \frac{1}{\Lambda} \sum_{i=1}^n \lambda_i \sigma_i^2$ — дисперсия распределения $B(t)$; $\Lambda = \sum_{i=1}^n \lambda_i$ — суммарный поток требований первой группы; $\rho_1 = \Lambda \mu_1$ — загрузка системы требованиями потоков первой группы.

На практике для инженерных расчетов распределение периода занятости $g(t)$ целесообразно аппроксимировать подходящей функцией распределения из числа наиболее распространенных, потребовав совпадения первых двух моментов [34].

Обслуживание произвольного требования второй группы происходит в интервалах между очередными требованиями. Эти интервалы имеют показательное распределение, поскольку суммарный поток требований первой группы является простейшим. Поэтому вероятность того, что за время обслуживания j -го требования произойдет ровно

k пребывания, определяется по формуле

$$P_{jk} = \int_0^{\infty} \frac{(\Delta t)^k}{k!} e^{-\Delta t} dB_j(t), \quad (3.90)$$

где $\sum_{k=0}^{\infty} P_{jk} = 1$.

Общая суммарная длительность пребывания j -го требования есть сумма n_j слагаемых, каждое из которых имеет одинаковое распределение $g(t)$, а число n_j — распределение (3.90).

Математическое ожидание и дисперсия общей длительности пребывания соответственно записываются в виде

$$\mu_{\text{лп}} = K_j \mu_j, \quad \sigma_{\text{лп}}^2 = \sigma_{kj}^2 \mu_j^2 + \bar{k}_j \sigma_j^2. \quad (3.91)$$

Увеличение длительности обслуживания t_y имеет тоже распределение $y(t)$. Общее увеличение длительности обслуживания требования j -го потока равно сумме одинаково распределенных слагаемых, число которых N_j с вероятностью $r_j p_{jk}$ равно $(1+k_j)$ и с вероятностью $(1-r_j) p_{jk}$ равно K_j .

Можно показать, что математическое ожидание \bar{N}_j и дисперсия $\sigma_{N_j}^2$ соответственно определяются по формулам

$$\bar{N}_j = \bar{K}_j + \bar{r}_j, \quad \sigma_{N_j}^2 = \sigma_{kj}^2 + r_j(1-r_j). \quad (3.92)$$

Исходя из этого, получим окончательные формулы для математического ожидания $\bar{\tau}_j$ и дисперсии σ_j^2 приведенного времени обслуживания

$$\begin{aligned} \bar{\tau}_j &= \bar{\tau}_{j0} + (\bar{K}_j + r_j) \mu_j; \\ \sigma_j^2 &= \sigma_{j0}^2 + [\sigma_{kj}^2 + r_j(1-r_j)] \mu_j^2 + (\bar{K}_j + r_j) \sigma_j^2, \end{aligned} \quad (3.93)$$

где τ_{j0} и σ_{j0} — математическое ожидание и дисперсия исходного распределения $B_j(t)$.

Для расчета вероятности r_j определим предварительно вероятность составного события, заключающегося в том, что очередной период занятости системы начнется с обслуживания произвольного требования первой группы; при этом первым обслуженным требованием второй группы будет требование j -го потока.

Последнее имеет место, когда за время обслуживания требований первой группы в систему поступило хотя бы одно требование j -го потока и не поступило ни одного требования второй группы более высокого относительного приоритета.

Очевидно,

$$r_j = \frac{\Lambda}{\Lambda + V} (1 - P_{0j}) \prod_{i=1}^{j-1} P_{0i}, \quad (3.94)$$

где $V = \sum_{j=1}^m v_j$; $P_{0j} = \int_0^{\infty} e^{-v_j t} g(t) dt$.

Введем плотность распределения длительности составного периода занятости $\gamma(t)$ и определим ее аналогично $g(t)$, но для суммарного потока требований с интенсивностью, равной $\Lambda + V$.

Поскольку за время, определяемое функцией $\gamma(t)$, в систему может поступить в общем случае S требований, а увеличивается время обслуживания лишь первого из них, то вероятность r_j может быть определена следующим образом:

$$r_j = \frac{\Lambda(1 - P_{0j})}{\Lambda + V} \prod_{r=1}^{j-1} P_{0r} \sum_{S=1}^{\infty} \frac{P_{Sj}}{S}, \quad (3.95)$$

где

$$P_{Sj} = \int_0^{\infty} \frac{(v_j t)^S}{S!} e^{-v_j t} \gamma(t) dt. \quad (3.96)$$

Заметим, что если функция $g(t)$ и $\gamma(t)$ аппроксимируются законом Эрланга вида

$$f(t) = \frac{\alpha^m}{(m-1)!} e^{-\alpha t} t^{m-1}, \quad (3.97)$$

то расчетные формулы существенно упрощаются, поскольку в этом случае

$$P_{Sj} = \frac{(1 - \beta_j)^m}{(m-1)!} \beta_j^S \frac{(S+m-1)!}{S!}, \quad (3.98)$$

где

$$\beta_j = \frac{v_j}{\alpha + v_j}.$$

При $m \geq 2$ для вычисления величин $R_j = \sum_{S=1}^{\infty} \frac{P_{Sj}}{S}$ можно использовать ряд

$$\sum_{S=0}^{\infty} \beta^S S^r = \frac{\beta(1 + \beta^{r-1})}{(1 + \beta)^{r-1}} + \frac{A_{r,2} \beta^r (1 - \beta^{r-2})}{(1 - \beta)^{r+2}}, \quad (3.99)$$

где $r \geq 2$; $A_{r,2} = 2^r - (r+1)$ — числа Эйлера.

Таким образом, изложенный метод расчета показателей функционирования многоприоритетной системы массового обслуживания с ожиданием практически позволяет получить введенные показатели при достаточно слабых ограничениях, налагаемых на функции распределения.

В заключение отметим, что применение методов теории массового обслуживания позволяет решать многие вопросы организации деятельности операторов в системе. К их числу относятся: определение необходимого числа операторов, организация взаимодействия операторов, расчет допустимой частоты поступления сигналов на устройства отображения и т. д.

Алгоритмическая модель деятельности оператора. Алгоритм работы оператора представляет собой логическую последовательность

действий оператора при решении задачи. Составление алгоритма предполагает прежде всего определение задачи в общем виде, а затем разбиение процесса ее решения на элементарные операции переработки информации и логические условия.

Элементарные операции представляют собой предел «дробления» деятельности оператора. Такими элементарными единицами могут быть оперативные единицы [17], которыми являются образы восприятия, использующиеся как нечто цельное операторами в своей работе: понятия, суждения и даже сложные движения, протекающие целостно под влиянием одного «пускового импульса». Состав и объем таких единиц устойчив лишь относительно, в пределах данного рода деятельности. Основным признаком элементарной операции является четкий критерий выполнения — начало и конец. В качестве примеров элементарных операций можно привести поиск объекта на экране, восприятие информации формуляра, ввод команды и т. д.

Анализируя работу оператора, всегда можно выделить некоторые совокупности элементарных операций, объединенные общей целью. Такая совокупность операций обычно называется задачей. Задача имеет более высокий порядок классификации, чем элементарная операция.

Конечной целью всякой деятельности является получение полезного эффекта. Конечная цель достигается поэтапно, через решение частных задач. Поэтому можно говорить не только об алгоритме решения отдельной задачи, но и об алгоритме деятельности оператора.

Существуют различные формы представления алгоритма: описательная, операторная, структурная, граф-схема, блок-схема и др. [17].

Описательная форма алгоритма представляет собой последовательное изложение всех действий оператора и их порядка в зависимости от тех или иных условий. Всякая подробная инструкция по эксплуатации или по боевой работе — это, по существу, такой алгоритм. Разница между инструкцией и описательной формой алгоритма, составляемой в процессе психофизиологического анализа деятельности оператора, заключается в том, что в последнем случае дробление задачи производится обязательно до выявления отдельных «оперативных единиц». Этот методический прием является переходным между профессиографией и составлением логических схем алгоритмов в операторной форме.

Членами логической схемы алгоритма в операторной форме являются операторы (0 — элементарные действия) и логические условия (1), которые в зависимости от того, выполняются они ($l = 1$) или не выполняются ($l = 0$), определяют тот или иной порядок срабатывания операторов. Операторы обозначаются большими латинскими буквами, а логические условия — малыми.

Введем некоторые дополнения при описании алгоритма. Будем различать операторы считывания показаний с приборов, поиска информации, восприятия команды нажатия кнопки, отдачи команды и запоминания, обозначая их соответственно индексами α , ϵ , μ . Кроме того, в ряде случаев требуется специальный оператор задержки τ .

Им будут обозначаться временные задержки в работе, например ожидание команды или окончание срабатывания автомата.

В табл. 7 дано описание одного простого алгоритма действий оператора.

В операторной форме алгоритм будет выглядеть следующим образом:

$$A^{(\alpha)} q^{(\alpha)} \uparrow R^{(e)} \downarrow D^{(e)}.$$

Структурная схема алгоритма представляет собой словесное описание алгоритма работы оператора с указанием логических связей между этапами его работы.

Операторная форма записи алгоритма компактна, но не наглядна и не позволяет проводить количественный анализ алгоритма. Структурная схема алгоритма хорошо обозрима, но громоздка и также позволяет получать количественные характеристики алгоритма [17].

Т а б л и ц а 7

| Члены алгоритма | Описание членов алгоритма | Символы |
|-----------------|--------------------------------|----------------|
| O_1 | Считывание показаний с прибора | $A^{(\alpha)}$ |
| I_1 | Показатель больше 100 | $q^{(\alpha)}$ |
| O_2 | Нажим кнопки | $R^{(e)}$ |
| O_3 | Доклад диспетчеру | $D^{(e)}$ |

Представление алгоритма в виде граф-схемы обладает тем неоспоримым преимуществом, что наряду с наглядностью позволяет проводить количественный анализ алгоритма.

Для построения граф-схемы алгоритма воспользуемся некоторыми понятиями теории сетей [80].

Пусть имеется конечное множество $V = \{v_1, \dots, v_n\}$ вершин и конечное множество $W = \{w_1, \dots, w_n\}$ весов вершин. Если установить однозначное соответствие между множествами V и W , то множество пар (v_i, w_i) представляет собой множество взвешенных вершин.

Пусть задано множество B , элементами которого являются упорядоченные пары элементов множества V , называемые направленными ветвями: $B = \{(v_i, w_i)\}$. Геометрической интерпретацией множества являются направленные ветви, соединяющие вершины v_i и v_j .

Обозначим через b_1, b_2, \dots, b_m элементы множества B . Далее рассмотрим конечное множество X весов ветвей. Элементы этого множества ставим в соответствие ветвям v_i, v_j и v_i, v_k , $j \neq k$, где v_i — вершина с двумя выходами.

Обозначим через γ_j элементы множества X . Тогда множество пар b_j, γ_j представляет собой множество взвешенных ветвей.

Совокупность множеств вершин и ветвей называется сетью. Это понятие сети совпадает с понятием взвешенного ориентированного графа.

Введем некоторые дополнительные определения.

Направленный путь длины r от вершины v_i к вершине v_j есть множество r ветвей вида $(v_i, v_{k1}), \dots, (v_{kr-1}, v_j)$, если в сети существуют такие ветви.

Если $i = j$, то путь замкнут и называется циклом длины r . Если все вершины $v_i, v_{k1}, \dots, v_{kr-1}, v_j$ различны, то путь называется правиль-

ным. Если $v_i = v_j$, то никакие другие вершины не совпадают, путь называется правильным циклом.

Алгоритм работы оператора удобно представлять в виде ориентированного графа. Граф алгоритма является графическим изображением последовательности выполнения отдельных операций алгоритма [17].

Соотнесем с каждой вершиной графа G_n трехмерный вектор $W_i = \{m_i, D_i, p_i\}$, составляющими которого являются: среднее время выполнения i -й операции m_i ; дисперсия времени выполнения D_i ; вероятность безошибочного выполнения p_i .

Сопоставим теперь каждой ветви логического оператора определенное число γ_i — вероятность перехода по этой ветви.

Совокупность множеств взвешенных вершин и ветвей графа совместно с множеством направленных ветвей образует сеть (граф-схему) алгоритма решения задачи оператором.

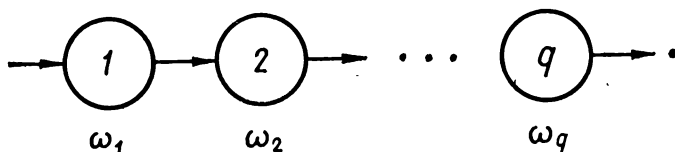
Временные и вероятностные характеристики алгоритма (среднее время m_A и его дисперсия D_A , вероятность безошибочного выполнения алгоритма P_A) могут быть получены на основе характеристик элементарных операций путем представления алгоритма в виде граф-схемы и использования методов ее преобразования.

Рассмотрим основные соотношения, позволяющие производить преобразования граф-схемы алгоритма.

1. Объединение путей без разветвления. Путь m_1, \dots, m_q назовем путем без разветвлений, если в вершинах m_1, \dots, m_{q-1} начинается и оканчивается в точности по одной ветви.

Удобно заменять путь разветвлений одной ветвью m_1, m_q . При этом происходит существенное сокращение размеров сети.

Граф пути без разветвлений имеет следующий вид:



Операция последовательного приведения преобразует путь без разветвлений в эквивалентную ветвь с одной вершиной. Определим составляющие веса вершины $\omega_s = \{m_s, D_s, p_s\}$. Очевидно, в предположении независимости выполнения отдельных операций справедливы соотношения

$$\begin{aligned} m_s &= \sum_{i=1}^q m_i, \\ D_s &= \sum_{i=1}^q D_i, \\ P_s &= \prod_{i=1}^q p_i. \end{aligned} \tag{3.100}$$

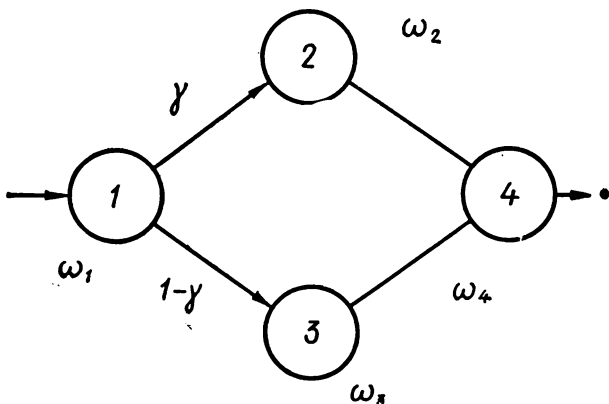
Если $g_i(t)$ — плотность распределения времени выполнения i -й операции, то путем многократного применения оператора свертки получим эквивалентную плотность распределения, т. е. справедливо соотношение

$$g_s(t) = \int_{-\infty}^{\infty} g_i(\tau) g_{i-1}(t - \tau) d\tau, \quad (3.101)$$

где $i = 1, 2, \dots, q$.

При достаточно большом q ($q > 5$) и примерно одинаковом порядке величин m_i согласно центральной предельной теореме теории вероятностей эквивалентный закон распределения можно считать нормальным. Числовые характеристики закона определяются в соответствии с выражениями (3.100).

2. Объединение путей с разветвлениями. Граф пути с разветвлениями выглядит следующим образом:



Операция параллельного приведения преобразует путь с разветвлениями в эквивалентную ветвь с одной вершиной. Составляющие веса вершины находятся из очевидных выражений

$$\begin{aligned} m_s &= m_1 + \gamma m_2 + (1 - \gamma) m_3 + m_4, \\ D_s &= D_1 + \gamma D_2 + (1 - \gamma) D_3 + D_4, \\ P_s &= p_1 p_4 [\gamma p_2 + (1 - \gamma) p_3]. \end{aligned} \quad (3.102)$$

Найдем выражение для эквивалентной функции распределения времени $F_s(t)$:

$$F_s(t) = P(T \leq t) = \gamma P(T \leq t_1 + t_2 + t_4) + (1 - \gamma) P(T \leq t_1 + t_3 + t_4).$$

Очевидно, что

$$\begin{aligned} P(T \leq t_1 + t_2 + t_4) &= F_{1,2,4}(t), \\ P(T \leq t_1 + t_3 + t_4) &= F_{1,3,4}(t), \end{aligned}$$

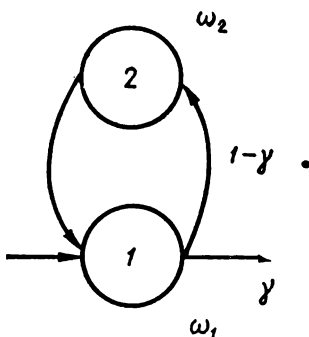
где $F_{1,2,4}(t)$, $F_{1,3,4}(t)$ — функции распределения времени выполнения операций 1, 2, 4-й или 1, 3, 4-й соответственно.

Функции $F_{1,2,4}(t)$ и $F_{1,3,4}(t)$ могут быть легко найдены с помощью формулы (3.101):

$$F_3(t) = \gamma F_{1,2,4}(t) + (1 - \gamma) F_{1,3,4}(t),$$

$$f_3(t) = \frac{dF_3(t)}{dt}. \quad (3.103)$$

3. Сведение контуров с одним выходом. Ветвь, связывающая вершину саму с собой, называется петлей. Когда в петлю включается более чем одна ветвь, ее называют контуром. Граф контура имеет следующий вид:



Для определения эквивалентных характеристик контура с одним выходом составим вспомогательную табл. 8.

Т а б л и ц а 8

| Число циклов | m_3 | D_3 | P_3 | Вероятности циклов |
|--------------|----------------------|----------------------|-------------------|-------------------------|
| 0 | m_1 | D_1 | p_1 | γ |
| 1 | $2m_1 + m_2$ | $2D_1 + D_2$ | $p_1^2 p_2$ | $\gamma (1 - \gamma)$ |
| 2 | $3m_1 + 2m_2$ | $3D_1 + 2D_2$ | $p_1^3 p_2^2$ | $\gamma (1 - \gamma)^2$ |
| ... | | | | |
| j | $j(m_1 + m_2) + m_1$ | $j(D_1 + D_2) + D_1$ | $p_1^{j+1} p_2^j$ | $\gamma (1 - \gamma)^j$ |

Очевидно, можно записать

$$m_3 = \sum_{j=0}^{\infty} \gamma (1 - \gamma)^j [m_1 + j(m_1 + m_2)],$$

$$D_3 = \sum_{j=0}^{\infty} \gamma (1 - \gamma)^j [D_1 + j(D_1 + D_2)], \quad (3.104)$$

$$P_3 = \sum_{j=0}^{\infty} \gamma (1 - \gamma)^j p_1^{j+1} p_2^j.$$

Преобразуем формулы (3.104) к виду, более удобному для использования.

Выражение для величины m_s запишем в следующем виде:

$$m_s = \gamma m_1 \sum_{j=0}^{\infty} (1 - \gamma)^j + (m_1 + m_2) \gamma \sum_{j=0}^{\infty} j (1 - \gamma)^j. \quad (3.105)$$

Для любых логических условий справедливо неравенство

$$0 \leq 1 - \gamma \leq 1.$$

Поэтому при преобразовании сумм, стоящих в скобках, получаем

$$\sum_{j=0}^{\infty} (1 - \gamma)^j = 1/\gamma; \quad \sum_{j=0}^{\infty} j (1 - \gamma)^j = (1 - \gamma)/\gamma^2. \quad (3.106)$$

Подставляя (3.106) в (3.105), окончательно имеем

$$m_s = (m_1 + m_2)/\gamma - m_2. \quad (3.107)$$

Аналогичные преобразования для дисперсии позволяют записать

$$D_s = (D_1 + D_2)/\gamma - D_2. \quad (3.108)$$

Далее, очевидно, выражение (3.104) для вероятности можно представить так:

$$P_s = \gamma p_1 \sum_{j=0}^{\infty} [p_1 p_2 (1 - \gamma)]^j. \quad (3.109)$$

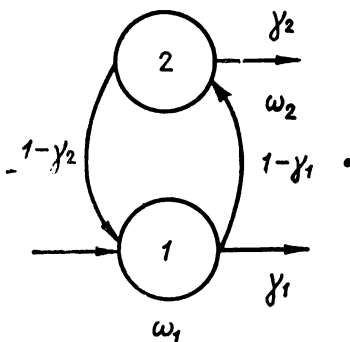
Поскольку всегда $0 \leq p_1 p_2 (1 - \gamma) \leq 1$, то

$$\sum_{j=0}^{\infty} [p_1 p_2 (1 - \gamma)]^j = \frac{1}{1 - p_1 p_2 (1 - \gamma)}. \quad (3.110)$$

В результате подстановки (3.110) в (3.109) получаем

$$P_s = \frac{\gamma p_1}{1 - p_1 p_2 (1 - \gamma)}. \quad (3.111)$$

4. Сведение контуров с двумя выходами. Граф-схема контура с двумя выходами имеет следующий вид:



Для получения эквивалентных характеристик контура составим вспомогательную табл. 9.

Усредняя эквивалентные характеристики с учетом возможных вероятностей исходов, получаем

$$\begin{aligned}
 m_s &= \sum_{j=0}^{\infty} \gamma_1 (1 - \gamma_1)^j (1 - \gamma_2)^j [j(m_1 + m_2) + m_1] + \\
 &+ \sum_{j=0}^{\infty} \gamma_2 (1 - \gamma_1)^{j+1} (1 - \gamma_2)^j (j+1)(m_1 + m_2), \\
 D_s &= \sum_{j=0}^{\infty} \gamma_1 (1 - \gamma_1)^j (1 - \gamma_2)^j [j(D_1 + D_2) + D_1] + \\
 &+ \sum_{j=0}^{\infty} \gamma_2 (1 - \gamma_1)^{j+1} (1 - \gamma_2)^j (j+1)(D_1 + D_2), \\
 P_s &= \sum_{j=0}^{\infty} \gamma_1 (1 - \gamma_1)^j (1 - \gamma_2)^j \rho_1^{j+1} \rho_2^j + \sum_{j=0}^{\infty} \gamma_2 (1 - \gamma_1)^{j+1} (1 - \gamma_2)^j \times \\
 &\times (\rho_1 \rho_2)^{j+1}.
 \end{aligned} \quad (3.112)$$

Т а б л и ц а 9

| Число циклов | Выходы | m_s | D_s | P_s | Вероятности циклов |
|--------------|--------|--------------------|--------------------|-------------------------|--|
| 0 | 1 | m_1 | D_1 | ρ_1 | γ_1 |
| | 2 | $m_1 + m_2$ | $D_1 + D_2$ | $\rho_1 \rho_2$ | $(1 - \gamma_1) \gamma_2$ |
| 1 | 1 | $2m_1 + m_2$ | $2D_1 + D_2$ | $\rho_1^2 \rho_2$ | $(1 - \gamma_1)(1 - \gamma_2) \gamma_1$ |
| | 2 | $2(m_1 + m_2)$ | $2(D_1 + D_2)$ | $\rho_1^2 \rho_2^2$ | $(1 - \gamma_1)^2 (1 - \gamma_2) \gamma_2$ |
| j | 1 | $(j+1)m_1 + j m_2$ | $(j+1)D_1 + j D_2$ | $\rho_1^{j+1} \rho_2^j$ | $(1 - \gamma_1)^j (1 - \gamma_2)^j \gamma_1$ |
| | 2 | $(j+1)(m_1 + m_2)$ | $(j+1)(D_1 + D_2)$ | $(\rho_1 \rho_2)^{j+1}$ | $(1 - \gamma_1)^{j+1} (1 - \gamma_2)^j \gamma_2$ |

Используя формулы для свертывания бесконечных рядов, после достаточно простых преобразований получаем

$$\begin{aligned}
 m_s &= \frac{1}{[1 - (1 - \gamma_1)(1 - \gamma_2)]^2} \{m_1 [\gamma_1 + (1 - \gamma_1) \gamma_2] + \\
 &+ m_2 (1 - \gamma_1) [\gamma_2 + \gamma_1 (1 - \gamma_2)]\}, \\
 D_s &= \frac{1}{[1 - (1 - \gamma_1)(1 - \gamma_2)]^2} \{D_1 [\gamma_1 + (1 - \gamma_1) \gamma_2] + \\
 &+ D_2 (1 - \gamma_1) [\gamma_2 + \gamma_1 (1 - \gamma_2)]\}, \\
 P_s &= \frac{\rho_1 [\gamma_1 + (1 - \gamma_1) \gamma_2 \rho_2]}{1 - \rho_1 \rho_2 (1 - \gamma_1)(1 - \gamma_2)}.
 \end{aligned} \quad (3.113)$$

Очевидно, формулы (3.107), (3.108), (3.111) могут быть получены из (3.113), если положить $\gamma_2 = 0$.

Вообще говоря, аналогично можно вывести формулы для преобразования контуров с любым числом выходов. Вывод этих формул достаточно прост, однако окончательные выражения громоздки.

Представление алгоритма работы оператора в виде ориентированного графа и использование правил преобразования графа позволяют решать целый ряд задач исследования деятельности операторов.

§ 4. Методы моделирования. Экспериментальные методы исследования деятельности оператора

Статистическое моделирование деятельности оператора.

Аналитические методы, в том числе методы теории массового обслуживания, используемые для описания деятельности оператора, позволяют проводить исследования в рамках жестких схем и ограничений. При формализации деятельности оператора трудно учесть многие психофизиологические закономерности: процессы памяти, возникновение напряженности и утомления, возникновение ошибок и их исправление, взаимодействие операторов и т. д. [30]. Основным препятствием здесь оказывается сложность непосредственной формализации и математического описания деятельности оператора на базе умозрительного анализа взаимозависимости составляющих эту деятельность операций и задач, тем более что не всегда для этой цели имеются подходящие математические средства.

Возможен другой путь исследования, когда умозрительный анализ позволяет выявить лишь операции, выполняемые человеком, и определять связи между ними. Конструирование деятельности оператора в целом возлагается на ЭВМ. Этот метод исследования получил название метода статистического (имитационного) моделирования.

По сравнению с экспериментальными методами он имеет преимущества, так как позволяет заменить дорогостоящие реальные эксперименты имитацией деятельности оператора с помощью ЭВМ.

При применении статистического моделирования работа оператора за пультом управления рассматривается как функционирование одноканальной системы массового обслуживания, на вход которой поступает множество потоков неоднородной структуры.

Рассмотрим основные этапы разработки статистической модели деятельности оператора.

1. Изучение и описание деятельности оператора. Основное содержание этого этапа заключается в определении круга обязанностей оператора и далее — в их детализации до уровня элементарных действий. При выполнении этого этапа можно рекомендовать использование алгоритмических методов описания деятельности оператора.

2. Описание деятельности оператора в терминах теории массового обслуживания. На данном этапе разработки модели определяются особенности оператора как системы массового обслуживания: потоки заявок и их вероятностные характеристики процесса обслуживания, критерии эффективности обслуживания и т. д. Эти исходные данные определяются спецификой работы оператора в системе, а их количественное выражение находится, как правило, экспериментальным (реже — расчетным) путем.

3. Учет психофизиологических особенностей операторской деятельности. Рассмотрим один из возможных методов учета психофизиологических особенностей операторской деятельности.

Учет напряженности в работе оператора может быть реализован в модели следующим образом. Ненапряженное (нормальное) состояние имеет место, когда число поступивших сигналов k не превышает объема кратковременной или оперативной памяти k_0 и не возникает дефицит времени при обработке информации. При нарушении любого из этих условий возникает напряженность. Условия возникновения этих состояний можно записать следующим образом:

$$а) k \leq k_0, \bar{\tau} \leq t_{\text{доп}} - \tau_{\text{ож}}; \quad (4.1)$$

$$б) k > k_0; \bar{\tau} > t_{\text{доп}} - \tau_{\text{ож}}, \quad (4.2)$$

где $\tau_{\text{ож}}$ — время ожидания начала обслуживания данного сигнала; $t_{\text{доп}}$ — допустимое время пребывания сигнала на обслуживании, или лимит времени, отводимый на обслуживание данного сигнала.

Из выражений (4.1) и (4.2) следует, что дефицит времени возникает тогда, когда оператор видит, что оставшегося времени ($t_{\text{доп}} - \tau_{\text{ож}}$) ему не хватает для своевременной обработки сигнала, если он будет работать с нормальной скоростью. Это требует от оператора либо более интенсивной работы, либо работы по «свернутому» алгоритму. В обоих случаях число ошибок возрастает.

При рассмотрении динамики работоспособности оператора учитывается влияние «вхождения» в работу (вработываемости) и утомления. Для учета этих явлений основные характеристики деятельности оператора задаются в функции времени его работы за пультом управления. Вид этих функций находится экспериментально.

Моделирование ошибок заключается в следующем. В результате экспериментальных исследований определяется вероятность совершения ошибок $p_{\text{ош}}$. В модели с помощью датчика случайных чисел выбирается результат для получения события «возникновение ошибки при обслуживании требования». Это имеет место при выполнении условия [4]

$$p_{\text{ош}} < \xi, \quad (4.3)$$

где ξ — число, равномерно распределенное в интервале $[0, 1]$, вырабатываемое датчиком случайных чисел.

Как правило, факт ошибки при выполнении операций за пультом индуцируется, и оператор повторяет ошибочное действие. В модели это учитывается путем реализации повторного обслуживания.

Моделирование уровня натренированности оператора производится изменением характеристик деятельности оператора. Как правило, устанавливается несколько градаций уровня натренированности (отлично, хорошо, удовлетворительно) и для них задаются таблично характеристики деятельности.

4. Разработка структурной схемы модели. Укрупненная структурная схема алгоритма деятельности оператора приведена на рис. 32.

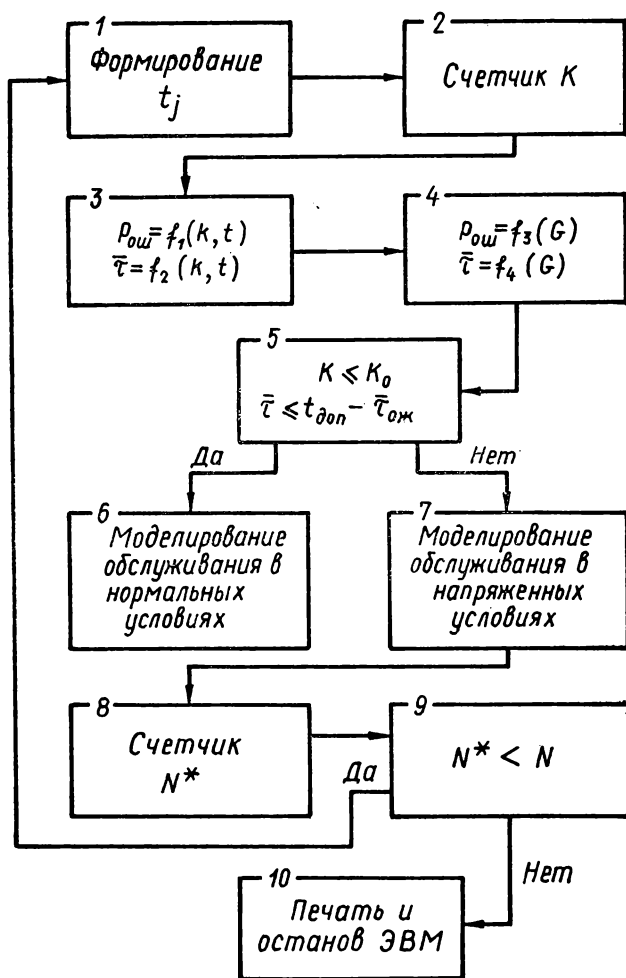


Рис. 32

Принцип работы модели заключается в следующем. Блок 1 формирует моменты поступления t_j очередных сообщений в соответствии с заданным законом распределения. Блок 2 производит подсчет числа k сообщений, одновременно находящихся на обслуживании. Блок 3 вычисляет основные характеристики деятельности оператора в зависимости от длины очереди и длительности работы оператора (учет динамики работоспособности). Блок 4 определяет характеристики деятельности оператора в зависимости от уровня натренированности. Блок 5 проверяет условия возникновения напряженности в соответствии с выражениями (4.1) и (4.2). В зависимости от результатов этой проверки блоки 6 и 7 моделируют процесс обслуживания в нормальных и напряженных условиях. Блок 8 подсчитывает количество проделан-

ных реализаций N^* , а блок 9 сравнивает его с числом N — количеством реализаций, необходимых для достижения заданной точности моделирования. При $N^* < N$ моделирование продолжается, при $N^* = N$ — заканчивается и блок 10 выдает результаты моделирования на печать. В зависимости от целей моделирования на печать могут быть выданы различные характеристики деятельности оператора: вероятность обслуживания, вероятность отказа, время ожидания и т. д.

5. Моделирование процессов в статистической модели. Моделирование случайных процессов в статистической модели включает:

- моделирование случайных моментов поступления заявок;
- анализ состояния системы массового обслуживания в момент поступления очередной заявки и формирование времени начала ее обслуживания;

- моделирование деятельности обслуживания заявки;
- формирование критериев эффективности обслуживания.

Моменты поступления заявок могут иметь произвольный закон распределения. Наиболее часто этот закон экспоненциальный. Любой закон распределения может быть получен с помощью последовательности чисел, подчиняющейся равномерному закону распределения в интервале $[0, 1]$. Способы таких преобразований достаточно развиты в настоящее время.

Экспериментальный метод исследования деятельности оператора. Сущностью экспериментального метода является установление связи между двумя или несколькими фактами или намеренное воспроизведение в специальных условиях установленной ранее связи. Экспериментальный метод складывается из двух фаз: наблюдения, которое позволяет обнаружить факты и сформулировать зависимости, и экспериментирования, целью которого является проверка гипотез.

При проведении эксперимента возможны следующие разновидности применения метода наблюдения:

- анализ поведения одного оператора (одной группы операторов) в одной и той же ситуации;

- наблюдение за одним оператором (одной группой) в разных ситуациях;

- наблюдение за поведением различных операторов в одинаковых условиях.

В зависимости от характера организации наблюдения различают наблюдение случайное и систематическое. Как правило, при эргономическом эксперименте используют систематическое наблюдение. Оно имеет определенную цель, разработанный план, методику фиксирования данных. Для получения более достоверных данных о поведении оператора исследователь обычно не вмешивается в ход эксперимента.

Эффективность наблюдения зависит от правильного сочетания всех видов и форм, их организации и методов фиксирования результатов эксперимента. Основным из этих факторов является организация наблюдения.

Для успешного наблюдения важно обеспечить применение различных методов регистрации наблюдаемых явлений. К числу основных

методов регистрации относятся: хронометраж, протокольная запись, алгоритмическое описание деятельности, кино- и фотосъемки, магнитофонные записи специальных бесед или словесных реакций по ходу выполнения действий и др. Основная ценность наблюдения заключается в создании условий для выдвижения гипотез о характере зависимостей между результатами деятельности оператора и различными внешними и внутренними условиями деятельности.

Метод наблюдения дополняется обычно разнообразными методами опроса оцениваемых операторов, которые могут применяться во многих случаях; когда фактор, который нужно учесть, плохо поддается внешнему контролю; когда изучаемый фактор легко выделяют операторы, но для его тщательного учета в наблюдении или в эксперименте требуется длительное и сложное исследование; когда другие методы не дают исчерпывающей информации.

К методам опроса относятся: анкеты, беседы, интервью, социометрический метод и метод полярных профилей.

Анкета представляет собой стандартный набор вопросов, подобранных таким образом, чтобы выявить искомый фактор и его особенности с помощью прямых и косвенных вопросов, на которые опрашиваемые дают письменные ответы.

Интервью представляет собой набор вопросов по анкетному типу, но проводится в форме живой беседы, содержание которой может меняться.

Методы социометрии и полярных профилей — тоже вопросного типа, но и с их помощью решается ограниченный круг вопросов, характеризующих особенности групповой деятельности.

Проведение опроса и статистическая обработка его результатов обычно завершаются разработкой плана и методики экспериментального изучения отдельных факторов и их зависимостей, обнаруженных в ходе опроса.

Рассмотрим более подробно основные вопросы поведения эргономического эксперимента применительно к исследованию деятельности оператора АССОРЛИ.

Организация эргономического эксперимента. Она включает: определение цели эксперимента; разработку его замысла; определение состава привлекаемых средств; определение состава операторов, участвующих в эксперименте; подготовку средств регистрации данных и разработку способов регистрации; определение порядка обработки и анализа результатов эксперимента.

По цели эргономические эксперименты можно условно разделить на частные и общие. В частных исследуются отдельные стороны деятельности оператора, определяются некоторые характеристики его работы, влияние на них различных факторов. Частные эргономические эксперименты представляют собой одиночные или групповые налеты имитируемого воздушного флота.

Общие эргономические эксперименты являются массированными налетами имитируемого воздушного флота. При таких экспериментах АССОРЛИ функционирует в условиях полной загрузки; операторы работают в условиях, максимально приближенных к рабочим. В экс-

перименте исследуются все стороны деятельности операторов, включая оценку влияния операторов на показатели эффективности АССОРЛИ.

Замысел эксперимента определяется его целью и представляет собой разработку тактики действия имитируемого флота. Замысел эксперимента включает определение количества целей в полете, их типов, плотности входа в зону ведения радиолокационной разведки АССОРЛИ, построение порядков, применение маневра, помех и т. д. Замысел реализуется в виде карты действий флота и легенды к ней.

Состав средств АССОРЛИ и количество операторов, участвующих в эксперименте, также определяются целями эксперимента и могут быть полными или сокращенными.

Средства регистрации данных должны обеспечить регистрацию всей необходимой информации, включая не только непосредственную информацию о деятельности операторов, но и информацию о всех этапах работы системы по каждой цели и о качестве выдаваемой информации.

Регистрация данных эксперимента. Основная цель регистрации данных эксперимента — зафиксировать следующие данные: общую воздушную обстановку в динамике; всю (основную) информацию, вырабатываемую ЭВМ и отбираемую на средствах отображения системы; действия операторов во время эксперимента.

Общая схема регистрации и обработки данных эксперимента приведена на рис. 33.

Регистрация данных осуществляется:

1) специальными (сервисными) программами, вводимыми в комплект программ ЭВМ исследуемой системы и обеспечивающими выдачу зарегистрированных данных на магнитную ленту или на ленту автоматической цифропечати (АЦП);

2) посредниками, рабочие места которых оснащены магнитофонами или специальными пультами для фиксации действий оператора с помощью шлейфового осциллографа (самописца);

3) регистрирующими фотокамерами.

Дополнительно к данным, фиксируемым посредниками, с помощью магнитофонов производится запись переговоров между операторами в ходе работы.

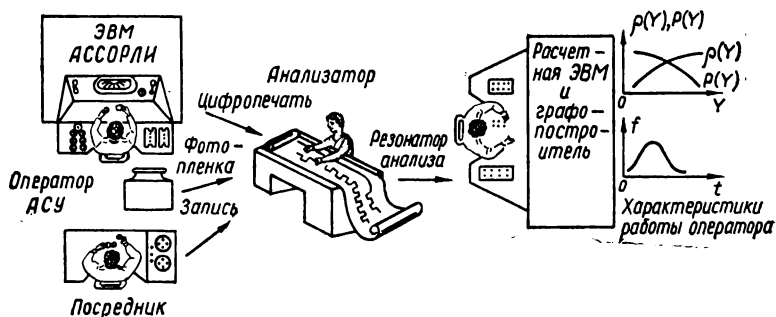


Рис. 33

Регистрация данных осуществляется с привязкой к единому времени, для чего используются соответствующие сигналы (импульсы начала и конца работы и секундные метки времени).

Характер информации, регистрируемой с помощью перечисленной выше аппаратуры, отражен в табл. 10.

После окончания эксперимента на каждом рабочем месте составляется протокол, в который заносятся замечания по работе аппаратуры; имевшие место отклонения от замысла полета; порядок работы системы, нарушения инструкций и т. д.

Т а б л и ц а 10

| Способ регистрации | Характер регистрируемых данных |
|---------------------------|--|
| «Сервисные» программы ЭВМ | <ol style="list-style-type: none"> 1. Моменты ввода оператором команд и распоряжений в ЭВМ 2. Моменты вызова оператором осведомительной информации на экран индикатора и табло 3. Моменты автоматического обновления информации на экране индикатора и табло 4. Решения, вырабатываемые ЭВМ автоматически 5. Обменная информация с управляемыми объектами |
| Посредник | <ol style="list-style-type: none"> 1. Моменты возникновения конфликтных ситуаций, требующих вмешательства оператора 2. Действия оператора при возникновении конфликтных ситуаций 3. Текущая оценка действий оператора посредником |
| Регистрирующая фотокамера | <ol style="list-style-type: none"> 1. Информационная модель 2. Моменты появления на экране индикаторов и табло информации, вызванной оператором 3. Моменты автоматического обновления информации на экране индикатора и табло |

Затем каждый оператор, без консультации с другими операторами, составляет отчет о своей работе или заполняет карточку отзывов, где отражает указанные вопросы.

Обработка результатов эксперимента. Особенности деятельности операторов АССОРЛИ объективно не позволяют полностью автоматизировать обработку и анализ результатов эргономических экспериментов. Наибольшую трудность представляет формализация определения момента фактического возникновения задачи, а также оценка правильности и своевременности решения оператором возникшей задачи.

Указанное обстоятельство обуславливает необходимость привлечения посредников для регистрации действий оператора, а также использования при анализе действий оператора специальных временных диаграмм, построенных анализатором.

После проведения каждого эксперимента строится временная диаграмма работы оператора (рис. 34), которая для удобства анализа содержит несколько осей.

Метки СЕВ

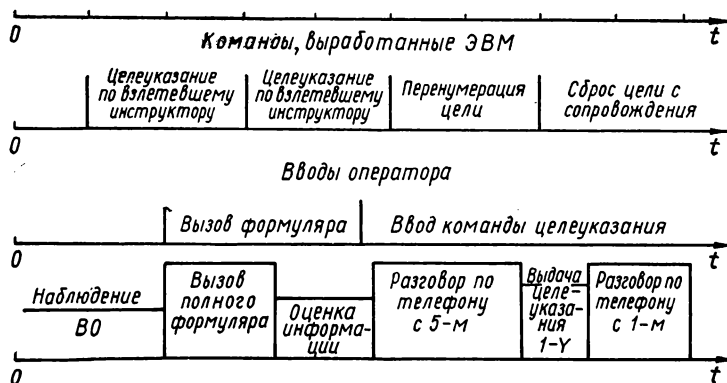


Рис. 34

На первую ось наносится последовательность работы АССОРЛИ в виде моментов появления эхо-сигналов воздушных объектов; команд, вырабатываемых ЭВМ автоматически (например, требование корректуры); команд и информации, поступающих от других операторов, и т. д. Наличие такой информации позволяет выявить задачи, которые объективно возникли в ходе функционирования системы и требовали вмешательства оператора.

На второй оси обозначаются моменты ввода оператором команд и распоряжений в ЭВМ, моменты вызова оператором информации и ее появление на экране индикатора и табло, а также переговоры оператора (по данным регистрации).

На третьей оси отмечаются действия оператора, зафиксированные посредником.

Данные, представленные на этих диаграммах, дополняют друг друга и повышают объективность анализа.

Анализ результатов эксперимента. Анализ результатов эксперимента проводится в такой последовательности.

На основе всех материалов обработки данных регистрации устанавливается последовательность работы АССОРЛИ: взлет имитируемых «целей»; обнаружение целей РЛС; обработка и выдача информации потребителям; выдача и обработка команд управления с вышестоящего КП и т. д.

Одновременно с этим выявляются задачи, которые объективно возникли в ходе работы АССОРЛИ и требовали вмешательства операторов. Для каждой задачи выявляются и отмечаются на временной диаграмме:

момент реального возникновения задачи, т. е. момент возникновения реальной ситуации, требующей вмешательства оператора;

момент появления на средствах отображения информации, свидетельствующей о возникновении задачи;

причина необнаружения задачи оператором и его последствия, т. е. как это повлияло на обработку информации.

Далее временная диаграмма дополнительно анализируется с целью выделения и группирования статистических данных по отдельным задачам.

Вся работа оператора во время эксперимента разделяется на интервалы времени, в течение которых имеют место практически неизменные условия работы оператора (по сложности воздушной обстановки, количеству сопровождаемых целей и т. д.). Все статистические данные группируются по интервалам с одинаковыми условиями и по типам задач.

Для однотипных задач и одинаковых условий считываются с временной диаграммы и последовательно заносятся в табл. 11 следующие

Т а б л и ц а 11

| | | |
|----|--|--------------------|
| 1 | Номер задачи | 17 |
| 2 | Наименование задачи | Оповещение |
| 3 | Момент появления задачи $t_{i\text{возн}}$ | 256с |
| 4 | Момент начала решения задачи $t_{i\text{нач}}$ | 286с |
| 5 | Момент окончания (прекращения) решения $t_{i\text{кон}}$ | 314с |
| 6 | Момент исчезновения задачи $t_{i\text{исч}}$ | 387с |
| 7 | Время ожидания начала решения $t_{i\text{ож}}$ | 30с |
| 8 | Время решения задачи τ_i (полного, частичного) | 28с |
| 9 | Время существования задачи $t_{i\text{сущ}}$ | 101с |
| 10 | Результат решения задачи (правильно, неправильно, частично) | Правильно |
| 11 | Влияние неправильного (частичного) решения на боевую задачу (привело к срыву задачи или нет) | |
| | Причина задержки начала решения | Загрузка оператора |
| 13 | Причина пропуска (частичного решения) задачи | — |
| 14 | Примечания | — |

данные: длительность задержки начала решения; причина задержки начала решения; время решения задачи; результат решения задачи (правильно или неправильно она решена); длительность существования задачи, не решавшейся оператором; причина пропуска задачи; длительность частичного решения задачи; причина прерывания решения задачи; последствия, к которым привели пропуск задачи или прерывание ее решения (как это сказалось на характеристиках информации).

Данные, занесенные в табл. 11, обрабатываются с целью расчета основных количественных характеристик работы оператора.

Обобщение экспериментальных данных. Обобщение экспериментальных данных включает следующие этапы, выполняемые последовательно: определение характеристик задач; оценка качества и напряженности работы оператора; оценка коллек-

тива операторов; оценка влияния операторов на показатели эффективности АССОРЛИ.

Остановимся на содержании указанных этапов.

1. По результатам анализа отдельных задач, возникших в эксперименте, должны быть определены:

полный перечень задач ($i = 1, \dots, i^*, i$ — индекс задачи);

условия возникновения i -й задачи;

условия обнаружения i -й задачи оператором;

плотность возникновения i -й задачи λ_i ;

длительность существования i -й задачи $t_{i\text{сущ}}$;

вероятность правильного решения i -й задачи $p_i(t)$;

закон распределения времени решения i -й задачи $\beta_i(t)$ с числовыми характеристиками τ_i и σ_i ;

влияние результата решения i -й задачи на характеристики радиолокационной информации.

Плотность возникновения i -й задачи λ_i определяется как количество задач i -го типа в единицу времени. Величина получается обобщением данных по результатам всех экспериментов с учетом условий работы оператора y (количество сопровождаемых целей, наличие и интенсивность помех, количество источников информации и т. д.):

$$\lambda_i(y) = \frac{\sum_{k=1}^N n_{ik}(y)}{\sum_{k=1}^N T_k(y)} = \frac{1}{N} \sum_{k=1}^N \lambda_{ik},$$

$$\lambda_{ik} = n_{ik}(y)/T_k(y), \quad (4.4)$$

где y — параметр условий работы оператора; $n_{ik}(y)$ — количество возникших задач i -го типа ($i = \overline{1, i^*}$) в k -м эксперименте ($k = \overline{1, N}$); $T_k(y)$ — длительность периода, в течение которого параметр условий работы оператора оставался неизменным.

При условии достаточного количества экспериментальных данных ($n \geq 30 \dots 50$) и «набора» условий обстановки известными методами обработки экспериментальных данных определяется вид функциональной зависимости $\lambda_i(y)$.

Длительность существования i -й задачи $t_{i\text{сущ}}$ определяется как отрезок времени от момента реального возникновения задачи (момента возникновения ситуации, требующей какого-либо вмешательства в процесс автоматизированной обработки информации) до момента ее «исчезновения» (момента, когда решение становится практически нецелесообразным, так как не приводит к положительному эффекту). Для определения этих моментов используется табл. 11.

Экспериментальное определение вероятности (частности) правильного решения i -й задачи $p_i(y)$ производится в соответствии с очевидным выражением

$$p_{ik}(y) = \sum_{k=1}^N m_{ik}(y) / \sum_{k=1}^N n_{ik}(y) = \frac{1}{N} \sum_{k=2}^N p_{ik}(y); \quad (4.5)$$

$$p_{ik} = m_{ik}/n_{ik},$$

где m_{ik} — число правильно и своевременно решенных задач i -го типа в k -м эксперименте.

Экспериментальное определение закона распределения времени решения i -й задачи $\beta_i(t)$ и его числовых характеристик τ_i и σ_i производится на основе статистики, объем которой должен быть не менее 50 точек. При обработке экспериментальных данных используется ЭВМ.

Влияние результата решения i -й задачи на характеристики радиолокационной информации можно оценить как условную вероятность q_{si} ухудшения S -го показателя качества информации при том, что i -я задач не решена вообще или неправильно решена:

$$q_{si} = \frac{1}{N} \sum_{k=1}^N q_{sik}; \quad q_{sik} = \frac{r_{sik}}{n_{ik} - m_{ik}}, \quad (4.6)$$

где r_{sik} — количество случаев, когда нерешение или неправильное решение i -й задачи оператора привело к ухудшению S -го показателя качества радиолокационной информации.

2. Качество работы оператора определяется следующими показателями:

вероятностью (частотой) правильного и своевременного решения i -й задачи в соответствии с выражением (4.5);

вероятностью (частотой) правильного и своевременного решения всех задач

$$p(y) = \frac{1}{N} \sum_{k=1}^N p_k(y);$$

$$p_k(y) = \frac{\sum_{i=1}^{i^*} m_{ik}(y)}{\sum_{i=1}^{i^*} n_{ik}(y)}. \quad (4.7)$$

Кроме качества работы оператора экспериментально определяется режим работы оператора, характеризуемый уровнем напряженности его работы или загрузкой. По результатам оценки режима работы оператора определяется, работает ли оператор в оптимальном диапазоне загрузки.

Напряженность работы оператора характеризуется следующим показателем:

в отдельном k -м эксперименте

$$\rho_k(y) = \frac{1}{T_k(y)} \sum_{i=1}^{i^*} \tau_{ik}(y), \quad (4.8)$$

где τ_{ik} — время решения i -й задачи в k -м эксперименте при усреднении по всем экспериментам

$$\rho(y) = \frac{1}{N} \sum_{k=1}^N \rho_k(y). \quad (4.9)$$

При оценке напряженности учитываются все решаемые оператором задачи, в том числе и решенные неправильно или частично.

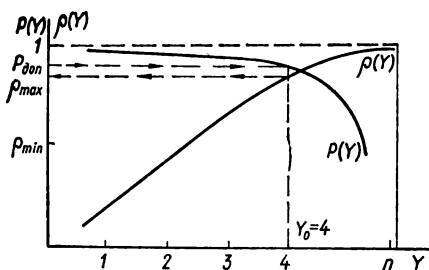


Рис. 35

На заключительной стадии экспериментальной оценки качества работы оператора определяется характер зависимостей $p_i(y)$, $P(y)$ и $\rho(y)$.

Для этого результаты отдельных экспериментов обобщаются по признаку одинаковых условий работы оператора y . Строятся графики зависимостей $p_i(y)$, $P(y)$ и $\rho(y)$ как функции параметра y . Пример построения этих функций приведен на рис. 35.

Для установления вида зависимостей $p_i(y)$, $P(y)$ и $\rho(y)$ производится сглаживание экспериментальных данных по методу наименьших квадратов. Подбор эмпирических формул осуществляется, как правило, из числа функций следующих классов: линейная функция, многочлен 2—3-й степени, экспоненциальная функция. Степень согласованности полученного уравнения с экспериментальными данными проверяется по F -критерию Фишера.

Наличие зависимостей $p_i(y)$, $P(y)$ и $\rho(y)$ позволяет определить комплексную характеристику оператора — его пропускную способность (производительность).

Под производительностью оператора будем понимать количество сопровождаемых целей, управляемых источников информации, обслуживаемой аппаратуры и т. д., при котором выполняются два условия: качество решения задач $p_i(y)$, $P(y)$ достаточно высокое; режим работы оператора, характеризуемый величиной $\rho(y)$, оптимален.

Формально указанные условия имеют вид

$$\begin{aligned} p_{i0} &\leq p_i(y) \leq 1; \\ P_{\text{доп}} &\leq P(y) \leq 1; \\ \rho_{\min} &\leq \rho(y) \leq \rho_{\max}. \end{aligned} \quad (4.10)$$

Граничные значения p_{i0} , $P_{\text{доп}}$, ρ_{\min} , ρ_{\max} задаются из следующих практических соображений:

1) для наиболее важных задач величина p_{i0} должна быть высокой (порядка 0,8—0,9);

2) значения ρ_{\min} , ρ_{\max} задаются с учетом специфики деятельности оператора;

3) если деятельность имеет в основном сенсомоторный характер $\rho_{\min} \approx 0,8$; $\rho_{\max} \approx 0,9$;

если деятельность связана с принятием решений, то $\rho_{\min} \approx 0,5$; $\rho_{\max} \approx 0,7 \dots 0,8$.

Производительность оператора оценивается с помощью графиков рис. 35 следующим образом.

По заданным значениям $P_{\text{доп}}(p_{i0})$ определяется параметр y , округляемый до ближайшего целого числа. При найденном значении параметра y по графику $\rho(y)$ находится загрузка оператора. Проверяется, находится ли величина $\rho(y)$ в пределах $\rho_{\min} \dots \rho_{\max}$. Если это условие

не выполняется, корректируется значение параметра y с последующей проверкой значений $P(y)$ и $p_i(y)$. Величина y_0 определяет производительность оператора.

3. При анализе деятельности коллектива операторов определяются: рациональные схемы организации работы коллектива; рациональность распределения функций между операторами; количество операторов.

Под схемой организации работы коллектива будем понимать логическую последовательность обработки информации операторами и функциональные связи между ними. Под функциональной связью между операторами понимается любое информационное взаимодействие в процессе обработки информации.

При оценке рациональности схемы организации работы коллектива проверяется выполнение следующих требований:

- минимальное число уровней управления;

- соблюдение норм управляемости (количество подчиненных, замкнутых непосредственно на руководителя, не должно превышать 5—8 человек);

- соблюдение единоначалия (у каждого подчиненного должен быть только один непосредственный начальник);

- соблюдение субординации (подчиненные одного уровня могут иметь только информативные, но не командные связи на своем уровне).

При оценке распределения функций между операторами должны приниматься во внимание следующие факторы:

- важность задачи, т. е. влияние на показатели качества радиолокационной информации;

- сложность решения задачи и объем используемой при этом информации;

- возможность выполнения функций каждым оператором по фактору времени;

- равномерность загрузки операторов;

- однотипность функций, выполняемых оператором (выполнение функций не должно требовать разнохарактерных знаний).

Важность и сложность задач учитываются следующим образом:

- более важные, ответственные задачи должны решаться на более высоких уровнях управления;

- простые задачи, которые связаны с выполнением большого числа моторных операций, решаются на низких уровнях;

- сложные задачи, требующие обработки (осмысливания) большого объема информации, решаются на верхних уровнях.

Оценка необходимого количества операторов M производится исходя из требований обеспечения необходимой производительности системы P_c и производительности оператора $P_{оп}$, т. е.

$$M = \left[\frac{P_c}{P_{оп}} \right] + 1, \quad (4.11)$$

где $[]$ — целая часть числа.

4. На заключительном этапе анализа определяется влияние коллектива операторов на системные характеристики: производительность и эффективность.

Для оценки влияния операторов на производительность системы определяется производительность коллектива, которая принимается равной сумме производительностей отдельных операторов. Полученный результат сравнивается с производительностью ЭВМ данной системы. Если производительность коллектива больше или равна производительности ЭВМ, то считается, что операторы не ограничивают производительность системы.

Для оценки влияния коллектива операторов на эффективность работы системы необходимо прежде всего установить, на какие показатели эффективности оказывает влияние работа каждого оператора. Очевидно, на общие показатели эффективности, как правило, влияет работа более широкого круга операторов, а на частные — работа ограниченного круга операторов.

Влияние коллектива на эффективность системы можно оценивать показателями следующего вида:

коэффициентом повышения эффективности системы за счет деятельности операторов

$$\Delta Q = \frac{Q_{\text{оп}} - Q_0}{Q_0}, \quad (4.12)$$

где Q_0 , $Q_{\text{оп}}$ — эффективность системы соответственно без учета и с учетом деятельности операторов;

коэффициентом использования возможностей коллектива операторов по повышению эффективности системы

$$\Delta S = \frac{Q_{\text{оп}}}{Q_{\Sigma}}, \quad (4.13)$$

где Q_{Σ} — эффективность системы при оптимальных условиях деятельности операторов.

Очевидно, если система практически не может функционировать без участия операторов или не может выполнять какая-то функция системы, то $\Delta Q = \infty$, т. е. сама постановка задачи оценки влияния операторов на показатели системы в таком виде не имеет смысла.

Применение комбинированного метода исследования деятельности оператора. Суть этого метода заключается в том, что при исследовании сложных систем вооружения методом моделирования часть функций операторов моделируется при помощи математических моделей, а остальные — «подыгрываются» операторами моделирующей установки, работающими за специальными рабочими местами. Структурная схема моделирующей установки приведена на рис. 36.

При разработке моделирующей установки сложной системы возникает проблема учета деятельности оператора при моделировании. Необходимо решить, каким образом учитывать деятельность операторов, чтобы обеспечить заданное соответствие выходных характеристик модели реальной системе. Это соответствие может быть достигнуто, во-первых, за счет полного учета в модели функций реальных операторов, во-вторых, за счет обеспечения условий деятельности операторов модели, соответствующих реальным условиям.

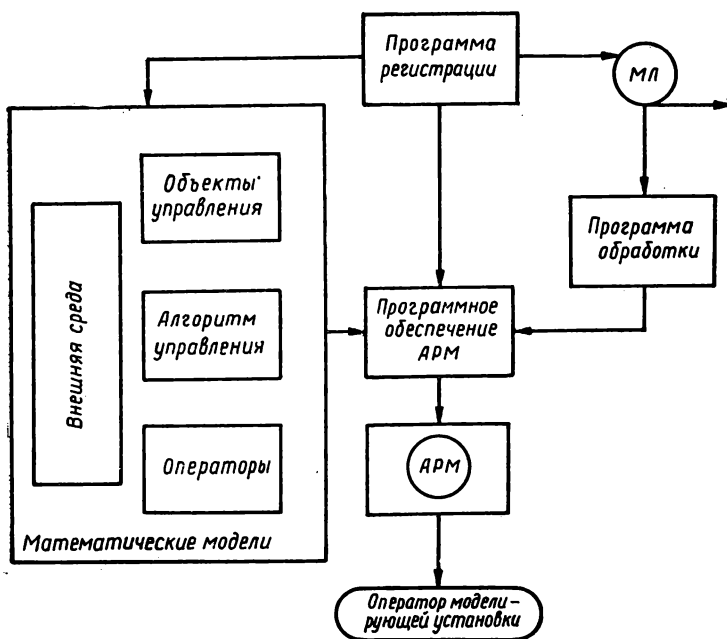


Рис. 36

При моделировании необходимо учитывать деятельность операторов различных рангов, включая командира. Содержание деятельности операторов существенно различно — от простой моторной деятельности до эвристической. Поэтому возникает вопрос, какие функции операторов имитировать при помощи ЭВМ и какие — при помощи операторов моделирующей установки, т. е. необходимо распределить функции операторов исследуемой системы для их имитации между ЭВМ и человеком.

Из интуитивных соображений очевидно, что функции операторов низших рангов, которые сравнительно легко поддаются формализации, могут быть реализованы в виде программы на ЭВМ. Моделирование на ЭВМ деятельности операторов более высоких рангов затруднительно из-за неформального характера решаемых ими задач. Поэтому имитация этих задач может быть обеспечена только при помощи операторов моделирующей установки.

Проектирование смешанной модели можно условно представить в виде трех этапов (рис. 37).

На первом этапе выбирается вариант имитации. В результате такого выбора множество имитируемых задач разбивается на два подмножества, одно из которых имитируется при помощи ЭВМ, а другое — при помощи операторов моделирующей установки.

На втором этапе определяются способы формализованного описания операторских задач, имитация которых возможна на ЭВМ, а также проектируются условия деятельности операторов моделирующей

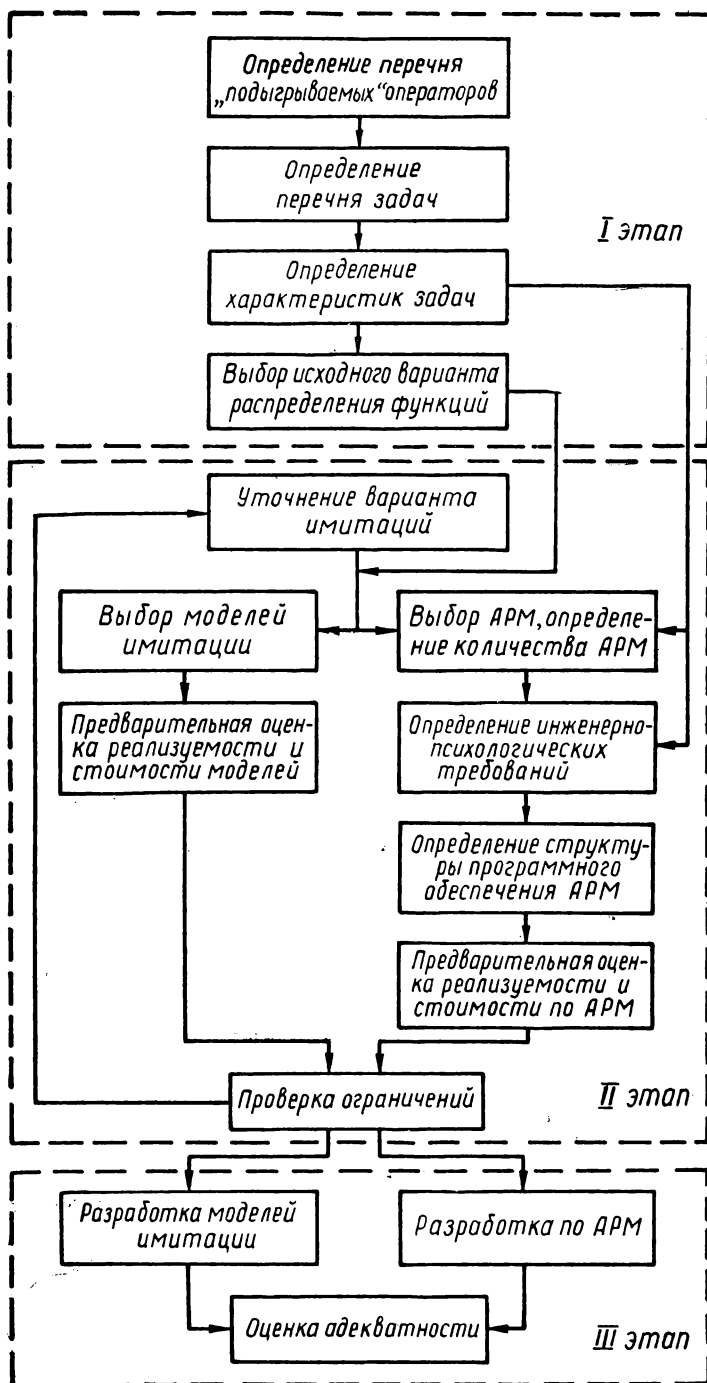


Рис. 37

установки. Для формализованного описания деятельности оператора, в частности, возможно использование математических моделей, рассмотренных в данной главе.

Для проектирования условий деятельности оператора моделирующей установки, соответствующих реальным условиям деятельности, проводятся эргономические исследования и определяются необходимое количество рабочих мест, требования к информационной модели, разрабатывается программное обеспечение рабочих мест для сопряжения с моделирующей установкой.

Третий этап включает разработку и отладку математических моделей и программ имитации деятельности операторов, а также программного обеспечения рабочих мест.

Применение комбинированного метода открывает широкие возможности для исследования деятельности операторов сложной системы. Комбинированный метод обладает всеми преимуществами экспериментального метода и метода моделирования. Таким образом, с одной стороны, деятельность оператора исследуется в условиях, близких к реальным; с другой, — имеется возможность проведения большого количества экспериментов, охватывающих весь диапазон условий деятельности. Кроме того, наличие эталонов при моделировании позволяет полностью автоматизировать процесс анализа деятельности оператора и получения количественных характеристик деятельности.

§ 5. Эргономический анализ деятельности оператора АССОРЛИ

Количественные характеристики выполнения элементарных операций. Среди множества операций, выполняемых операторами АССОРЛИ, наиболее типичны сенсорные, перекодировочные, моторные, а также операции смешанного типа (например, сенсомоторные, сенсоперекодировочные и т. д.). Эвристические операции, связанные с принятием решения, в дальнейшем рассматривать не будем ввиду сложности их изучения.

К сенсорным операциям можно, к примеру, отнести поиск объекта, обладающего определенным признаком, в информационном поле; обнаружение полезного сигнала; отождествление объектов по каким-то признакам и т. д. Перекодировочными операциями можно назвать операции, связанные с восприятием информации табло, формуляров воздушных объектов. Моторные операции включают вызов информации, ввод ее в ЭВМ при помощи пульта, отдачу приказаний голосом.

Как уже указывалось в § 2, основными характеристиками выполнения операций являются время выполнения, безошибочность и точность. Рассмотрим характеристики выполнения оператором АССОРЛИ некоторых типичных операций.

В работе исследовалась зависимость времени поиска τ_n от влияющих на него факторов. Оказалось, что время поиска существенно зависит от двух факторов информационного поля: общего объема отображения (числа элементов информационного поля) N и оперативного объема отображения M (числа критических элементов информа-

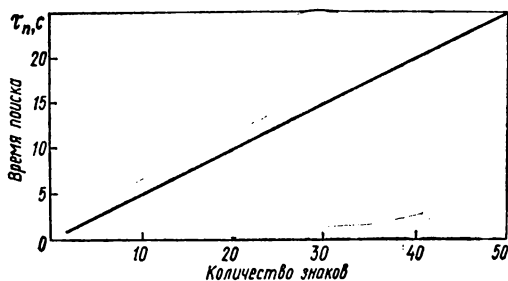


Рис. 38

Пример зависимости времени зрительного поиска от общего объема отображаемой информации приведен на рис. 38.

Основным фактором, влияющим на время выполнения моторных операций τ_m , является количество установок переключателей и кнопок l при выполнении операции. Для определения характера связи величин τ_m и l использовался корреляционно-регрессионный анализ. Оказалось, что эта зависимость имеет следующий вид:

$$\tau_m = \bar{\tau}_m + \alpha (l - \bar{l}), \quad (5.2)$$

где $\bar{\tau}_m$ — среднее значение времени выполнения; \bar{l} — среднее количество кнопок и переключателей; α — коэффициент регрессии.

Вид зависимостей для одного из автоматизированных рабочих мест приведен на рис. 39.

Рассмотрим некоторые характеристики оператора в процессе полуавтоматической первичной обработки при выполнении операций обнаружения воздушных объектов и съема их координат.

Отметки реальных воздушных объектов на экранах индикаторов имеют определенную форму, их угловой размер зависит от технических параметров РЛС и изменяется под влиянием помех. Шумовые отметки отличаются от отметок реальных целей. Чаще всего они имеют

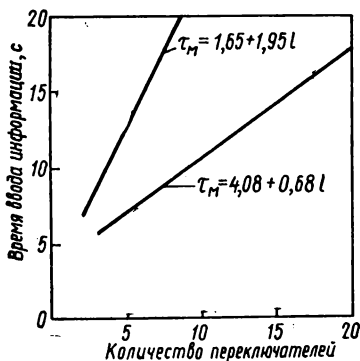


Рис. 39

ционного поля, которые обладают искомым признаком).

Полученная экспериментальным путем зависимость имеет следующий вид:

$$\tau_n = \frac{N+1}{M+1} T_{\text{фп}}, \quad (5.1)$$

где $T_{\text{фп}}$ — средняя длительность фиксации глаза при поиске ($T_{\text{фп}} \approx 0,28$ с).

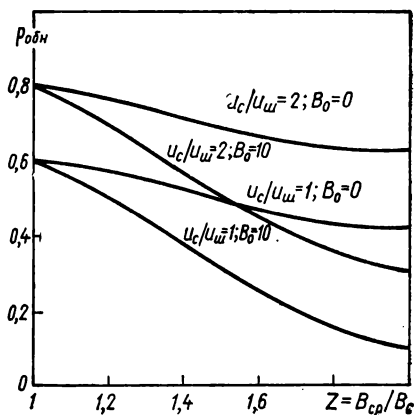


Рис. 40

круговую, реже — эллиптическую форму и подвержены значительно большим изменениям во времени.

Информация о размерах и форме отметки используется оператором в процессе решения, когда, получив от зрительной системы сообщение о превышении контрастирующими яркостями порога контрастной чувствительности глаза (обнаружив отметку) и информацию о размерах и форме, центральная нервная система решает задачу опознавания образа и выносит решение.

В соответствии с гипотезой, высказанной рядом авторов, решение оператора об обнаружении сигнала основывается на сравнении яркости отметки и яркости узкой области, прилегающей к отметке, т. е. на контрасте между отметкой и фоном. На рис. 40 приведены графики вероятности обнаружения одиночного сигнала от порогового значения контраста $z_c = B_{\phi}/B_c$ для различных значений отношения сигнала к шуму $U_c/\sigma_{\text{ш}}$ и различных уровней освещенности B_c .

Обнаружив цель, оператор приступает к решению последующей задачи первичной обработки — съему координат. Точный съем координат осуществляется совмещением электрического маркера с центром отметки цели. Точность совмещения маркера с центром отметки наряду с психофизиологическими характеристиками оператора, его опытом определяется качеством фокусировки, размером отметки вдоль линии развертки, масштабом, диаметром электронно-лучевой трубки и т. д.

Ошибки измерения дальности и угловых координат за счет инерционности оператора возникают при быстрых изменениях положения отметки, на которые оператор не успевает реагировать. Их величина зависит от субъективных особенностей оператора и возрастает с увеличением скоростей и ускорений изменения текущих координат цели.

Рассмотрим субъективные ошибки оператора при изменении дальности и угловой координаты.

Ошибка съема дальности при пользовании электронным маркером определяется в основном точностью совмещения маркера с центром отметки и качеством фокусировки.

Ошибка в нахождении отметки зависит от размера отметки и определяется в конечном счете параметрами РЛС и индикатора. Экспериментально было вычислено среднеквадратическое значение этой ошибки

$$\sigma_u = 0,03 \frac{v\tau + 0,5d}{m_R}, \quad (5.3)$$

где v — скорость развертки (мм/кмс); τ — длительность импульса от цели (мкс); m_R — масштаб шкалы дальности (мм/км); d — диаметр пятна (мм).

При недостаточно хорошей фокусировке ошибка в определении центра отметки возрастает. Считается, что величина ошибки из-за недостаточно хорошей фокусировки не превышает диаметра пятна в масштабе шкалы, а ее среднеквадратическое значение может быть вычислено по формуле

$$\sigma = \frac{0,5d}{2\sqrt{3}m_R}. \quad (5.4)$$

Суммарная ошибка оператора определяется очевидным соотношением

$$\sigma_{\Sigma} = \sqrt{\sigma_{\Sigma}^2 + \sigma_{\text{ср}}^2}. \quad (5.5)$$

Рассмотрим пример расчета величины ошибки съема дальности. Исходные данные: $D_{\text{э}} = 450$ мм; $R_{\text{РЛС}} = 300$ км; $d = 1$ мм; $\tau = 1$ мкс; период развертки $T = 2500$ мкс. Определяем масштаб шкалы дальности m_R и скорость развертки v :

$$m_R = \frac{D_{\text{э}}}{R_{\text{РЛС}}} = 1,5 \frac{\text{мм}}{\text{км}}; \quad v = \frac{D_{\text{э}}}{T} = 0,18 \frac{\text{мм}}{\text{мкс}}.$$

Имеем

$$\sigma_{\Sigma} \approx 14 \text{ м}; \quad \sigma_{\Phi} \approx 96 \text{ м}; \quad \sigma_{\Sigma} \approx 100 \text{ м}.$$

Среднеквадратическая ошибка измерения угловой координаты за счет неточности совмещения маркера с центром отметки вычисляется по формуле

$$\sigma_{\beta} \approx 0,58 \left(\frac{\varphi_0}{2} + 57,3 \frac{R_{\text{max}}}{R} \frac{1}{\lambda} \right), \quad (5.6)$$

Т а б л и ц а 12

| Содержание операции | Временные характеристики | | Вероятность произвольного выполнения P_{0n} |
|---|--------------------------|-------------------|---|
| | $\bar{\tau}_{0,n}$, с | σ_{0n} , с | |
| Сопровождение цели | | | |
| Первый ввод координат | 2,5 | 0,5 | 0,999 |
| Второй ввод (запрос + ввод) | 3,5 | 0,7 | 0,994 |
| Присвоение характеристик | 2,2 | 0,4 | 0,991 |
| Корректурa | 2 | 0,4 | 0,999 |
| Сброс цели с сопровождения | 2,3 | 0,4 | 0,997 |
| Отождествление информации разных источников | 4,7 | 0,6 | 0,85 |
| Управление обработкой информации | | | |
| Оценка воздушной обстановки по селектуемому параметру | 14,5 | 5,3 | 0,93 |
| Оценка характеристик цели по формуляру | 11 | 3,8 | 0,79 |
| Оценка состояния источника по формуляру | 10,3 | 3,5 | 0,83 |
| Гашение информации фотосъемником | 8 | 2,7 | 0,91 |

где φ_0 — ширина диаграммы направленности антенны по точкам половинной мощности, град; R_{max} — максимальная дальность обнаружения цели РЛС, км; R — текущая дальность, км; $\lambda = 2D_{\text{э}}/d$ — разрешающая способность трубки.

Оценим значение ошибки измерения азимута цели при применении той же электронно-лучевой трубки и значениях других исходных данных, равных $R = 0,5 R_{\text{max}}$; $\varphi_0 = 1^\circ$. Тогда

$$\sigma_{\beta} \approx 22'.$$

В табл. 12 приведены характеристики наиболее типичных операций при сопровождении целей и управлении обработкой информации, полученные экспериментально для операторов одной из АССОРЛИ.

Эргономический анализ задач. Задача управления источниками радиолокационной информации является одной из центральных. Суть ее заключается в выборе источника, который может обеспечить

Т а б л и ц а 13

| Описание членов алгоритма | Обозначение | Количественные характеристики | | | |
|---|-------------|-------------------------------|--------------|----------|----------|
| | | $\bar{\tau}_{0n}$, с | D_{0n} , с | P_{0n} | γ |
| Выбор очередной цели для назначения источнику | O_1 | 3,1 | 2,2 | 0,97 | — |
| Уяснение содержания рекомендации ЭВМ по назначению источника | O_2 | 1,7 | 0,5 | 1 | — |
| Логическое условие: необходимо оценить характеристики цели по ее формуляру | I_1 | 0,5 | 0,3 | 1 | 0,62 |
| Вызов формуляра цели | O_3 | 4,1 | 2,3 | 0,98 | — |
| Уяснение содержания формуляра | O_4 | 1,1 | 0,9 | 0,99 | — |
| Логическое условие: необходимо оценить состояние источника информации | I_2 | 0,5 | 0,3 | 1 | 0,56 |
| Оценка состояния источника по табло | O_5 | 3,3 | 1,6 | 0,99 | — |
| Логическое условие: необходимо оценить подлетное время цели до источника | I_3 | 0,5 | 0,3 | 1 | 0,33 |
| Вызов упрежденной точки цели | O_6 | 4,0 | 1,1 | 0,98 | — |
| Оценка подлетнего времени | O_7 | 1,4 | 0,7 | 0,98 | — |
| Назначение цели источнику информации | O_8 | 5,2 | 3,9 | 0,96 | — |
| Передача распоряжения о назначении цели на источник (выдача целеуказания) | O_9 | 7,4 | 3,1 | 1 | — |

качественное сопровождение цели, путем оценки характеристик цели и состояния источника.

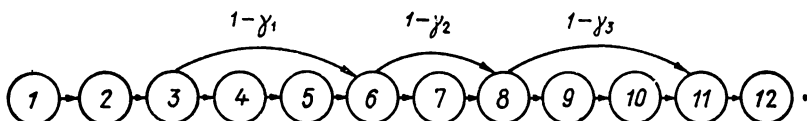
Проведем исследование этой задачи с помощью алгоритмического метода применительно к одной АССОРЛИ.

Первым шагом исследования является составление алгоритма решения задачи. Для составления алгоритма задача была разбита на операции и логические условия. Содержание операций и логических условий, а также их количественные характеристики приведены в табл. 13.

Алгоритм решения задачи, записанный в операторной форме, имеет следующий вид:

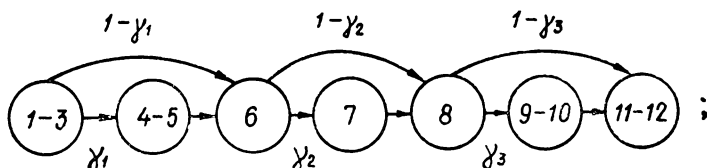
$$O_1 O_2 I_1 \uparrow^2 O_3 O_4 \downarrow^1 I_2 \uparrow^3 O_5 \downarrow^2 I_3 \uparrow^8 O_6 O_7 \downarrow^3 O_8 O_9.$$

Граф-схема алгоритма выглядит следующим образом:



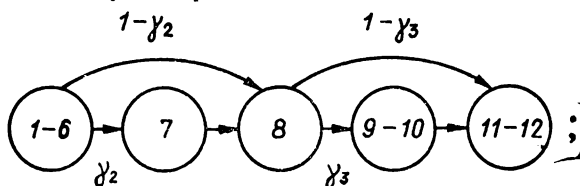
Следующим шагом исследования является получение количественных характеристик задачи управления источниками путем применения правил преобразования граф-схемы алгоритма, изложенных в § 3.

1. Объединение путей без разветвления



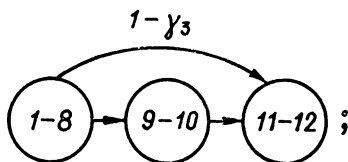
$$\begin{aligned}\bar{\tau}_{1-3} &= \bar{\tau}_1 + \bar{\tau}_2 + \bar{\tau}_3 = 5,3 \text{ с}; & \bar{\tau}_{4-5} &= \bar{\tau}_4 + \bar{\tau}_5 = 5,2 \text{ с}; \\ D_{1-3} &= D_1 + D_2 + D_3 = 3 \text{ с}^2; & D_{4-5} &= D_4 + D_5 = 3,2 \text{ с}^2; \\ p_{1-3} &= p_1 p_2 p_3 = 0,97; & p_{4-5} &= p_4 p_5 = 0,97; \\ \bar{\tau}_{9-10} &= \bar{\tau}_9 + \bar{\tau}_{10} = 5,4 \text{ с}; & \bar{\tau}_{11-12} &= \bar{\tau}_{11} + \bar{\tau}_{12} = 12,6 \text{ с}; \\ D_{9-10} &= D_9 + D_{10} = 1,8 \text{ с}^2; & D_{11-12} &= D_{11} + D_{12} = 7 \text{ с}^2; \\ p_{9-10} &= p_9 p_{10} = 0,96; & p_{11-12} &= p_{11} p_{12} = 0,96;\end{aligned}$$

2. Объединение путей с разветвлениями



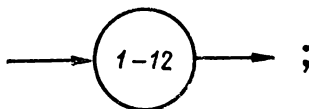
$$\begin{aligned}\bar{\tau}_{1-6} &= \bar{\tau}_{1-3} + \bar{\tau}_6 + \gamma_1 \bar{\tau}_{4-5} = 9 \text{ с}; \\ D_{1-6} &= D_{1-3} + D_6 + \gamma_1 D_{4-5} = 5,3 \text{ с}^2; \\ p_{1-6} &= p_{1-3} p_6 [(1 - \gamma_1) + \gamma_1 p_{4-5}] = 0,95.\end{aligned}$$

3. Объединение путей с разветвлениями



$$\begin{aligned}\bar{\tau}_{1-8} &= \bar{\tau}_{1-6} + \bar{\tau}_8 + \gamma_2 \bar{\tau}_7 = 11,4 \text{ с}; \\ D_{1-8} &= D_{1-6} + D_8 + \gamma_2 D_7 = 6,5 \text{ с}^2; \\ p_{1-8} &= p_{1-6} p_8 [(1 - \gamma_2) + \gamma_2 p_7] = 0,94.\end{aligned}$$

4. Объединение путей с разветвлениями



$$\bar{\tau}_{1-12} = \bar{\tau}_{1-8} + \bar{\tau}_{11-12} + \gamma_3 \bar{\tau}_{9-10} = 25,8 \text{ с};$$

$$D_{1-12} = D_{1-8} + D_{11-12} + \gamma_3 D_{9-10} = 14,1 \text{ с}^2;$$

$$p_{1-12} = p_{1-8} p_{11-12} [(1 - \gamma_3) + \gamma_3 p_{9-10}] = 0,89.$$

Таким образом, задача управления источниками информации имеет следующие характеристики: среднее время решения $\bar{\tau} = 25,8 \text{ с}$; среднеквадратическое отклонение времени решения $\sigma = 3,8 \text{ с}$; вероятность безошибочного решения $P = 0,89$.

Значение этих характеристик позволило оценить, насколько условия работы расчета удовлетворяют требованиям, предъявляемым к этим характеристикам реальной обстановкой.

Применение корреляционно-регрессионного метода при исследовании процесса восстановления аппаратуры. Основными функциями обслуживающего персонала АССОРЛИ являются контроль за исправностью функционирования аппаратуры и оперативное устранение неисправностей в случае выхода ее из строя. Общепринятым показателем качества выполнения этой функции является среднее время восстановления после отказа или время ремонта.

Задача заключается в том, чтобы найти зависимость времени восстановления аппаратуры от характеристик аппаратуры и обслуживающего персонала. Здесь мы имеем дело с многофакторной зависимостью, в которой можно предположить наличие взаимной связи факторов. Для получения этой зависимости воспользуемся методами корреляционно-регрессионного анализа, изложенными в § 3.

Психофизиологический анализ трудовой деятельности персонала, обслуживающего АССОРЛИ, проведенный по методике, и обобщение опыта эксплуатации ряда конкретных образцов аппаратуры позволили выявить факторы, которые в общем случае можно считать влияющими на время восстановления. Все эти факторы можно условно разбить на четыре группы:

- 1) характеристики обслуживающего персонала;
- 2) уровень сложности аппаратуры;
- 3) качество системы контроля;
- 4) степень удобства ремонта.

Перечень всех факторов и их возможная числовая интерпретация приведены в табл. 14.

Для исследования влияния факторов на время восстановления аппаратуры были собраны статистические данные по времени восстановления и определяющим его фактора для 177 участков аппаратуры, входящих в состав 19 систем.

Предварительный анализ исходных данных показал необходимость разбиения полученного статистического материала на две совокупности. В первую включаются данные по времени восстановления

Т а б л и ц а 14

| Группа факторов | Название фактора | Условное обозначение | Единица измерения | Возможная числовая интерпретация |
|---|---|----------------------|-------------------|--|
| Характеристики обслуживающего персонала | Квалификационный уровень специалиста | x_1 | балл | 1 балл — специалист низкой квалификации с малым опытом работы 2 балла — специалист низкой квалификации с большим опытом работы 3 балла — специалист средней квалификации (техник) 4 балла — специалист высокой квалификации (инженер) |
| | Общий стаж работы специалиста | x_2 | год | |
| | Стаж работы на исследуемой системе | x_3 | » | |
| | Уровень сложности аппаратуры | x_4 | штука | |
| Качество системы контроля | Количество функциональных частей | x_5 | » | |
| | Количество типов функциональных частей | x_6 | » | |
| | Количество уровней схемно-конструктивного построения | x_7 | » | 1 балл — вычислительные машины, дискретная аппаратура кодирования и передачи данных |
| | Тип аппаратуры (по элементной базе) | x_8 | | 2 балла — видеотракты РЛС, индикаторная аппаратура 3 балла — высоковольтная аппаратура, антенно-фидерный тракт |
| | Коэффициент внутри-проектной унификации | x_9 | 0—1 | $x_9 = \ln K_n / \ln x_5$ $K_n = x_5 / x_8$ |
| | Вид контроля (по уровню автоматизации) | x_{10} | балл | 1 балл — аппаратный 2 балла — программно-тестовый |
| | Количество элементов индикации | x_{11} | штука | |
| | Количество контрольных гнезд | x_{12} | » | |
| Степень удобства ремонта | Количество регулировок переключателей | x_{13} | » | |
| | Масштаб типового элемента замены | x_{14} | балл | 1 балл — деталь 2 балла — модуль 3 балла — блок |
| | Качество конструктивного исполнения аппаратуры (удобство сборки, разборки, доступа и пр.) | x_{15} | » | 2 балла — плохое |
| | | | | 3 балла — удовлетворительное |
| | | | | 4 балла — хорошее 5 баллов — отличное |

| Группа факторов | Название фактора | Условное обозначение | Единица измерения | Возможная числовая интерпретация |
|--------------------------|--|----------------------|-------------------|--|
| Степень удобства ремонта | Качество эксплуатационной документации (полнота и четкость указаний по ремонту, перечень характерных неисправностей, трактовок схем и пр.) | x_{16} | балл | оценивается аналогично качеству конструктивного исполнения |

аппаратуры при стаже работы специалистов, эксплуатирующих эту аппаратуру, менее трех лет (период обучения), во вторую — более трех лет (период обученности). Необходимость такого разбиения объясняется тем, что в период обучения специалиста основное значение имеет такой фактор, как стаж работы, из-за чего влияние аппаратурных факторов становится малозаметным. В период обученности, когда время восстановления стабилизируется, сильнее проявляется влияние аппаратурных факторов.

Для периода обучения было получено следующее уравнение регрессии:

$$T_{B(\min)} = 80,2 + 14x_1 - 1,64x_2 - 3,8x_3 - 8,1x_{10} - 2,5x_9 - 28,3x_{14}. \quad (5.7)$$

Полученное уравнение значимо при 5-процентном уровне значимости. Проверка по t -критерию показала, что остальные факторы незначимы. Коэффициент множественной корреляции $R \approx 0,7$, т. е. вариация времени восстановления на 70 %, объясняется вариацией факторов, входящих в уравнение (5.7).

Таким образом, на время восстановления в период обучения влияют квалифицированный уровень оператора, общий стаж работы специалиста, стаж работы на исследуемой системе, вид контроля аппаратуры (по уровню автоматизации), масштаб поиска неисправности и коэффициент внутрипроектной унификации.

Физическая сущность влияния на время восстановления аппаратуры характеристики обслуживающего персонала очевидна: чем больше опыт и квалификация специалиста, тем быстрее он устраняет неисправность.

При увеличении автоматизации контроля время восстановления существенно уменьшается за счет уменьшения времени поиска адреса неисправности.

С увеличением масштаба поиска неисправности процесс поиска значительно упрощается. Кроме того, как правило, сокращается трудоемкость замены отказавшей составной части. Все это вместе взятое уменьшает время восстановления аппаратуры.

Степень единообразия схемных и конструктивных решений, определяемая коэффициентом внутрипроектной унификации, влияет на

время освоения аппаратуры. Чем больше одинаковых по устройству и принципам действий составных частей имеет аппаратура, тем быстрее она осваивается и тем меньше время восстановления.

Для периода обученности было получено следующее уравнение регрессии:

$$T_{B(\min)} = 54 - 0,66x_2 - 0,0056x_8 - 4,14x_{11} - 6,7x_{14}. \quad (5.8)$$

Полученное уравнение значимо при 5-процентном уровне значимости. Коэффициент корреляции $R = 0,84$, что говорит о хороших прогностических возможностях уравнения.

Существенное влияние на время восстановления в период обученности специалиста оказывают общий стаж работы, тип аппаратуры (по элементарной базе), количество элементов индикации, масштаб поиска неисправности.

На участке обученности общий стаж работы по-прежнему оказывает (хотя и незначительное) влияние на время восстановления, что свидетельствует о непрекращающемся процессе совершенствования навыков специалиста.

Влияние типа аппаратуры на время восстановления связано с тем, что тип аппаратуры определяет степень трудности устранения неисправности: устранить неисправность антенно-фидерного тракта, например, значительно труднее, чем полупроводниковой аппаратуры.

Существенно влияет на время восстановления количество элементов индикации. Наличие индикации неисправности составных частей аппаратуры значительно облегчает поиск неисправностей. Однако следует отметить, что неоправданно большое количество элементов индикации создает трудности индикации соответствующим частям аппаратуры.

Таким образом, корреляционно-регрессионный анализ позволил вскрыть физическую сущность влияния характеристик обслуживающего персонала на время восстановления аппаратуры. Это позволяет решать задачу оптимизации технического обслуживания аппаратуры, обоснованно выбирая ее характеристики.

Исследование производительности оператора сопровождения методами теории массового обслуживания. Основная задача оператора пункта обработки радиолокационной информации заключается в своевременной завязке и обеспечении непрерывного сопровождения трасс воздушных объектов.

В процессе сопровождения объектов оператором возможны случаи, когда радиолокационный пост (РЛП) прекращает выдачу информации, по которой производится расчет трассы. В этих случаях оператор, чтобы обеспечить непрерывное сопровождение воздушного объекта, осуществляет операцию «переприем». Если время переприема превышает некоторую величину $\tau_{\text{ожл}}^{\text{оп}}$ (константа алгоритма), то происходит разрыв трассы воздушного объекта.

Основной характеристикой оператора является **производительность**, определяемая количеством трасс в единицу времени, которые оператор может своевременно завязать и непрерывно сопровождать.

Для экспериментально-теоретической оценки производительности оператора и выявления основных факторов, влияющих на производительность, используем интерпретацию оператора одноканальной системой массового обслуживания с ограниченным временем ожидания и относительными приоритетами, описанную в § 3.

Дадим краткую характеристику оператора как системы массового обслуживания.

На вход системы массового обслуживания поступает два потока заявок.

Поток заявок первого типа («переприем») характеризуется плотностью потока λ_1 , законом распределения времени обслуживания $f_1(t)$ и допустимым временем ожидания $\tau_{ож1}^{доп}$.

Поток заявок второго типа («заявка трассы») характеризуется плотностью потока λ_2 , законом распределения времени обслуживания $f_2(t)$ и допустимым временем ожидания $\tau_{ож2}^{доп}$.

Если оператор не успел осуществить «переприем», то он обязан произвести повторную заявку трассы. Это обусловило две особенности рассматриваемой системы массового обслуживания:

1) заявки первого типа обладают относительным приоритетом по отношению к заявкам второго типа;

2) необслуженные заявки первого типа превращаются в заявки второго типа.

Учитывая сказанное, плотности потоков заявок λ_1 и λ_2 можно определить следующим образом:

$$\lambda_1 = \lambda_u N_p;$$

$$\lambda_2 = \lambda_u + \lambda_1(1 - p_{01}) = \lambda_u [1 + N_p(1 - p_{01})], \quad (5.9)$$

где λ_u — плотность входа целей в зону ответственности ПОРИ; N_p — среднее количество разрывов (отсутствие информации от приоритетного РЛП) в пересчете на одну трассу; p_{01} — вероятность своевременного обслуживания заявок первого типа.

Характер законов распределения времени обслуживания $f_1(t)$ и $f_2(t)$ был определен экспериментальным путем. Законы распределения $f_1(t)$ и $f_2(t)$ имеют следующий вид (рис. 41, 42):

$$\begin{aligned} f_1(t) &= A_1 v_1^2 t \exp(-v_1 t), \\ f_2(t) &= \\ &= A_2 v_2^2 (t - t_{02}) \exp(-v_2 (t - t_{02})), \end{aligned} \quad (5.10)$$

где $A_1 = 1,23$; $A_2 = 1,1$ — параметры усечения; $t_{1 \min} = 3$ с; $t_{2 \min} = 7$ с — минимальное время обслужи-

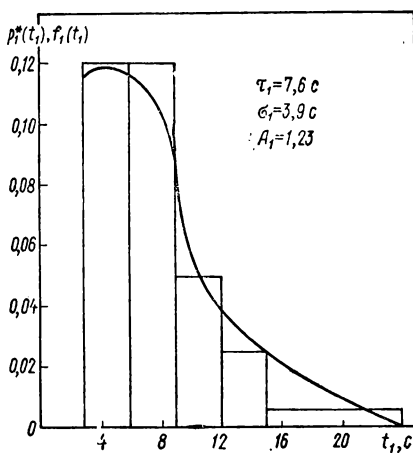


Рис. 41

живания заявок первого и второго типа; $t_{02} = 3$ с — величина сдвига закона распределения $f_2(t)$; $v_1 = 2/m_1$; $v_2 = 2/m_2$; $m_1 = 7,6$ с; $m_2 = 16,1$ с.

Параметры усечения A_1 и A_2 находились из условия нормировки

$$\int_{t_{\min}}^{\infty} A f(t) dt = 1, \quad (5.11)$$

откуда

$$A_1 = \frac{1}{(v_1 t_{1 \min} + 1) \exp(-v_1 t_{1 \min})}, \quad (5.12)$$

$$A_2 = \frac{1}{[v_2 (t_{2 \min} - t_{02}) + 1] \exp[v_2 (t_{2 \min} - t_{02})]}.$$

Основными показателями рассматриваемой системы массового обслуживания являются вероятности обслуживания заявок каждого типа

$$p_{01} = \lambda_1^0 / \lambda_1; \quad p_{02} = \lambda_2^0 / \lambda_2, \quad (5.13)$$

где λ_1 и λ_2 — интенсивность потока обслуженных заявок первого и второго типа соответственно.

Общий подход к расчету показателей p_{01} и p_{02} описан в § 3.

В нашем случае при $i^* = 2$ имеем систему двух уравнений:

$$p_{01} = \rho_1^0 / \rho_1 = 1 - \{\rho_0 + \tau_2 \lambda_2 [1 + N_p (1 - \rho_1^0 / \rho_1)]\} \exp\{-(1 - \rho_0^0) \varepsilon_{01}\}. \quad (5.14)$$

$$p_{02} = \rho_2^0 / \rho_2 = 1 - (\rho_1^0 + \rho_2^0) \exp\{-(1 - \rho_1^0)(1 - \rho_1^0 + \rho_2^0) \varepsilon_{02}\}.$$

Решение системы трансцендентных уравнений производится графически, методом последовательных приближений, начиная со старшего приоритета при фиксированных параметрах заявок второго типа.

Графический метод состоит

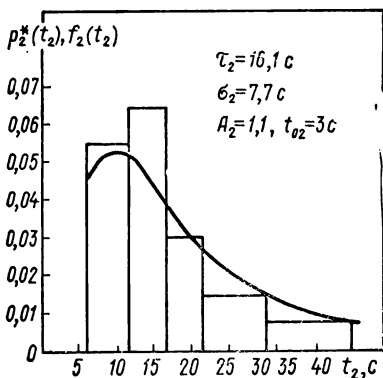


Рис. 42

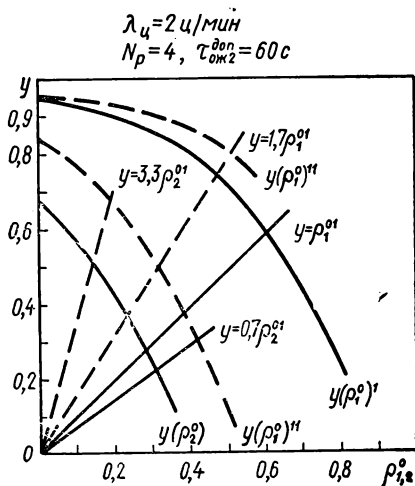


Рис. 43

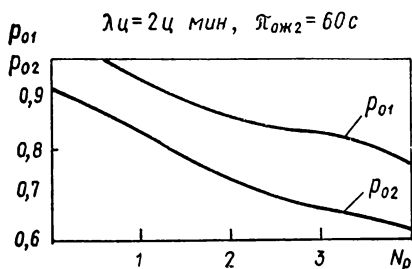


Рис. 44

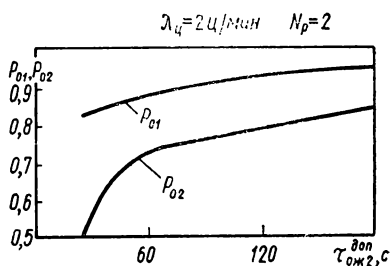


Рис. 45

в непосредственном определении координат точки пересечения луча $y = \rho_1^0 / \rho_1$ с кривой вида

$$y = 1 - \{\rho_1^0 + \tau_2 \lambda_u [1 + N_0 (1 - \rho_1^0 / \rho_1)]\} \exp \{-(1 - \rho_1^0) \varepsilon_{01}\}. \quad (5.15)$$

Полученное значение величины ρ_1^0 подставляется во второе уравнение, после чего оно решается аналогично.

Закончив первый цикл приближения, получаем набор величин $\rho_1^0(1)$, $\rho_2^0(1)$, $\rho_{01}(1)$, $\rho_{02}(1)$. Второй цикл приближения производится по той же методике, но с использованием полученных величин ρ_1^0 и ρ_2^0 . Практически для получения величин ρ_{01} и ρ_{02} достаточно провести два цикла.

Пример графического решения системы уравнений приведен на рис. 43.

При заданной аппаратуре и алгоритмах величины τ_1 , τ_2 и $\tau_{ож1}^{доп}$ можно считать фиксированными. Тогда значения вероятностей обслуживания ρ_{01} и ρ_{02} зависят от величины N_p , $\tau_{ож2}^{доп}$ и λ_u . Для исследования характера указанных зависимостей были произведены расчеты, результаты которых представлены на графиках (рис. 44—46). Указанные графики использовались для оценки производительности оператора.

Производительность оператора определялась плотностью потока целей, при которой качество решения задач достаточно высокое, т. е.

$$\rho_{01} \geq \rho_1^*; \quad \rho_{02} \geq \rho_2^*; \quad \rho_1^* = \rho_2^* = 0,8 \dots 0,9.$$

Анализ зависимостей (рис. 44, 45) позволил выявить влияние величин N_p и $\tau_{ож}^{доп}$ на производительность оператора. По результатам анализа можно сделать следующий вывод:

производительность оператора существенным образом зависит от качества информации, поступающей с РЛС (непрерывности проводки целей); например, если при $N_p = 0,5$ производительность оператора составила 2 ц/мин, при $N_p = 2$ она оказалась равной 1 ц/мин ($\rho_1^* = \rho_2^* = 0,9$; $\tau_{ож2}^{доп} = 60$ с);

автоматизация решения задачи «переприемка» позволяет увеличить

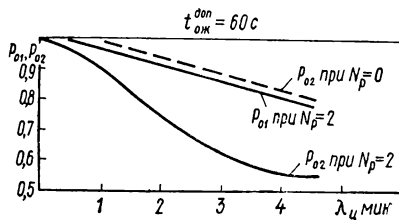


Рис. 46

производительность оператора до величины, равной примерно 3 ц/мин;

влияние времени задержки завязки трассы $\tau_{\text{ож2}}^{\text{доп}}$ на производительность оператора значительно менее существенно, чем влияние величины N_p ; поэтому производительность оператора не может быть существенно увеличена за счет снижения требования к времени завязки трассы.

Исследование работы оператора сопровождения методом статистического моделирования. Содержанием деятельности оператора сопровождения является прием информации, поступающей в виде отдельных заявок, ее переработка, принятие решения и осуществление исполнительных действий путем двигательных актов. Рассмотрим кратко основные рабочие функции оператора сопровождения. Он осуществляет захват цели на сопровождение, выполняя операции «первый ввод», «запрос при корректуре», «второй ввод», присваивает признаки государственной принадлежности и количественного состава цели и помех. Последующее сопровождение целей оператор обеспечивает периодическим вводом корректур. При выходе сопровождаемой цели из назначенного сектора он осуществляет условный сброс с сопровождения.

Деятельность оператора сопровождения относится к сенсомоторному виду деятельности, который достаточно легко поддается моделированию. Целью моделирования реальной работы оператора является получение характеристик качества сопровождения целей (времени завязки трассы цели и среднего темпа ввода корректур, определяющего точность сопровождения). Для решения этой задачи используем представление оператора в виде системы массового обслуживания. Оператор сопровождения рассматривается как одноканальная система массового обслуживания с ожиданием и относительными приоритетами.

Заявки поступают на обслуживание в случайные моменты времени. При отсутствии заявок оператор находится в состоянии ожидания. Заявки могут быть шести видов (первый ввод, запрос и второй ввод, присвоение характеристик, простая корректура, сложная корректура, сброс цели с сопровождения). Времена обслуживания заявок оператором — величины случайные, с нормальным законом распределения. Если во время обслуживания заявки поступает еще одна или несколько новых заявок, то они образуют очередь, ожидая обслуживания.

На основе анализа особенностей деятельности оператора можно предположить, что в алгоритме оператора сопровождения должны быть реализованы три основные функции:

- формирование заявок оператором;
- выбор оператором заявок для обслуживания;
- обслуживание заявок.

Заявки формируются по следующим признакам:

«первый ввод» — цель видна и не сопровождается;

«второй ввод» — цель видна, первый ввод воспринят ЭВМ, число обзоров после первого ввода менее трех;

«присвоение характеристик» — второй ввод воспринят ЭВМ, цель видна, нет характеристик;

«простая корректура» — цель видна, корректура не вводилась в течение трех обзоров;

«сложная (специальная) корректура» — между эхо-сигналом и экстраполяционной точкой имеется большое рассогласование;

«сброс цели с сопровождения» — корректура не вводилась в течение восьми обзоров, второй ввод не вводился в течение пяти обзоров после осуществления первого ввода, а также по команде старшего оператора.

Выбор заявок на обслуживание определяет очередность обслуживания заявок оператором. В первую очередь обслуживаются заявки,

Т а б л и ц а 15

| Выполняемые операции | Временные характеристики оператора | | | | | |
|--------------------------|------------------------------------|-----|-----|--------|-----|-----|
| | МОЖ, с | | | СКО, с | | |
| | Оценка | | | | | |
| | 5 | 4 | 3 | 5 | 4 | 3 |
| Первый ввод | 2 | 2,5 | 2,7 | 0,3 | 0,5 | 0,6 |
| Второй ввод | 2,8 | 3,5 | 4,1 | 0,4 | 0,7 | 0,9 |
| Присвоение характеристик | 1,7 | 2,2 | 2,5 | 0,3 | 0,4 | 0,5 |
| Простая корректура | 1,7 | 2 | 2,3 | 0,3 | 0,4 | 0,5 |
| Сложная корректура | 2,2 | 2,6 | 2,8 | 0,4 | 0,6 | 0,8 |
| Сброс с сопровождения | 1,8 | 2,3 | 2,5 | 0,3 | 0,4 | 0,5 |

невыполнение которых может привести к исключению цели из обработки. Заявка, обслуживание которой не закончено в данном обзоре, обслуживается в следующем обзоре. Таким образом, начатое обслуживание заявки доводится до конца. Установлен следующий относительный приоритет заявок: первый ввод, второй ввод, присвоение характеристик, ввод корректуры, сброс цели с сопровождения. Функцию обслуживания заявок имитирует датчик случайных чисел, воспроизводящий время обслуживания заявок. Обработка данных испытаний показала, что с достаточной для практики точностью можно считать время решения частных задач распределенным по нормальному закону. Ошибки съема координат оператором сопровождения также распределены по нормальному закону, параметры которого зависят в основном от загрузки и уровня подготовки оператора.

В качестве исходных данных в алгоритме задаются параметры нормального закона, по которому распределены время обслуживания заявок и ошибки по координатам, допущенные оператором при обслуживании заявок. Математическое ожидание и среднеквадратическое отклонение зависят от подготовки оператора сопровождения. Экспериментальным путем получены характеристики времени выполнения оператором сопровождения (табл. 15).

Характеристики точности выполнения операций, зависящие от количества сопровождаемых целей и уровня подготовки оператора, приведены в табл. 16. Точность определяется максимальной ошибкой снятия координат цели по дальности ΔD и азимуту $\Delta \beta$.

Поток входных заявок разных типов принят пуассоновским с различными значениями параметра потока λ . Плотность потока заявок типа «первый ввод» определяется плотностью входа целей в сектор оператора. Заявки типа «второй ввод» формируются после выполнения первого ввода, а заявки «присвоение характеристики» — после второго ввода. Заявки типа «простая и сложная корректура» формируются один раз в три обзора. Наконец, сброс целей с сопровождения

реализуется при выходе цели из сектора оператора или по указанию старшего оператора.

С учетом содержания и особенностей деятельности оператора был разработан алгоритм, моделирующий работу оператора сопровождения.

В соответствии с разработанным алгоритмом была реализована комплекс-программа и проведе-

но моделирование. Варианты моделирования отличались количеством одновременно сопровождаемых целей, а также уровнем подготовки оператора. В результате моделирования были получены зависимости показателей качества сопровождения (времени завязки трассы и среднего темпа ввода корректур) от количества сопровождаемых целей и уровня подготовки оператора. При этом средний темп ввода корректур определялся из соотношения

$$n_{\text{ср}} = \sum_{i=1}^M N_{ik} / (nM), \tag{5.16}$$

где N_{ik} — количество корректур по i -й цели ($i = 1, \dots, M$); n — количество обзоров; M — количество целей, находящихся на сопровождении.

Вид указанных зависимостей представлен на графиках (рис. 47, 48). Из этих графиков, в частности, следует, что оператор с отличной подготовкой может сопровождать до шести целей. При этом среднее время завязки трассы не превышает 90 с, а корректура вводится примерно раз в два обзора.

Экспериментальная оценка работы оператора управления. Оператор управления, как правило, входит в состав группы управления КП и выполняет следующие основные функции:

распределяет цели между источниками информации и выдает целеуказания для своевременного взятия целей на сопровождение;

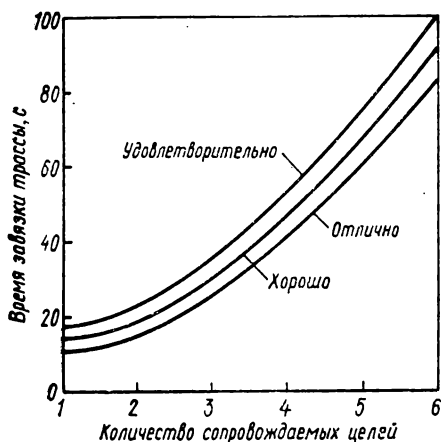


Рис. 47

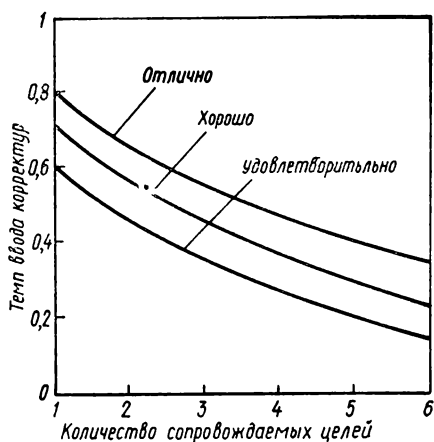


Рис. 48

контролирует качество и непрерывность сопровождения целей источником путем выдачи целеуказания и команд на восстановление трасс;

контролирует своевременность взятия на сопровождение взлетевших перехватчиков путем выдачи целеуказаний и команд;

контролирует правильность привязки характеристик, принадлежности, состава и высоты; разрешает конфликтные ситуации по государственной принадлежности;

регулирует загрузку источников путем выдачи команды на сброс отдельных целей с сопровождения.

Основной целью экспериментального исследования было определение производительности оператора управления, т. е. количества одновременно управляемых источников информации в условиях достаточно сложной воздушно-помеховой обстановки.

Эксперименты отличались количеством источников информации, управляемых с автоматизированного рабочего места оператором управления; оно изменялось от одного до четырех.

По результатам каждого эксперимента, как это описано в § 4, строилась временная диаграмма работы оператора управления. Далее временная диаграмма анализировалась с целью выделения и группирования статистических данных по отдельным задачам. Для каждой задачи ($i = 1, \dots, 5$) в каждом из экспериментов ($k = 1, \dots, 4$) определялись:

количество возникших задач n_{ik} ;

количество правильно и своевременно решенных задач τ_{ik} ;

среднее время решения задачи в эксперименте τ_{ik} .

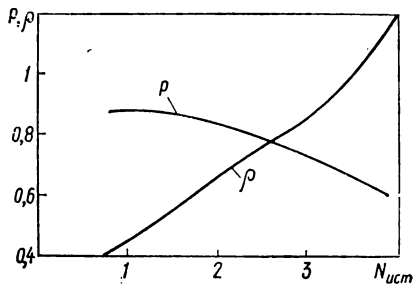


Рис. 49

Т а б л и ц а 17

| Решаемая задача ($i = 1, \dots, 5$) | $k = 1$ | | | $k = 2$ | | | $k = 3$ | | | $k = 4$ | | |
|--|----------------------|----------|--------------|----------------------|----------|--------------|----------------------|----------|--------------|----------------------|----------|--------------|
| | $N_{\text{ист}} = 1$ | | | $N_{\text{ист}} = 2$ | | | $N_{\text{ист}} = 3$ | | | $N_{\text{ист}} = 4$ | | |
| | n_{ik} | m_{ik} | τ_{ikc} | n_{ik} | m_{ik} | τ_{ikc} | n_{ik} | m_{ik} | τ_{ikc} | n_{ik} | m_{ik} | τ_{ikc} |
| Распределение целей между источниками | 11 | 9 | 30 | 16 | 14 | 30 | 22 | 19 | 30 | 31 | 23 | 30 |
| Восстановление трасс целей | 14 | 12 | 40 | 21 | 17 | 40 | 29 | 20 | 40 | 42 | 17 | 40 |
| Обеспечение захвата перехватчиков | 5 | 5 | 20 | 8 | 8 | 20 | 11 | 10 | 20 | 14 | 10 | 20 |
| Разрешение конфликтных ситуаций по ОГП | 9 | 7 | 50 | 13 | 10 | 50 | 18 | 12 | 50 | 21 | 12 | 50 |
| Регулирование загрузки источников | 3 | 3 | 20 | 7 | 6 | 20 | 11 | 9 | 20 | 18 | 12 | 20 |

Т а б л и ц а 18

| Решаемая задача | Эксперимент | | | | | | | |
|--|----------------------|----------|----------------------|----------|----------------------|----------|----------------------|----------|
| | $N_{\text{ист}} = 1$ | | $N_{\text{ист}} = 2$ | | $N_{\text{ист}} = 3$ | | $N_{\text{ист}} = 4$ | |
| | p_{i1} | p_{i1} | p_{i2} | p_{i2} | p_{i3} | p_{i3} | p_{i4} | p_{i4} |
| Распределение объектов между источниками | 0,09 | 0,82 | 0,13 | 0,87 | 0,18 | 0,26 | 0,26 | 0,74 |
| Восстановление трасс объектов | 0,16 | 0,86 | 0,23 | 0,81 | 0,32 | 0,69 | 0,47 | 0,41 |
| Обеспечение захвата | 0,03 | 1 | 0,05 | 1 | 0,06 | 0,91 | 0,08 | 0,71 |
| Разрешение конфликтных ситуаций | 0,12 | 0,78 | 0,18 | 0,77 | 0,25 | 0,67 | 0,29 | 0,57 |
| Регулирование загрузки источников | 0,02 | 1 | 0,04 | 0,86 | 0,06 | 0,82 | 0,1 | 0,67 |
| Итоговые данные | 0,42 | 0,86 | 0,63 | 0,84 | 0,87 | 0,77 | 1,2 | 0,59 |

Статистические данные по отдельным задачам в разных экспериментах сведены в табл. 17. Они позволяют рассчитать загрузку оператора управления решением возникающих задач p_{ik} и вероятность правильного решения каждой задачи p_{ik} , а также определить суммарную загрузку ρ и вероятность безошибочной работы P при разном количестве управляемых источников.

Результаты расчетов приведены в табл. 18. Предполагалось, что среднее время эксперимента составляет около 60 мин.

Для определения производительности оператора управления был построен график зависимости величин ρ и P в функции от количества управляемых источников (рис. 49). В соответствии с рекомендациями оператор управления может управлять не более чем тремя источниками информации. При этом его загрузка не превышает допустимую ($\rho \leq 0,9$), а качество решения задач достаточно высокое ($P \geq 0,8$).

Определив производительность оператора управления, можно дать рекомендации по их количеству в составе расчета, а также по необходимому числу автоматизированных рабочих мест данного типа.

КОМПЛЕКС ПРОГРАММЫ «СОКИ-Н» СТАТИСТИЧЕСКОЙ ОБРАБОТКИ РЕЗУЛЬТАТОВ КАЧЕСТВЕННЫХ ИЗМЕРЕНИЙ В НОМИНАЛЬНЫХ ШКАЛАХ

Измерение в номинальной шкале (шкале наименований, назывной шкале, шкале классификаций) применяется в том случае, когда можно установить отличие одного состояния объекта измерения от других. Максимально возможная группа различающихся состояний объекта измерения и образует номинальную шкалу. Количество разных элементов шкалы (количество градаций) задается исследователем. Элементы номинальной шкалы можно кодировать с помощью произвольного набора символов. Наиболее удобен для ввода в ЭВМ числовой код. Процесс измерения в номинальной шкале осуществляется стандартным способом: измеряемая величина, т. е. состояние объекта измерения, сравнивается со шкалой возможных значений. В качестве результата измерения фиксируется код того элемента шкалы, с которым установлено отношение эквивалентности. Никаких других содержательных отношений, кроме «эквивалентно» и «неэквивалентно», между элементами номинальной шкалы нет. В номинальную шкалу всегда можно преобразовать измерения, произведенные по количественным (метрическим) шкалам и шкале рангов. Каждому объекту измерения соответствует своя номинальная шкала (арифметические операции для нее не имеют смысла). Статистически положение центра выражает мода.

Статистическая обработка результатов измерений в номинальных шкалах для комплекса «СОКИ-Н» включает:

- статистическую обработку одной выборки;
- оценку различия двух независимых выборок;
- оценку различных групп M независимых выборок;
- оценку различия двух связанных выборок (опыт «до — после»);
- оценку различия M связанных выборок (M условий опыта);
- оценку корреляционной связи двух выборок.

Для ввода в ЭВМ результаты измерений представляются в виде числовых кодов (в форме целых чисел) типа *INTEGER*. В случае оценки корреляционной связи между двумя выборками, одна из которых — результат измерения в номинальной шкале с двумя градациями, а другая — результат количественных измерений по шкале отношений, элементы последней вводятся в виде чисел с плавающей точкой типа *REAL*4*.

Основные исходные параметры для комплекса программ «СОКИ-Н» такие:

код типовой задачи;
 выборка числовых кодов — результаты измерений в номинальной шкале;
 длина выборки;
 количество градаций;
 массив уровней градаций — констант, определяющих правило распределения числовых кодов по градациям;
 уровень значимости риска.

Дополнительно задается служебная информация — идентификационные коды выборок, коды печати исходных данных и результатов, номер логического канала вывода, числовые параметры.

Структура комплекса программ «СОКИ-Н» представлена на рис. 50.

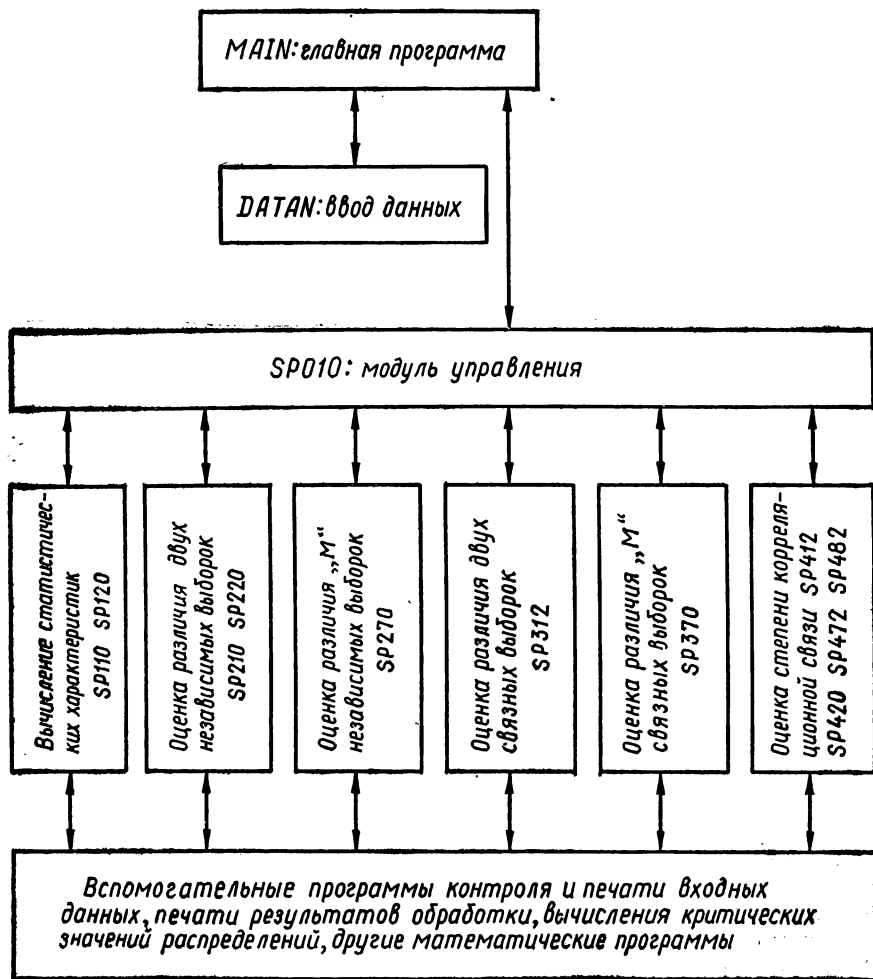


Рис. 50

Главная программа комплекса резервирует пространство, запрашивает имя файла со входной информацией, вызывает подпрограмму ввода данных и подпрограмму управления комплексом. После окончания обработки запрашивает условие продолжения работы, новое имя файла и т. д. или завершает выполнение программы.

Подпрограмма ввода данных реализует динамическое распределение рабочей памяти для данных, параметров и рабочих массивов комплекса, осуществляет ввод данных из начального файла («поток»). Информация из «потока» выбирается по ключевым словам, анализируется. В случае обнаружения ошибок выдаются диагностические сообщения. В «потоке» возможно использовать ссылки на имена вторичных файлов, содержащих элементы выборки или таблиц данных.

Модуль управления комплексом организует вызов функциональных модулей в соответствии с указанием на решение той или иной типовой задачи.

Каждому функциональному модулю соответствует своя подпрограмма ввода и контроля входных данных и подпрограмма печати результатов.

Из принятой структуры комплекса «СОКИ-Н» со всей очевидностью вытекают три варианта использования комплекса. Первый вариант — использование отдельных функциональных модулей (с их обрамлением) в программах пользователей, так как каждый модуль является самостоятельной единицей и выполняет одну из перечисленных задач. Второй вариант — использование всей группы функциональных модулей в целом под управлением модуля *SPC10*, через который и происходит вызов комплекса с указанием кода функции обработки. Оба варианта требуют участия программиста для организации вызываемой программы. Третий вариант — использование «автономного» варианта комплекса «СОКИ-Н», дополнительно включающего главную программу и подпрограмму ввода данных. Этот вариант не требует знания программирования, необходимы только навыки по заполнению стандартной формы подготовки информации.

Программа «СОКИ-Н» предусматривает твердый порядок ввода исходных данных. В соответствии с указанной в ней типовой задачей статистической обработки вводятся только те данные, которые требуются по программе, остальные пропускаются (рассматриваются как комментарии). Идентификация данных производится по ключевым словам. Исходные данные могут находиться в одном или нескольких файлах. Основной источник данных — «поток» — создаст я по стандартной форме — макету подготовки данных. Остальные («вторичные») файлы могут содержать только массивы данных и список спецификаций оператора *FORMAT* в круглых скобках.

Управление выводом результатов обработки осуществляется с помощью массива ключей управления, кодирующих контроль и печать входных данных, а также результатов обработки для каждой функциональной подпрограммы.

В комплексе программ «СОКИ-Н» не предусмотрена программная обработка аварийных ситуаций, поэтому исходные данные желательно контролировать перед обработкой с целью повышения «живучести»

комплекса. Защита по исходным данным заключается в организации проверок значений параметров и исходных данных, предотвращающих возникновение аварийных ситуаций. В случае обнаружения неправильных значений комплекс прекращает работу и формирует сообщение об ошибке с указанием конкретных параметров.

Программа «СОКИ-Н» выполняется под управлением операционной системы РАФОС (RT-11). Для работы этой программы требуется 20k слов оперативной памяти, накопитель на гибком магнитном диске, дисплей. Исполняемая программа комплекса занимает на диске около 210 блоков. Объем обрабатываемых выборок — порядка 1000 элементов.

```

СС: SP010 - МОДУЛЬ УПРАВЛЕНИЯ КОМПЛЕКСА "СОКИ - Н"
С
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ УПРАВЛЕНИЯ РЕЖИМАМИ РАБОТЫ
С; КОМПЛЕКСА "СОКИ - Н" / НОМИНАЛЬНЫЕ ШКАЛЫ /
С
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : Н Е Т
С
СТ; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
С
СА П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ,
С; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
С
SUBROUTINE SP010(KEY,NFWR,IRP,NS,KSQN,MD,N,X,Z,T,K,M,C,Q
* J,Y,P,IND,IER,IERW )
С
С ОПИСАНИЕ ПАРАМЕТРОВ :
С -----
С
СР ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
С
С KEY I*4 * (2,8,4) МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ КОМПЛЕКСОМ
С NFWR I*4 * 1 НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
С IRP I*4 * 1 К О Д ВЫПОЛНЯЕМОЙ ФУНКЦИИ
С NS I*4 * 1 КОЛИЧЕСТВО ВЫБОРОК
С KSQN R*8 * (4,NS) МАССИВ К О Д О В ВЫБОРОК
С MD I*4 * 1 ПАРАМЕТР
С N I*4 * (NS) МАССИВ ДЛИН ВЫБОРОК
С X I*4 * (1) МАССИВ ЗНАЧЕНИЙ ЭЛЕМЕНТОВ ВЫБОРОК
С Z R*4 * (1) МАССИВ ЗНАЧЕНИЙ ЭЛЕМЕНТОВ ВЫБОРОК
С T R*4 * * (K,M) ТАБЛИЦА ИСХОДНЫХ ДАННЫХ , ВЫХОДНАЯ
С ТАБЛИЦА МУЛЬТИНОМИНАЛЬНЫХ ЧАСТОТ
С K I*4 * 1 КОЛИЧЕСТВО ГРАДАЦИИ
С M I*4 * 1 НОМЕР ИССЛЕДУЕМОЙ ГРАДАЦИИ
С C I*4 * (1) МАССИВ ЗНАЧЕНИЙ УРОВНЕЙ ДЛЯ ГРАДАЦИИ
С Q R*4 * 1 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
С J I*4 * (1) МАССИВ ВЫЧИСЛЕННЫХ ЗНАЧЕНИЙ ПАРАМЕТРОВ
С Y R*4 * (1) МАССИВ ВЫЧИСЛЕННЫХ ЗНАЧЕНИЙ ПАРАМЕТРОВ
С P R*4 * (1) МАССИВ ВЫЧИСЛЕННЫХ ЗНАЧЕНИЙ ДОЛЕЙ
С IND I*4 * (1) МАССИВ ЗНАЧЕНИЙ ИНДЕКСОВ ОТВЕТОВ
С IER I*4 * (1) МАССИВ ЗНАЧЕНИЙ ПРИЗНАКОВ ОШИБОК
С IERW I*4 * (1) МАССИВ ЗНАЧЕНИЙ ОШИБОК
С
С DIMENSION KEY(2,8,4), KL(8), KSQN(4,NS), N(NS),T(K,M)
С * ,X(1), Z(1), P(1), J(1), Y(1), C(1)
С * ,IND(1), IER(1), IERW(1)
С
С INTEGER X,C
С REAL*8 NAME, KSQN
С DATA NAME /' SP010 '/, EPS/0.001/, CONST/5.0/
С
С РАЗРЕШЕННЫЕ КОДЫ ВЫПОЛНЯЕМОЙ ФУНКЦИИ - IRP:
С
С КОДЫ ЗАДАЧА
С X01,X11,X21 1 - СТАТИСТИЧЕСКАЯ ОБРАБОТКА ВЫБОРКИ
С 102,112,122,132,142 2 - ОЦЕНКА СВЯЗИ 2-Х НЕЗАВИС. ВЫБОРОК
С 002,012,022,032,042 2 +ВЫЧИСЛ.КОЭФФ. КОРРЕЛЯЦИОННОЙ СВЯЗИ
С 0X3,1X3 3 - ОЦЕНКА СВЯЗИ "М" НЕЗАВИС. ВЫБОРОК
С XX4 4 - ОЦЕНКА СВЯЗИ 2-Х СВЯЗАННЫХ ВЫБОРОК
С XX5 5 - ОЦЕНКА СВЯЗИ "М" СВЯЗАННЫХ ВЫБОРОК
С 006,016,026 6 - ВЫЧИСЛЕНИЕ КОЭФФ. КОРРЕЛЯЦ. СВЯЗИ
С 106,116,126 6 2-Х ВЫБОРОК, ИЗМЕР. В НОМ. ШКАЛАХ
С X07,X17,X27,X37 7 - ВЫЧ. КОРР.СВЯЗИ ВЫБОРОК,ИЗМЕРЕННЫХ
С ОДНА - В НОМИНАЛЬНЫХ ШКАЛАХ,
С ДРУГАЯ - В ШКАЛЕ ОТНОШЕНИЙ.
С
С
С

```

```

С      Л О К А Л Ь Н Ы Е   П Е Р Е М Е Н Н Ы Е
С
С      'IS' - МЕТКА ВОЗВРАТА В УПРАВЛЯЮЩУЮ ПОСЛЕДОВАТЕЛЬНОСТЬ
С      'KV' - КОЛИЧЕСТВО ВАРИАНТОВ ПОПАРНЫХ ПРОВЕРОК В ДИСП. АН.
С      'KY' - ПОСЛЕДНИЙ ЗАНЯТЫЙ ЭЛЕМЕНТ МАССИВА 'Y'
С      'KJ' - ПОСЛЕДНИЙ ЗАНЯТЫЙ ЭЛЕМЕНТ МАССИВА 'J'
С      'J0' - ИНДЕКС НАЧАЛА ПОДМАССИВА 'P' В 'Y'
С      'J1' - ИНДЕКС НАЧАЛА ПОДМАССИВА 'ST' В 'Y'
С      'J2' - ИНДЕКС НАЧАЛА ПОДМАССИВА 'TS' В 'Y'
С      'J3' - ИНДЕКС НАЧАЛА ПОДМАССИВА 'TIN' В 'Y'
С      'J4' - ИНДЕКС НАЧАЛА ПОДМАССИВА 'NST' В 'J'
С      'J5' - ИНДЕКС НАЧАЛА ПОДМАССИВА 'INDST' В 'J'
С      'KL' - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ ПОДПРОГРАММАМИ
С      'INL' - УКАЗАТЕЛЬ НАЧАЛА ПОДМАССИВА В 'IERW'
С      'NY' - УКАЗАТЕЛЬ ТЕКУЩЕГО ЭЛЕМЕНТА ДЛЯ 'Y'
С      'NL' - СЧЕТЧИК НАКОПЛЕНИЯ ДЛЯ 'IND', 'IER', 'IERW'
С      'IP' - УКАЗАТЕЛЬ ПЕРВОГО ЭЛЕМЕНТА МАССИВА ДОЛЕЙ
С      'IX' - УКАЗАТЕЛЬ ПЕРВОГО ЭЛЕМЕНТА МАССИВА ИСХ. ДАННЫХ
С      'IR' - КОД ТИПОВОЙ ЗАДАЧИ
С      'IPR' - КОД ВАРИАНТА ТИПОВОЙ ЗАДАЧИ
С      'ISR' - КОД СПОСОБА ПРЕДСТАВЛЕНИЯ ДАННЫХ
С
С      IRP = IR + 10*IPR + 100*ISR
С
С      ISR = IRP / 100
С      IPR = MOD(IRP,100) / 10
С      IR = IRP - 100*ISR - 10*IPR
С
С      NL = 1
С      INL = 1
С      NJ = 1
С      NY = 1
С      IP = 1
С      IX = 1
С
С      GO TO ( 100, 200, 300, 400, 500, 600, 700 ), IR
С      IR = 1
С      ---- ВЫЧИСЛЕНИЕ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК ВЫБОРОК ----
С      IR = 1 : IPR=1,11; ISR=21.
100 ASSIGN 9999 TO IS
      IF ( IPR .EQ. 2 ) GO TO 1120
      GO TO 1110
С      IR = 2
С      ---- ОЦЕНКА РАЗЛИЧИЯ 2-Х НЕЗАВИСИМЫХ ВЫБОРОК ----
С      IR = 2: IPR=X02,X12,X22,X32,X42; ГДЕ X=0,1,2,3.
С      *** ВЫЧИСЛЕНИЕ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК ВЫБОРОК ***
200 IF ( MOD(ISR,2) .EQ. 0 ) GO TO 202
      DO 201 I=1,K
      P(I) = Z(I)
201 P(I+K) = Z(I+K)
202 ASSIGN 203 TO IS
      GO TO 1120
203 IF (IABS(IER(1)) .GE. 3 .OR. IABS(IER(2)) .GE. 3) GO TO 299
С      +++++ ПА Р А М Е Т Р И Ч Е С К И Е   К Р И Т Е Р И И +++++
205 IV = IPR+1
      GO TO ( 206, 206, 206, 230, 240 ), IV
С      IPR = X02; X12; X22; X32; X42;
С      ПРОВЕРКА НА ДОПУСТИМОСТЬ Т-КРИТЕРИЯ И ИНТЕРВАЛЬНОГО ОЦЕИВАНИЯ
206 DO 207 I=1,K
      IF ((P(I) *N(1)+EPS) .LT. CONST) GO TO 225
      IF ((P(K+I)*N(2)+EPS) .LT. CONST) GO TO 225
207 CONTINUE
С      *** Т-КРИТЕРИИ *** ОЦЕНКА РАЗЛИЧИЯ 2-Х ВЫБОРОК ***
210 ASSIGN 220 TO IS
      IF( IPR .NE. 0 .AND. IPR .NE. 1 ) GO TO 1215
      GO TO 1210

```

```

C *** ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ *** ОЦЕНКА РАЗЛИЧИЯ 2-Х ВЫБОРОК **
220 ASSIGN 230 TO IS
    IF( IPR .NE. 0 .AND. IPR .NE. 2 ) GO TO 1225
    GO TO 1220
C +++ Н Е П А Р А М Е Т Р И Ч Е С К И Е   К Р И Т Е Р И И +++
225 NY = NY+K+5
    NJ = NJ+K
    NL = NL+2
    INL = INL+16
C *** КРИТЕРИЙ ХИ-КВАДРАТ *** ОЦЕНКА РАЗЛИЧИЯ 2-Х ВЫБОРОК **
C      В ТАБЛИЦЕ СОПРЯЖЕННОСТИ R*C / K*2 /
230 ASSIGN 240 TO IS
    IF( IPR .NE. 0 .AND. IPR .NE. 3 ) GO TO 1235
    GO TO 1230
C *** КРИТЕРИЙ Н12 *** - ОЦЕНКА РАЗЛИЧИЯ 2-Х ВЫБОРОК ***
C      ПРИ ЧИСЛЕ ГРАДАЦИИ > 3
C ПРОВЕРКА НА ДОПУСТИМОСТЬ ПРИМЕНЕНИЯ
240 IF( IPR .NE. 0 .AND. IPR .NE. 4 ) GO TO 290
    IF (K.LT.4) GO TO 290
    ASSIGN 290 TO IS
    GO TO 1240
C
290 IF(ISR .GE. 2) GO TO 299
    IF( IABS( IER(1) ) .GE. 2 .OR. IABS( IER(2) ) .GE. 2 ) GO TO 299
    GO TO 603
299 RETURN
C      IR = 3
C ---- ОЦЕНКА РАЗЛИЧИЯ "М" НЕЗАВИСИМЫХ ВЫБОРОК ----
C      IR = 3: IRP=XX3.
300 CONTINUE
C *** ДИСПЕРСИОННЫЙ ОДНОФАКТОРНЫЙ АНАЛИЗ ***
C ПРОВЕРКА НА ДОПУСТИМОСТЬ ПРИМЕНЕНИЯ ДИСПЕРСИОННОГО АНАЛИЗА
    IF ( NS .LT. 3 ) GO TO 390
    KV = NS*(NS-1)/2
    KY = 0
    IF(ISR .EQ. 0) GO TO 304
    DO 302 I=1,NS
302  Y(I)=Z(K*(I-1)+N)
    GO TO 320
304  KY = NS*8
    I=0
305  I=I+1
    ASSIGN 310 TO IS
    GO TO 1110
310  CONTINUE
    IF ( IND(NL-1) .GT. 1 ) GO TO 390
    Y(KY+I) = P(K*(I-1)+M)
    IF (I.LT.NS) GO TO 305
320  J0 = KY + 1
    KY = KY + NS
    J1 = KY + 3
    J2 = J1+KV*3
    J3 = J2+KV
C ВСЕГО ИСПОЛЬЗУЕТСЯ ((NS*8)+NS+2+KV*6) ЭЛЕМЕНТОВ МАССИВА "Y"
    J4 = 1
    J5 = J4+KV*2
C ВСЕГО ИСПОЛЬЗУЕТСЯ (KV*4) ЭЛЕМЕНТОВ МАССИВА "J"
    ASSIGN 390 TO IS
    GO TO 1300
C
390 RETURN
C      IR = 4
C ---- ОЦЕНКА РАЗЛИЧИЯ 2 - Х СЕЗЯАННЫХ ВЫБОРОК ----
C      ИЗМЕРЕННЫХ В ШКАЛЕ НАИМЕНОВАНИИ /ОПЫТ "ДО-ПОСЛЕ"/
C      IR = 4: IRP=XX4.
C

```

```

C *** КРИТЕРИИ МАКНИМАРА *** ДЛЯ 2-Х ВЫБОРОК -РЕЗУЛЬТАТОВ
C ОПЫТА "ДО-ПОСЛЕ", ПРИ ДВУХВАРИАНТНОМ ХАРАКТЕРЕ ОТВЕТА
400 ASSIGN 490 TO IS
GO TO 1400
C
490 RETURN
C IR = 5
C ---- ОЦЕНКА РАЗЛИЧИЯ "М" СВЯЗАННЫХ ВЫБОРОК ----
C IR = 5; IRP=XX5.
C
C *** КРИТЕРИИ КОКРЕНА ** ДЛЯ "М" ВЫБОРОК - РЕЗУЛЬТАТОВ
C ОПЫТА "ДО-ПОСЛЕ" ПРИ "М" УСЛОВИЯХ НАД "N" ОБЪЕКТАМИ
C / ПРИ ДВУХВАРИАНТНОМ ХАРАКТЕРЕ ОТВЕТОВ /
500 ASSIGN 590 TO IS
GO TO 1500
C
590 RETURN
C IR = 6
C ---- ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ 2-Х ВЫБОРОК ----
C IR = 6; IRP=006,106; IRP=016,026,116,126.
600 IF ( ISR.NE. 0 ) GOTO 605
C ФОРМИРОВАНИЕ ТАБЛИЦЫ МУЛЬТИНОМИНАЛЬНЫХ ЧАСТОТ ПРИ IRP=0XX
C ИЗ ЭЛЕМЕНТОВ ВХОДНЫХ ВЫБОРОК
ASSIGN 603 TO IS
GO TO 1120
C
603 IF(IABS(IER(1)).GE.2 .OR. IABS(IER(2)).GE.2) GO TO 690
C
605 IV = IPR+1
GO TO ( 610, 610, 620 ), IV
C IRP = X06; X16; X26.
C *** КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ (АССОЦИАЦИИ) ЮЛА ***
610 IF ( K.NE. 2 .OR. M.NE. 2 ) GO TO 620
615 ASSIGN 690 TO IS
GO TO 1610
C *** КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ЧУПРОВА-КРАМЕРА ***
620 IF ( K.LT. 3 .AND. M.LT. 3 ) GO TO 615
ASSIGN 690 TO IS
GO TO 1620
C
690 RETURN
C IR = 7
C ---- ВЫЧИСЛЕНИЕ КОЭФФ. КОРРЕЛЯЦИИ МЕЖДУ 2-Я ВЫБОРКАМИ: ----
C ОДНА - В ШКАЛЕ НАИМЕНОВАНИИ, ДРУГАЯ - В ШКАЛЕ ОТНОШЕНИИ
C IR = 7; IRP=7; IRP=17,27.
700 IV = IPR +1
GO TO ( 710, 715, 720 ), IV
C IRP = X07; X17; X27.
C *** БИСЕРИАЛЬНЫЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ***
C В ПРЕДПОЛОЖЕНИИ ДИХОТОМИИ НА ОСНОВЕ
C НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ ДЛЯ "X"
710 ASSIGN 720 TO IS
GO TO 1710
C
715 ASSIGN 790 TO IS
GO TO 1710
C
C IRP = X07, IRP = X27
C *** ТОЧЕЧНЫЙ БИСЕРИАЛЬНЫЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ***
C В ПРЕДПОЛОЖЕНИИ ДИХОТОМИИ ДЛЯ "X"
720 ASSIGN 790 TO IS
GO TO 1720
C
790 RETURN
C
C GO TO 1990

```

```

C
C ----- КОНЕЦ ----- АЛГОРИТМОВ ----- УПРАВЛЕНИЯ -----
C
C ----- ОБРАЩЕНИЯ ----- К ----- ПОДПРОГРАММАМ -----
C
C
C      СТАТИСТИЧЕСКАЯ ОБРАБОТКА ВЫБОРОЧНОЙ СОВОКУПНОСТИ
C      IR=1: IRP=1,11. IR=3: IRP=3.
C
1110  KL(1)=KEY(1,1,1)
      KL(2)=KEY(2,1,1)
C
      CALL SP110(KL,NFWR,KSQN(1,NL),X(IX),N(NL),K,C,MD,Q,P(IP)
      * ,Y(NY), IND(NL), IER(NL), IERW(INL) )
CL; SP110 - ВЫЧИСЛЕНИЕ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК ВЫБОРКИ
      GO TO (9999,9999,1112), IR
1112  IX = IX + N(NL)
      IP = IP + K
      NY = NY + 8
      NL = NL + 1
      INL = INL + 8
      GO TO IS, ( 310 )
C
C *** СТАТИСТ. ОБРАБОТКА 2-Х ВЫБОРОК, ФОРМИРОВАНИЕ МАТРИЦЫ ***
C      IR = 1: IRP=21. IR = 6: IRP = 006,016,026.
C      IR = 2: IRP=2,12,22,32,42; IRP=202,212,222,232,242
1120  KL(1)=KEY(1,2,1)
      KL(2)=KEY(2,2,1)
      NSUM = N(1)+N(2)
      LS = K + M
C
      CALL SP120(KL,NFWR,KSQN,X,NSUM,N,LS,K,M,C,MD,Q
      * ,T,P(IP),Y(NY),IND(NL),IER(NL),IERW(INL))
CL; SP120 - ОБРАБОТКА 2-Х ВЫБОРОК, ФОРМИРОВАНИЕ МАТРИЦЫ ЧАСТОТ
      IP = IP + LS
      NY = NY + 16
      NL = NL + 2
      INL = INL + 16
      GO TO IS, ( 203, 603, 9999)
C
C ***Т-КРИТЕРИЙ *** ОЦЕНКА РАЗЛИЧИЯ 2-Х ВЫБОРОК
C      IR = 2: IRP = 2, 12
1210  KL(1)=KEY(1,1,2)
      KL(2)=KEY(2,1,2)
C
      CALL SP210(KL,NFWR,KSQN,K,N,P(1),P(K+1),Q
      * ,Y(NY),Y(NY+K), IND(NL), IER(NL), IERW(INL))
CL; SP210 - Т-КРИТЕРИЙ СТУДЕНТА ОЦЕНКИ РАЗЛИЧИЯ 2-Х НЕЗ. ВЫБ.
C
1215  NY = NY + K + 1
      NL = NL + 1
      INL = INL + 8
      GO TO IS, ( 220 )
C
C ****ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ*** - ОЦЕНКА РАЗЛИЧИЯ 2-Х ВЫБОРОК
C      IR = 2: IRP = 2, 22
1220  KL(1)=KEY(1,2,2)
      KL(2)=KEY(2,2,2)
C
      CALL SP220(KL,NFWR,KSQN,K,N,P(1),P(K+1),Q
      * ,Y(NY),Y(NY+1),J(NJ),Y(NY+2),Y(NY+3),IND(NL),IER(NL)
      * ,IERW(INL))
CL; SP220 - ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ РАЗЛИЧИЯ 2-Х НЕЗ. ВЫБ.
C
1225  NY = NY + 4
      NJ = NJ + K

```



```

      NL = NL + 1
      INL = INL + 8
      GO TO IS, ( 230 )

C
C *** КРИТЕРИЙ ХИ-КВАДРАТ *** НЕПАРАМЕТРИЧЕСКИЙ ***
C ОЦЕНКА РАЗЛИЧИЯ 2-Х НЕЗАВИСИМЫХ ВЫБОРОК
C В ТАБЛИЦЕ СОПРЯЖЕННОСТИ R*C /NS*NT/
C IR = 2; IRP = 2, 32
C
1230 KL(1)=KEY(1,3,2)
      KL(2)=KEY(2,3,2)
C
      CALL SP230(KL,NFWR,KSQN, 2, K, M, T, Q
      *,Y(NY+4),Y(NY+4+K),Y(NY),Y(NY+1),IND(NL),IER(NL),IERW(INL))
CL SP230 - КРИТЕРИЙ ХИ-КВАДРАТ ОЦЕНКИ РАЗЛИЧИЯ 2-Х НЕЗ. ВЫБ.
C
1235 J0 = NY + 2
      NY = NY + 4 + K + M
      NL = NL + 1
      INL = INL + 8
      GO TO IS, ( 240 )
C
C *** КРИТЕРИЙ Н12 *** - ОЦЕНКА РАЗЛИЧИЯ
C 2-Х НЕЗАВИСИМЫХ ВЫБОРОК ПРИ ЧИСЛЕ ГРАДАЦИИ > 3
C IR = 2; IRP = 2, 42
C ПРОВЕРКА НА ДОПУСТИМОСТЬ ПРИМЕНЕНИЯ Н12-КРИТЕРИЯ
C
1240 IF (K.LT.4) GO TO 1290
      KL(1)=KEY(1,4,2)
      KL(2)=KEY(2,4,2)
C
      CALL SP240(KL,NFWR,KSQN,K,N,P(1),P(K+1),Q
      *,Y(J0),Y(J0+1),IND(NL),IER(NL),IERW(INL))
CL SP240 - КРИТЕРИЙ ХИ-КВАДРАТ ОЦЕНКИ РАЗЛИЧИЯ 2-Х
C; НЕЗАВИСИМЫХ ВЫБОРОК ПРИ ЧИСЛЕ ГРАДАЦИИ > 3
C
      NL = NL + 1
      INL = INL + 8
1290 CONTINUE
C
      GO TO IS, ( 290 )
C
C ----- ОЦЕНКА РАЗЛИЧИЯ "М" НЕЗАВИСИМЫХ ВЫБОРОК -----
C IR = 3; IRP=3
C 1300 CONTINUE
C
C *** ДИСПЕРСИОННЫЙ ОДНОФАКТОРНЫЙ АНАЛИЗ ***
C IR = 3; IRP=3
C ПРОВЕРКА НА ДОПУСТИМОСТЬ ПРИМЕНЕНИЯ ДИСПЕРСИОННОГО АНАЛИЗА
      IF ( NS .LT. 3 ) GO TO 1390
      KL(1)=KEY(1,7,2)
      KL(2)=KEY(2,7,2)
C
C
      CALL SP270(KL,NFWR,KSQN,M,NS,N,Y(J0),KV,Q
      *,Y(KY+1),Y(KY+2),IND(NL)
      *,Y(J1),Y(J2),Y(J3),J(J4),J(J5)
      *,IER(NL),IERW(INL) )
CL SP270 - ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ "М" НЕЗ. ВЫБОРОК
C
1390 CONTINUE
C
      NL = NL + 1
      INL = INL + 8
      GO TO IS, ( 390 )
C
C ----- ОЦЕНКА РАЗЛИЧИЯ МЕЖДУ СВЯЗАННЫМИ ВЫБОРКАМИ -----

```

```

C           ИЗМЕРЕННЫМИ В ШКАЛЕ НАИМЕНОВАНИЯ /ОПЫТ "ДО-ПОСЛЕ"/
C
C 1400 CONTINUE
C
C *** КРИТЕРИИ МАКНИМАРА ***
C   для 2-Х СВЯЗАННЫХ ВЫБОРОК, ЯВЛЯЮЩИХСЯ РЕЗУЛЬТАТОМ
C   ОПЫТА "ДО-ПОСЛЕ" ПРИ ДВУХВАРИАНТНОМ ХАРАКТЕРЕ ОТВЕТА
C   IR = 4: IRP=4
C       KL(1)=KEY(1,1,3)
C       KL(2)=KEY(2,1,3)
C
C       CALL SP312(KL, NFWR, KSON, NS, T, Q
C   *           ,Y(1), Y(2), IND(NL), IER(NL), IERW(INL))
C CL; SP312 - КРИТЕРИИ МАКНИМАРА ОЦЕНКИ РАЗЛИЧИЯ 2-Х СВЯЗ. ВЫБОРОК
C
C       NY = NY + 2
C       NL = NL + 1
C       INL =INL + 8
C       GO TO IS, ( 490 )
C
C 1500 CONTINUE
C
C *** КРИТЕРИИ КОКРЕНА **
C   для "М" СВЯЗАННЫХ ВЫБОРОК, ЯВЛЯЮЩИХСЯ РЕЗУЛЬТАТОМ
C   ОПЫТА "ДО-ПОСЛЕ" ПРИ "М" УСЛОВИЯХ НАД "N" ОБЪЕКТАМИ
C   / ХАРАКТЕР ОТВЕТОВ ДВУХВАРИАНТНЫЙ /
C   IR = 5: IRP=5
C       KL(1)=KEY(1,7,3)
C       KL(2)=KEY(2,7,3)
C
C       CALL SP370 (KL, NFWR, KSON, NS, K, M, T, Q
C   *           ,Y(3),Y(M+3),Y(1),Y(2),IND(NL),IER(NL),IERW(INL))
C CL; SP370 - КРИТЕРИИ КОКРЕНА ОЦЕНКИ РАЗЛИЧИЯ "М" СВЯЗ. ВЫБОРОК
C
C       NY = NY + 2 + K + M
C       NL = NL + 1
C       INL =INL + 8
C       GO TO IS, ( 590 )
C       IR = 6
C ----- ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ -----
C   для 2-Х ВЫБОРОК, ИЗМЕРЕННЫХ В ШКАЛЕ НАИМЕНОВАНИЙ
C
C 1600 CONTINUE
C
C *** КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ (АССОЦИАЦИИ) ЮЛА ***
C   IR = 6: IRP=6,16
C 1610   KL(1)=KEY(1,1,4)
C       KL(2)=KEY(2,1,4)
C       NSUM = N(1)+N(2)
C
C       CALL SP412(KL,NFWR,KSON,NS, T, NSUM, Q
C   *           ,Y(NY),Y(NY+1),Y(NY+2),Y(NY+3),IND(NL),IER(NL),IERW(INL))
C CL; SP412 - КОЭФФИЦИЕНТ АССОЦИАЦИИ ( КОРРЕЛЯЦИИ ) ЮЛА
C
C       NY = NY + 4
C       NL = NL + 1
C       INL =INL + 8
C       GO TO IS, ( 690 )
C
C *** КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ЧУПРОВА/КРАМЕРА ***
C   IR = 6: IRP=6,26
C 1620   KL(1)=KEY(1,2,4)
C       KL(2)=KEY(2,2,4)
C
C       CALL SP420(KL,NFWR,KSON,NS,K,M,T,Q
C   *           ,Y(NY+3),Y(NY+3+M),Y(NY),Y(NY+1),Y(NY+2) ,

```

```

      *          ,IND(NL),IER(NL),IERW(INL))
CL; SP420 - КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ЧУПРОВА-КРАНЕРА
C
      NY = NY + 3 + K + M
      NL = NL + 1
      INL = INL + 8
      GO TO IS, ( 690 )
C      IR = 7
C ----- ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ -----
C      МЕЖДУ 2-Я ВЫБОРКАМИ :
C      ОДНА - В ШКАЛЕ НАИМЕНОВАНИЙ, ДРУГАЯ - В ШКАЛЕ ОТНОШЕНИЙ.
C
C1700 CONTINUE
C
C *** БИСЕРИАЛЬНЫЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ***
C      В ПРЕДПОЛОЖЕНИИ ДИХОТОМИИ НА ОСНОВЕ
C      НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ ДЛЯ "X"
C      IR = 7; IRP=7
C1710 KL(1)=KEY(1,7,4)
C      KL(2)=KEY(2,7,4)
C
C      CALL SP472(KL,NFWR, KSQN, N(1), Z, X
C      *          , Y(1), Y(2), Y(3), IND(NL), IER(NL), IERW(INL))
CL; SP472 - БИСЕРИАЛЬНЫЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ
C
      NL = NL + 1
      INL = INL + 8
      GO TO IS, ( 720, 790 )
C
C1750 CONTINUE
C
C *** ТОЧЕЧНЫЙ БИСЕРИАЛЬНЫЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ***
C      В ПРЕДПОЛОЖЕНИИ ДИХОТОМИИ ДЛЯ "X"
C
C1720 KL(1)=KEY(1,8,4)
C      KL(2)=KEY(2,8,4)
C
C      CALL SP482(KL,NFWR, KSQN, N(1), Z, X, Q
C      *          , Y(4), IND(NL), IER(NL), IERW(INL) )
CL; SP482 - ТОЧЕЧНЫЙ БИСЕРИАЛЬНЫЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ
C
      NL = NL + 1
      INL = INL + 8
      GO TO IS, ( 790 )
C
C      DATA : НОМЕР :                      КОРРЕКЦИЯ 2 ИЮЛЯ 1987 ГОДА
C      9999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
      END

```

§ 1. Статистическая обработка выборки

1.1. Статистические характеристики выборки. Проверка нуль-гипотезы H_0 о равенстве долей. Статистическая обработка выборки заключается в вычислении статистических характеристик выборки, построении диаграмм распределения долей градаций качественного признака и проверке нуль-гипотезы H_0 о равенстве долей (частот) в генеральной совокупности по критерию χ^2 .

Для статистического описания одной выборки анализируемого признака (параметра) допускаются следующие статистические характеристики:

1) n_i — выборочная (эмпирическая) частота, т. е. число единиц, попавших в i -ю градацию;

2) p_i — выборочная доля i -й градации (N — объем всей выборки)

$$p_i = n_i/N,$$

3) S — показатель изменчивости, характеризующий варьирование данных относительно друг друга

$$S = \sqrt[k]{p_1 p_2 \dots p_k}.$$

где k — количество градаций качественного признака;

4) S_p — стандартная ошибка доли

$$S_p = S/\sqrt{N};$$

5) V — коэффициент варьирования

$$V = S/S_M,$$

где S_M — максимальное значение S для заданного числа градаций

$$S_M = 1/k;$$

6) M_0 — мода, выражающая значение максимальной доли,

$$M_0 = \max(p_1, p_2, \dots, p_k).$$

Осуществляется проверка нуль-гипотезы H_0 о равенстве долей в генеральной совокупности по критерию χ^2 , т. е. вычисляется статистическая величина

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - n'_i)^2}{n'_i},$$

где n_i — частота i -й градации; n_i — ожидаемая (теоретическая на основе H_0) частота в i -й градации,

$$n'_i = \frac{1}{k} \sum_{i=1}^k n_i.$$

Вычисляется критическое значение χ^2_{qv} распределения χ^2 при числе степеней свободы $v = k - 1$ и заданном уровне значимости риска q (обычно $q = 0,05$ или $q = 0,1$, т. е. 5 или 10 %);

$H_0 : p_1 = p_2 = \dots = p_k$ не отвергается, если $\chi^2 < \chi^2_{qv}$; отвергается при $\chi^2 \geq \chi^2_{qv}$.

1.2. Программная реализация решения задачи статистической обработки одной выборки. Блок-схема программной реализации задачи статистической обработки выборки (вычисления статистических характеристик выборки) представлена на рис. 51.

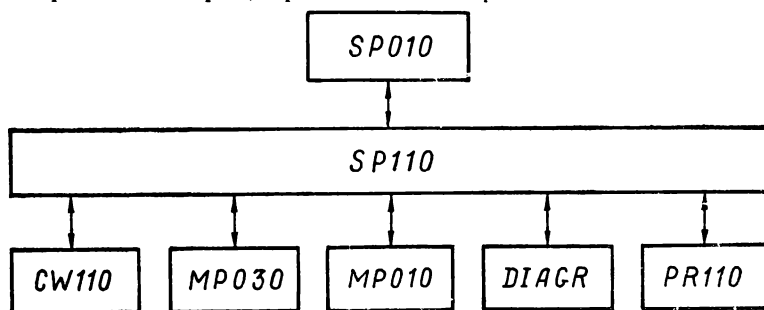


Рис. 51

Обращение к подпрограмме вычисления статистических характеристик выборки *SP11θ* производится из подпрограммы *SPθ1θ*.

Назначение остальных подпрограмм следующее:

CW11θ — контроль и печать входных данных;

MPθ3θ — вычисление критических значений распределения — квадрат; эта программа основана на подпрограмме *PPH12*;

MPθ1θ — вычисление критических значений *t*-распределения Стьюдента;

DIAGR — печать диаграммы распределения долей в выборке, результата проверки гипотезы о равенстве долей, статистических характеристик выборки;

PR110 — печать результатов работы *SP11θ*.

Тексты подпрограмм, необходимых для статистической обработки одной выборки *SP11θ*, *CW11θ*, *PR11θ* приведены ниже, после примера.

1.3. Пример.

Приведем статистическую обработку модельной выборки объемом 200 элементов при девяти градациях качественного признака. Элементы выборки даны в виде вектора числовых кодов.

| № п/п | Код | N | Код | N | Код | N | Код | N | Код | N | Код | N | Код | N | Код |
|----------|-----|----|-----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1 | 1 | 26 | 2 | 51 | 3 | 76 | 4 | 101 | 5 | 126 | 6 | 151 | 7 | 176 | 8 |
| 2 | 1 | 27 | 2 | 52 | 3 | 77 | 4 | 102 | 5 | 127 | 6 | 2 | 7 | 7 | 8 |
| 3 | 1 | 28 | 2 | 53 | 3 | 78 | 4 | 103 | 5 | 128 | 6 | 3 | 7 | 8 | 8 |
| 4 | 1 | 29 | 2 | 54 | 3 | 79 | 4 | 104 | 5 | 129 | 6 | 4 | 7 | 9 | 8 |
| 5 | 1 | 30 | 2 | 55 | 3 | 80 | 4 | 105 | 5 | 130 | 6 | 5 | 7 | 180 | 8 |
| 6 | 1 | 31 | 2 | 56 | 3 | 81 | 4 | 106 | 5 | 131 | 6 | 6 | 7 | 1 | 8 |
| 7 | 1 | 32 | 2 | 57 | 3 | 82 | 4 | 107 | 5 | 132 | 6 | 7 | 7 | 2 | 8 |
| 8 | 1 | 33 | 2 | 58 | 3 | 83 | 4 | 108 | 5 | 133 | 6 | 8 | 7 | 3 | 8 |
| 9 | 1 | 34 | 2 | 59 | 4 | 84 | 4 | 109 | 5 | 134 | 6 | 9 | 7 | 4 | 8 |
| 10 | 1 | 35 | 2 | 60 | 4 | 85 | 4 | 110 | 6 | 135 | 6 | 160 | 7 | 5 | 8 |
| 11 | 1 | 36 | 2 | 61 | 4 | 86 | 4 | 111 | 6 | 136 | 6 | 1 | 7 | 6 | 8 |
| 12 | 1 | 37 | 2 | 62 | 4 | 87 | 4 | 112 | 6 | 137 | 6 | 2 | 7 | 7 | 8 |
| 13 | 1 | 38 | 2 | 63 | 4 | 88 | 4 | 113 | 6 | 138 | 6 | 3 | 8 | 8 | 8 |
| 14 | 1 | 39 | 2 | 64 | 4 | 89 | 4 | 114 | 6 | 139 | 6 | 4 | 8 | 9 | 8 |
| 15 | 2 | 40 | 2 | 65 | 4 | 90 | 5 | 115 | 6 | 140 | 6 | 5 | 8 | 190 | 9 |
| 16 | 2 | 41 | 3 | 66 | 4 | 91 | 5 | 116 | 6 | 141 | 6 | 6 | 8 | 1 | 9 |
| 17 | 2 | 42 | 3 | 67 | 4 | 92 | 5 | 117 | 6 | 142 | 6 | 7 | 8 | 2 | 9 |
| 18 | 2 | 43 | 3 | 68 | 4 | 93 | 5 | 118 | 6 | 143 | 6 | 8 | 8 | 3 | 9 |
| 19 | 2 | 44 | 3 | 69 | 4 | 94 | 5 | 119 | 6 | 144 | 6 | 9 | 8 | 4 | 9 |
| 20 | 2 | 45 | 3 | 70 | 4 | 95 | 5 | 120 | 6 | 145 | 6 | 170 | 8 | 5 | 9 |
| 21 | 2 | 46 | 3 | 71 | 4 | 96 | 5 | 121 | 6 | 146 | 6 | 1 | 8 | 6 | 9 |
| 22 | 2 | 47 | 3 | 72 | 4 | 97 | 5 | 122 | 6 | 147 | 7 | 2 | 8 | 7 | 9 |
| 23 | 2 | 48 | 3 | 73 | 4 | 98 | 5 | 123 | 6 | 148 | 7 | 3 | 8 | 8 | 9 |
| 24 | 2 | 49 | 3 | 74 | 4 | 99 | 5 | 124 | 6 | 149 | 7 | 4 | 8 | 9 | 9 |
| 25 | 2 | 50 | 3 | 75 | 4 | 100 | 5 | 125 | 6 | 150 | 7 | 5 | 8 | 200 | 9 |

Массив уровней градации и соответствующие им числовые коды задаются следующим образом:

| Градация | Уровень | Числовые коды |
|----------|---------|------------------|
| 1 | 1 | код ≤ 1 |
| 2 | 2 | 1 < код ≤ 2 |
| 3 | 3 | 2 < код ≤ 3 |
| 4 | 4 | 3 < код ≤ 4 |
| 5 | 5 | 4 < код ≤ 5 |
| 6 | 6 | 5 < код ≤ 6 |
| 7 | 7 | 6 < код ≤ 7 |
| 8 | 8 | 7 < код ≤ 8 |
| 9 | 9 | 8 < код ≤ 9 |

При таком способе задания правила, по которому элемент со своим кодом попадает в соответствующую ему градацию, легко можно уменьшить количество градаций шкалы, если в процессе измерения окажется, что в какой-то градации нет элементов. Так, например, задав массив уровней градаций

1, 2, 5, 6, 8, 9

и количество градаций, равное шести, получаем следующее правило соответствия между числовыми кодами результатов измерений и элементами номинальной шкалы:

| Градация | Уровень | Числовые коды |
|----------|---------|----------------------------------|
| 1 | 1 | код ≤ 1 , т. е. 1 |
| 2 | 2 | 1 < код ≤ 2 , т. е. 2 |
| 3 | 5 | 2 < код ≤ 5 , т. е. 3, 4, 5 |
| 4 | 6 | 5 < код ≤ 6 , т. е. 6 |
| 5 | 8 | 6 < код ≤ 8 , т. е. 7, 8 |
| 6 | 9 | 8 < код ≤ 9 , т. е. 9 |

При подготовке исходного файла для обработки комплексом программ «СОКИ-Н» необходимо задать дополнительные параметры:

IRP-011 — код функции обработки одной выборки;

Q-0.05 — уровень значимости риска (равный 5 %);

MD-40 — максимальный размер (в позициях) диаграммы.

ТЕСТОВОЕ ЗАДАНИЕ НА ВЫЧИСЛЕНИЕ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК
ВЫБОРОЧНОЙ СОВОКУПНОСТИ

```

---- СТАТИСТИЧЕСКАЯ ОБРАБОТКА ОДНОЙ ВЫБОРКИ :
****   МАКЕТ ПОДГОТОВКИ ДАННЫХ ДЛЯ IR= 1,11:   ****SP110****
IRP -011                                     (I3)- КОД ВЫПОЛНЯЕМОЙ ФУНКЦИИ
NS  -01 /ПО УМОЛЧАНИЮ/ (I2)- КОЛИЧЕСТВО ВЫБОРОК
K   -09                                     (I2)- КОЛИЧЕСТВО ГРАДАЦИЙ
M   -01                                     (I2)-
      (I2)-НОМЕР; (4A8)- К О Д  ВЫБОРКИ;      (I4)-ДЛИНА
      XX  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  XXXX
KSON-01  ТЕСТОВАЯ /9 ГРАДАЦИЙ/ ВЫБОРКА 1 0200
      (I2, I3,I2) [ 4A8 ]:НОМЕР ВЫБ.,КОД,КАНАЛ-ИМЯ ФАЙЛА
FILX-01 000 08-INF:DATAXS.DAT
      ( СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) )
FORX-01 (4X,30I2)
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
      2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 4 4
      4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 5
      5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 6
      6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
      7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
      8 8 8 8 8 8 8 8 8 8 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9
      (I2, I3,I2) [ 4A8 ]:НОМЕР ВЫБ.,КОД,КАНАЛ;ИМЯ ФАЙЛА
FILC-01 000 08-INF:DATAXS.DAT
      ( СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) )
FORC-01 (4X,9I2)
      1 2 3 4 5 6 7 8 9
Q   -0.05 - (G14.7)- УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
NFWR-02 (I2) - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
      (I1,I2)-ТИП КЛЮЧЕЙ ПЕЧАТИ: 0-ПОРАЗРЯДНЫЙ; И ИХ КОЛИЧЕСТВО
KEYC-0 02 1-ИНТЕРВАЛЬНЫЙ.
      ИНДЕКСЫ/ИНТЕРВАЛЬНОЕ ЗНАЧЕНИЕ КЛЮЧА ПЕЧАТИ/ПОРАЗРЯДНОЕ/
      12345678 - РАЗРЯДЫ КЛЮЧА ПЕЧАТИ :
KEY - 1 1 1/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SP110 :
KEY - 2 1 1/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP110 .
MD  -40 (I2)- МАКСИМАЛЬНАЯ ВЫСОТА ДИАГРАММЫ
      (I1)-РАСПРЕДЕЛЕНИЕ ПАМЯТИ:=1-ВЫДАЕТСЯ ВСЕГДА;
END -1 КОНЕЦ ДАННЫХ; =0-ВЫДАЕТСЯ ТОЛЬКО АВАРИЙНО.

```

РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ ЗАДАНИЯ (РАСПЕЧАТКА РЕШЕНИЯ):

```

IRP - 11
K - 9
M - 1
KSGN NN XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX NNNN
KSGN- 1 ТЕСТОВАЯ /9 ГРАДАЦИЙ/ ВЫБОРКА 1 200
FILX- 1 0 8-INF:DATAXC.DAT
FORX- 1: (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 4 4
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 5 5
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 6 6
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 7 7
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 8 8
8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 9 9
FILC- 1 0 8-INF:DATAXC.DAT
FORC- 1: (4X,9I2)
1 2 3 4 5 6 7 8 9
Q 0.5000000E-01
NFWR 2
KEYC 0 2 2
KEY 1 1 1/8/11111111/
KEY 2 1 1/8/11111111/
MD 40
ОБЪЕМ РАБОЧЕЙ ПАМЯТИ - 4000 ЕДИНИЦ INTEGER
ТРЕБУЕМЫЙ ОБЪЕМ ПАМЯТИ - 347 ЕДИНИЦ INTEGER

```

ИНДЕКСЫ НАЧАЛА ПЕРЕМЕННЫХ В РАБОЧЕЙ ПАМЯТИ :

| | | | | | | | |
|-----|---|------|----|------|-----|------|-----|
| 1 - | 1 | 6 - | 6 | 11 - | 90 | 16 - | 152 |
| 2 - | 2 | 7 - | 7 | 12 - | 290 | 17 - | 161 |
| 3 - | 3 | 8 - | 36 | 13 - | 299 | 18 - | 338 |
| 4 - | 4 | 9 - | 19 | 14 - | 150 | 19 - | 339 |
| 5 - | 5 | 10 - | 89 | 15 - | 151 | 20 - | 340 |

SW110 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА Q = 0.050

КОД ВЫБОРКИ :

ТЕСТОВАЯ /9 ГРАДАЦИЙ/ ВЫБОРКА 1

КОЛИЧЕСТВО ГРАДАЦИЙ - K = 9

ЗНАЧЕНИЯ ЭЛ. ВЫБОРОЧНОЙ СОВОКУПНОСТИ

ОБЪЕМ :
200

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 |
| 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 5 |
| 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 6 |
| 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| 6 | 6 | 6 | 6 | 6 | 6 | 7 | 7 | 7 | 7 |
| 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| 7 | 7 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 9 |
| 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |

DIAGR K=9

[illegible]

К О Д В Ы Б О Р К И : Т Е С Т О В А Я / 9 Г Р А Д А Ц И Я / В Ы Б О Р К А 1

| | | |
|--|---|---------|
| ОБЪЕМ ВЫБОРКИ | - | 200 |
| ЧИСЛО ГРАДАЦИЙ | - | 9 |
| ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ | - | 0.10377 |
| КОЭФФИЦИЕНТ ВАРИИРОВАНИЯ | - | 0.93391 |
| КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ | - | 0.00734 |
| ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ (D) | - | 0.01453 |
| МОДА | - | 0.18500 |
| КРИТЕРИЙ ХИ-КВАДРАТ К СПРАВЕДЛИВОЙ ГИПОТЕЗЕ H1 | | |


```

CC; SP110 - ОБРАБОТКА ОДНОЙ ВЫБОРКИ ИЗМЕР. В НОМИНАЛЬНОЙ ШКАЛЕ
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫЧИСЛЕНИЯ СТАТИСТИЧЕСКИХ
C ХАРАКТЕРИСТИК ОДНОЙ ВЫБОРОЧНОЙ СОВОКУПНОСТИ
C1 (НОМИНАЛЬНАЯ ШКАЛА)
C
CR ЛИТЕРАТУРА ПО АЛГОРИТМУ : Г.ШВАРЦ, ВЫБОРОЧНЫЙ МЕТОД,
C1 М., СТАТИСТИКА, 1978.
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММО,
C1 Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE SP110(KEY,NFWR,KSEQ,X,N,K,C,MD,Q,P,
* Y,IND,IER,IERW)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
CР ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C
C KEY I*4 * * (8) МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 * * 1 НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<NFWR<=99/
C KSEQ R*8 * * (4) КОД ВЫБОРКИ
C X I*4 * * (N) МАССИВ ЗНАЧЕНИЙ ВЫБОРОЧНОЙ СОВОКУПНОСТИ
C / НЕ РАВНЫХ ВСЕ ТОЖДЕСТВЕННО КОНСТАНТЕ /
C N I*4 * * 1 ДЛИНА МАССИВА 'X' /N>=2/
C K I*4 * * 1 КОЛИЧЕСТВО ГРАДАЦИЙ(ДЛИНА 'C', 'P')/K>=2/
C C I*4 * * (K) МАССИВ ЗНАЧЕНИЙ УРОВНЕЙ (ДЛЯ ГРАДАЦИЙ)
C / НЕ РАВНЫХ ВСЕ ТОЖДЕСТВЕННО КОНСТАНТЕ /
C MD I*4 * * 1 МАКСИМАЛЬНАЯ ВЫСОТА ДИАГРАММЫ, ПОЗИЦИИ
C Q R*4 * * 1 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА /.01<Q<=.5/
C P R*4 * * (K) МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ /СУММА =1.0+-EPS/
C Y R*4 * * (8) МАССИВ ВЫЧИСЛЕННЫХ СТАТ. ХАРАКТЕРИСТИК
C Y(1)=S - ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ;
C Y(2)=V - КОЭФФИЦИЕНТ ВАРИИРОВАНИЯ;
C Y(3)=SP - КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ;
C Y(4)=ZMD - МОДА;
C Y(5)=HI2 - ЗНАЧЕНИЕ КРИТЕРИЯ ХИ-КВАДРАТ;
C Y(6)=TH1 - КРИТИЧЕСКОЕ ЗНАЧЕНИЕ ХИ-КВАДРАТ;
C Y(7)=0 - РЕЗЕРВИРОВАНО;
C Y(8)=D - ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ
C IND I*4 * * 1 ИНДЕКС ПРОВЕРКИ ПО ХИ-КВАДРАТ ;
C IND=0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
C IND=1 - НУЛЕВАЯ ГИПОТЕЗА ОТВЕРГАЕТСЯ.
C IER I*4 * * 1 ВЫРАБОТАННЫЙ КОД ОШИБКИ
C IERW I*4 * * (8) КОД ОШИБКИ, ПРИНЯТЫЙ ИЗ CW110
C
C DIMENSION KEY(8),X(N),C(K),P(K),Y(8),IERW(8),KSEQ(4)
C INTEGER X,C
C REAL*8 NAME,KSEQ
C DATA NAME /' SP110 '/,ZERO/0.0/,ONE/1.0/, EPS/0.001/
C IND = 11
C IER = 0
C
C ЕСЛИ KEY(1) > 0, ПРОВЕРКА И ПЕЧАТЬ ИСХОДНЫХ ДАННЫХ
C KLP=KEY(2)
C IF (KEY(1).EQ.0) GO TO 120
C CALL CW110(KEY,NFWR,KSEQ,X,N,K,C,Q,IERW,MWC)
C CW110 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ;
C
C АНАЛИЗ КОДОВ ОШИБОК "KF" И "MWC"
C KF = IERW(1)
C IF (MWC.LT.3) GO TO 120
C ЕСЛИ MWC >= 3, TO IER=MWC

```

```

IER = MWC
  IF (MWC.GE.3) GO TO 999
120 DO 200 I=1,K
200 P(I) = ZERO
  L = 1
  DO 600 I=1,N
    J = X(I)
    IF (K - 2) 999, 250, 300
250 L = 1
    IF ( J .GT. C(1) ) L = 2
    GO TO 500
300 IF ( J - C(L) ) 350,500,400
350 M = L-1
    IF ( M .EQ. 0 ) GO TO 500
    IF ( J .GT. C(M) ) GO TO 500
    L = M
    GO TO 350
400 IF ( L .GE. K ) GO TO 600
    L = L + 1
    IF ( J .LE. C(L) ) GO TO 500
    GO TO 400
500 P(L)=P(L)+ONE
600 CONTINUE
    D=FLOAT(N)
    S=ONE
    M = K
    ZMD=ZERO
    DO 700 I=1,K
      IF (P(I).EQ.ZERO) M=M-1
      P(I)=P(I)/D
      IF (P(I).GT.ZMD) ZMD=P(I)
700 S = S*P(I)
    IF (M .EQ. K) GO TO 710
    IER=2
    GO TO 810
710 SMAX=ONE/FLOAT(K)
    S=S**SMAX
    V=S/SMAX
    SP=S/SQRT(D)
    ZMN=ZERO
C  ПРОВЕРКА ГИПОТЕЗЫ О РАВЕНСТВЕ ЗНАЧЕНИЙ ЧАСТОТ ( ИЛИ ДОЛЕЙ )
C  ПО КРИТЕРИЮ ХИ-КВАДРАТ
    D = ZERO
    DO 750 I=1,K
750 D=D+(P(I)-SMAX)**2
    HI2=D/SMAX*FLOAT(N)
    Y(1)=S
    Y(2)=V
    Y(3)=SP
    Y(4)=ZMD
    Y(5)=HI2
    Y(7)=ZMN
C
C  ЕСЛИ БЫЛИ ОШИБКИ, ТО IER = 2
C
    CALL MP030( Q,K-1,EPS,THI,IERTHI)
CL; MP030 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ ХИ-КВАДРАТ;
    IF (IERTHI.EQ.0) GO TO 760
    IER = 4+IERTHI
    THI = ZERO
    GO TO 770
760 IND = 0
    IF (HI2.GE.THI) IND = 1
770 Y(6)=THI
    CALL MP010(Q,N-1,EPS,TGV,IERT)
CL; MP010 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ Т-РАСПРЕДЕЛЕНИЯ;

```

```

        IF (IERT.EQ.0) GO TO 780
        IER = 4+IERT
        Y(8) = ZERO
        GO TO 810
780 Y(8)=SP*TQV
C
810 IF (KF .NE. 3) GO TO 830
    IER = -IER
    GO TO 999
830 IF (KLP.EQ.0) GOTO 999
    KEY(1)=KLP
    CALL DIAGR (KEY,NFWR,KSEQ,0,N,K,C,P,Y,MD,IND,IER)
CL; DIAGR - ПЕЧАТЬ ДИАГРАММЫ РАСПРЕДЕЛЕНИЯ ДОЛЖЕ
    KEY(1)=KLP
    CALL PR110(KEY,NFWR,KSEQ,K,P,Y,IND,IER)
CL; PR110 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP110
999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
END
02-FEB-88

```

```

CS; SW110 - КОНТРОЛЬ И/ИЛИ ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP110
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ КОНТРОЛЯ ВХОДНЫХ ДАННЫХ ДЛЯ SP110
C И ВЫДАЧИ В ФАЙЛ С НОМЕРОМ 'NFWR' РЕЗУЛЬТАТА ПРОВЕРКИ
C И ВХОДНЫХ ДАННЫХ
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP110
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧМФФ,
C Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE SW110(KEY,NFWR,KSEQ,X,N,K,C,Q,IERW,MNC)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
CР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА/1<=NFWR<=99/
C KSEQ R*8 (4) 4A8 7 К О Д ВЫБОРКИ
C X I*4 (N) 15 8 МАССИВ ЗНАЧЕНИЙ ВЫБОРОЧНОИ
C СОВОКУПНОСТИ / НЕ РАВНЫХ ВСЕ ТОЖДЕСТВЕННО КОНСТАНТЕ /
C N I*4 1 15 7 ДЛИНА МАССИВА 'X' / N >= 2 /
C K I*4 1 15 7 КОЛИЧЕСТВО ГРАДАЦИИ
C (ДЛИНА МАССИВА 'C') / K >= 2 /
C C I*4 (K) 15 8 МАССИВ ЗНАЧЕНИЙ УРОВНЕЙ
C (ДЛЯ ГРАДАЦИИ) / НЕ РАВНЫХ ВСЕ ТОЖДЕСТВЕННО КОНСТАНТЕ /
C Q R*4 1 F8.3 3 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
C .005<=Q<=.5 /
C IERW I*4 (8) 612 [2] ВЫРАБОТАННЫЙ КОД ОШИБКИ : 6 РАЗРЯД
C;MNC I*4 1 .11 2 КОД МАКСИМАЛЬНОЙ ОШИБКИ
C
C DIMENSION KEY(8),X(N),C(K),IERW(8),KSEQ(4)
C INTEGER X,C,T,XMIN,XMAX
C REAL*8 NAME,KSEQ
C DATA NAME / ' SW110 ' /
C
C KLU=KEY(1)
C IF (KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 11 KLU=IABS(KLU)
C DO 13 J=1,8
C I=9-J
C KEY(I)=MOD(KLU,2)
C 13 KLU=KLU/2
C GOTO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 15 DO 17 I=1,8
C KEY(I)=1
C IF (KLU.LT.I) KEY(I)=0
C 17 CONTINUE
C 19 CONTINUE
C
C - ПРОВЕРКА 'NFWR' ( 1-й РАЗРЯД )
C НАЧАЛЬНАЯ УСТАНОВКА : ОШИБОК НЕТ
C KF = 0
C ЕСЛИ 'NFWR' БОЛЕЕ 99 ИЛИ МЕНЬШЕ 1, ЭТО ГРУБАЯ ОШИБКА
C IF (NFWR.GT.99) KF = 3
C IF (NFWR.LT.1) KF = 3
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
C IERW(1) = KF
C
C - ПРОВЕРКА 'N' ( 3-й РАЗРЯД )

```

```

C KN=0 ОЗНАЧАЕТ, ЧТО 'N' - НОРМАЛЬНОЕ, Т.Е. БОЛЬШЕ
  KN = 0
C ЕСЛИ 'N' МЕНЬШЕ ЧЕМ 10, ТО ЭТО ЛОГИЧЕСКАЯ ОШИБКА
  IF (N.LT.10) KN = 1
C ЕСЛИ 'N' МЕНЬШЕ ЧЕМ 2, ЭТО ГРУБАЯ ОШИБКА
  IF (N.LT. 2) KN = 3
C ПО ВЕРХНЕМУ ПРЕДЕЛУ ДЛЯ 'N' ПРОВЕРКИ НЕТ
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
  IERW(3) = KN

C
C-#- ПРОВЕРКА 'K' ( 4-й РАЗРЯД )
C НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'K' : 2 <= K <= 99
  KK = 0
C ЕСЛИ K>=100 - ЭТО ЛОГИЧЕСКАЯ ОШИБКА
  IF (K.GE.100) KK = 1
C НЕПРАВИЛЬНОЕ ЗНАЧЕНИЕ 'K' : K < 2
  IF (K.LT.2) KK=3
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
  IERW(4) = KK

C
C-#- ПРОВЕРКА 'C' ( 5-й РАЗРЯД )
  KC = 3
  IF (KK.GT.1) GO TO 40
C ПРОВЕРКА НЕВОЗМОЖНА, ПЕЧАТЬ НЕЖЕЛАТЕЛЬНА
  KC = 0
C НАЧИНАЕМ ПРОВЕРКУ 'C'
  33 DO 35 IK=2,K
    IF (C(IK-1).GE.C(IK)) GO TO 37
  35 CONTINUE
  GO TO 40
C ГРУБАЯ ОШИБКА В 'C'
  37 KC = 3
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
  40 IERW(5) = KC

C
C-#- ПРОВЕРКА 'X' ( 2-й РАЗРЯД )
  KX = 3
  IF (KN.GT.1) GO TO 50
C ПРОВЕРКА 'X' НЕВОЗМОЖНА, ПЕЧАТЬ НЕЖЕЛАТЕЛЬНА
  KX = 0
C 'N' ПРАВИЛЬНОЕ : ПРОВЕРКА ВОЗМОЖНА, НАЧАЛИ
  41 CONTINUE
  IF (KC.GT.1) GO TO 48
C 'C' ПРАВИЛЬНОЕ : ПОЛНАЯ ПРОВЕРКА
C KIN - СЧЕТЧИК ПО ПРЕДЫДУЩЕМУ ПРОХОДУ ; KI - ПО НОВОМУ
C KIK - СЧЕТЧИК НЕПУСТЫХ ПРОХОДОВ
  KIN = 0
  KIK = 0
  DO 47 IK=1,K
    T = C(IK)
    KI = 0
    DO 43 IC=1,N
      IF (X(IC).GT.T) GO TO 43
      KI = KI + 1
    43 CONTINUE
    IF (KI.GT.KIN) GO TO 45
C ГРУБАЯ ОШИБКА В 'X' : НЕТ ТОЧЕК В ДАННОМ СЛОЕ
  KX = 3
  GO TO 47
C НЕПУСТОЙ СЛОЙ /УРОВЕНЬ/
  45 KIK = KIK + 1
  KIN = KI
  47 CONTINUE
C НАДО ПРОВЕРИТЬ, ОДИН ИЛИ НЕ ОДИН НЕПУСТОЙ СЛОЙ
  IF (KIK.GT.1) GO TO 50
C ПОЛУЧИЛИ ТОЛЬКО ОДИН НЕПУСТОЙ СЛОЙ ПРИ ПОЛНОЙ ПРОВЕРКЕ

```

```

C      ИЛИ НЕЛЬЗЯ СДЕЛАТЬ ПОЛНУЮ ПРОВЕРКУ /ПЛОХОЙ 'С'/
C      НЕОБХОДИМО ПРОВЕРИТЬ ДИАПАЗОН ИЗМЕНЕНИЯ 'X'.
48 XMIN = X(1)
   XMAX = X(1)
   IF (KX.EQ.0) KX = 1
   DO 49 IC=2,N
     T = X(IC)
     IF (T.GT.XMAX) XMAX = T
     IF (T.LT.XMIN) XMIN = T
49   CONTINUE
   IF (XMIN.NE.XMAX) GO TO 50
C      ГРУБАЯ ОШИБКА В 'X' : ВСЕ ЭЛЕМЕНТЫ 'X' РАВНЫ CONST
   KX = 3
C      ФОРМИРОВАНИЕ КОДА ОШИБКИ
50 IERW(2) = KX
C
C*-   ПРОВЕРКА 'Q' ( 6-й РАЗРЯД
C      НОРМАЛЬНОЕ ЗНАЧЕНИЕ Q : 0.025 < Q < 0.2
   Q = 0
C      ЕСЛИ Q БОЛЬШЕ ЧЕМ 0.2 ЛИБО МЕНЬШЕ ЧЕМ 0.025,
C      ТО СЧИТАЕМ ЧТО ДОПУЩЕНА ЛОГИЧЕСКАЯ ОШИБКА
   IF (Q.LT.0.025)
     IF (Q.GT.0.2)
C      ВВОДИМ ОГРАНИЧЕНИЕ : 0.005 < Q < 0.5
     IF (Q.LT.0.005) KQ = 3
     IF (Q.GT.0.5)   KQ = 3
C      ФОРМИРОВАНИЕ КОДА ОШИБКИ
     IERW(6) = KQ
     IERW(8) = 6
C      ОПРЕДЕЛЕНИЕ КОДА МАКСИМАЛЬНОЙ ОШИБКИ
C
   MWC = MAX0 ( KN,KF,KC,KX
C
C      ШКАЛА КОДИРОВАНИЯ С ЛЕВЫХ ПАРАМЕТРОВ :
C
C      РАЗРЯД КОДА ОШИБКИ 1 2 3 4 5 6
C      ПАРАМЕТР В ОБРАЩЕНИИ I NFWR I X I N I K I C I Q
C      ВОЗМОЖНЫЕ ЗНАЧЕНИЯ
C      1 0 I 0 I 0 I 0 I 0 I 0
C      I - I 1 I 1 I 1 I 1 I 1
C      I - I - I - I - I - I -
C      I 3 I 3 I 3 I 3 I 3 I 3
C
   IF (KF.GT.0) GO TO 999
   IF (KEY(2).EQ.0) GO TO 225
C
   WRITE (NFWR,220) NAME, MWC
220 FORMAT (//',AB,4X,'КОД МАКСИМАЛЬНОЙ ОШИБКИ' = ',11)
C
225 IF (MWC.LT.1) GO TO 310
   J = IERW(8)
   WRITE (NFWR,227) (IERW(I),I=1,J)
227 FORMAT (//',8X,'ИМЕНА ПАРАМЕТРОВ :
   *      ',NFWR, X, N, K, C, Q
   *      //',8X,'К О Д Ы ОШИБОК : ',6(3X,12)/)
310 IF (KEY(3).EQ.0) GOTO 710
   WRITE (NFWR,330) Q
330 FORMAT (//',8X,'УРОВЕНЬ ЗНАЧИМОСТИ РИСКА Q = ',F8.3)
C
710 IF (KEY(7).EQ.0) GO TO 810
   WRITE (NFWR,740) KSEQ,N,K
740 FORMAT (//',20X,'К О Д ВЫБОРКИ : ',8X,' ОБЪЕМ :
   *      //',12X,4A8,6X,15
   *      //',8X,'КОЛИЧЕСТВО ГРАДАЦИИ - K = ',15)
}

```



```

810 IF (KEY(8).EQ.0) GO TO 999
IF (KN.GT.1) GO TO 880
WRITE (NFWR,870)
870 FORMAT(' ',8X,'ЗНАЧЕНИЯ ЭЛ. ВЫБОРОЧНОЙ СОВОКУПНОСТИ')
WRITE (NFWR,875) X
875 FORMAT(' ',8X,10I5)
C
880 IF (KK.GT.1) GO TO 999
WRITE (NFWR,890)
890 FORMAT(' ',8X,'МАССИВ ЗНАЧЕНИЙ УРОВНЕЙ:')
WRITE (NFWR,895) C
895 FORMAT(' ',8X,10I5)
C
999 RETURN
CV) ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
END

```

```

CC; PR110 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP110
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫВОДА В ФАЙЛ 'NFWR'
C: РЕЗУЛЬТАТОВ РАБОТЫ П/П SP110
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММО,
C: Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE PR110(KEY,NFWR,KSEQ,K,P,Y,IND,IER)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЯ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА/1<=NFWR<=99/
C KSEQ R*8 (4) 4A8 7 К О Д ВЫБОРКИ
C K I*4 1 -- - КОЛ-ВО ГРАДАЦИЙ (ДЛИНА МАССИВА 'P')
C P R*4 (K) F12.5 8 МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ
C Y R*4 (8) G20.5 3 МАССИВ ВЫЧИСЛЕННЫХ СТАТИСТИЧЕСКИХ
C ХАРАКТЕРИСТИК (СМ. ОПИСАНИЕ SP110)
C IND I*4 1 I1 2 ИНДЕКС ПРОВЕРКИ ПО ХИ-КВАДРАТ:
C IND=0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
C IND=1 - НУЛЕВАЯ ГИПОТЕЗА ОТКЛОНЯЕТСЯ.
C IER I*4 1 I3 2 КОД ОШИБКИ, ВЫРАБОТАННЫЙ SP110
C
DIMENSION KEY(8),P(K),Y(8),KSEQ(4)
REAL*8 TNK(3),TS1(4),TS2(4),TS3(4),NAME,KSEQ
DATA TNK /' STAT.OB',' РАБОТКА',' ВЫБОРКИ' /
*,TS1 /' ',' ВЫЧИС',' ЛЕНИЕ' ,'/
*,TS2 /' СТАТИС',' ТИЧЕСКИХ',' ХАРАКТЕ',' РИСТИК' /
*,TS3 /' ОДНОЙ',' ВЫБОРОЧН',' ОЙ СОВОК',' УПНОСТИ' /
DATA NAME /' PR110' /
C
IF (NFWR.LE.0.OR.NFWR.GE.100) GO TO 999
C
KLU=KEY(1)
IF (KLU) 11,999,15
C - ПОРЯЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
13 KLU=KLU/2
GO TO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF (KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
C
IF (KEY(1).EQ.0) GO TO 210
WRITE (NFWR,120) TS1,TNK,TS2,TS3
120 FORMAT(/ 1H ,26(1H*)
* /2H *,24(1H ),3H* ,4A8
* /2H *,3A8 ,3H* ,4A8
* /2H *,24(1H ),3H* ,4A8
* / 1H ,26(1H*) )
210 IF (KEY(2).EQ.0) GO TO 310
C
WRITE (NFWR,230) NAME, IND, IER
230 FORMAT(/' ',A8,' - СПРАВЕДЛИВА ГИПОТЕЗА Н',I1
* ,'; КОД ОШИБКИ =' ,I3)

```

```

C 310 IF (KEY(3).EQ.0) GOTO 710
    WRITE (NFWR,340) Y(1),Y(2),Y(3),Y(4),Y(5),Y(6)
340 FORMAT(/' ',8X,'ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ      ='',620.5
*         //' ',8X,'КОЭФИЦИЕНТ ВАРЬИРОВАНИЯ      ='',620.5
*         //' ',8X,'КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ    ='',620.5
*         //' ',8X,'МОДА                            ='',620.5
*         //' ',8X,'ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ HI2       ='',620.5
*         //' ',8X,'КРИТИЧЕСКОЕ ЗНАЧЕНИЕ HI2      ='',620.5)
710 IF (KEY(7).EQ.0) GOTO 810
    WRITE (NFWR,750) KSEQ
750 FORMAT(/' ',8X,12X,'К О Д          ВЫБОРКИ  '
*         //' ',8X, 4X,4A8)

C 810 IF (KEY(8).EQ.0) GO TO 999
C
    WRITE (NFWR,860)
860 FORMAT(/' ',8X,'МАССИВ ЗНАЧЕНИЙ ВЫЧИСЛЕННЫХ ДОЛЕЙ: ')
    WRITE (NFWR,870) P
870 FORMAT(' ',8X,4F12.5)

C 999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
    END

```

```

CC; DIAGR - ПЕЧАТЬ ДИАГРАММ РАСПРЕДЕЛЕНИЯ ДОЛЕЙ

C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫВОДА В ФАЙЛ С НОМЕРОМ 'NFWR'
C ДИАГРАММЫ РАСПРЕДЕЛЕНИЯ ДОЛЕЙ В ВЫБОРКИ
C И СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК ВЫБОРКИ
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : Н Е Т
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA ПРОГРАММА ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММО,
C Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE DIAGR (KEY,NFWR,KSQ,IST,N,K,C,P,T,MD,IND,IER)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
C ИМЯ ТИП РАЗМ.ФОРМАТ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
CP
C KEY I*4 (8) -- МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<NFWR<99/
C KSQ R*8 (4) 4A8 К О Д ВЫБОРКИ
C IST I*4 1 -- ТИП ШКАЛЫ: 0-НОМИНАЛЬНАЯ, 1-РАНГОВАЯ
C N I*4 1 I8 ОБЪЕМ (ДЛИНА) ВЫБОРКИ
C K I*4 1 I8 КОЛИЧЕСТВО ГРАДАЦИЙ / 2<= K <=9
C C I*4 (K) G11.4 МАССИВ ЗНАЧЕНИЙ УРОВНЕЙ ДЛЯ ГРАДАЦИЙ
C P R*4 (K) F10.5 МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ / СУММА =1 /
C T R*4 (8) F8.5 МАССИВ СТАТИСТИЧ. ХАРАКТЕРИСТИК ВЫБОРКИ
C (СМ. ОПИСАНИЕ "Y" В SP110, SP510)
C
C MD I*4 1 -- МАКСИМАЛЬНАЯ ВЫСОТА ДИАГРАММЫ, ПОЗИЦИЯ
C IND I*4 1 I1 ИНДЕКС ПРОВЕРКИ ПО ХИ-КВАДРАТ :
C IND= 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
C IND= 1 - НУЛЕВАЯ ГИПОТЕЗА ОТВЕРГАЕТСЯ.
C IER I*4 1 -- ВЫРАБОТАННЫЙ КОД ОШИБКИ
C
C DIMENSION C(K),P(K),T(8),KEY(8),GRAD(9),OUT(59),PER(9)
C * ,KSQ(4),DP(9),DM(9),LINE(4,9)
C INTEGER C,VP1,VP2,VSD,SD,CONST,COL
C REAL*8 NAME,KSQ
C LOGICAL*1 BLANK,MINUS,MINA,VRG,GRAD,OUT,PER
C DATA BLANK/1H /, MINUS/1H-/, MINA/1H8/, VRG/1H1/
C * ,GRAD/1H1,1H2,1H3,1H4,1H5,1H6,1H7,1H8,1H9/
C * ,OUT/59*1H /,PER/9*1H /,VSD/1/,CONST/13/,COL/59/
C EQUIVALENCE (Y(1),S),(Y(2),V),(Y(3),SP),(Y(4),A)
C * ,(Y(5),HI2),(Y(6),THI),(Y(7),DN),(Y(8),D)
C DATA NAME/' DIAGR '/,ZERO,HALF,ONE/0.0,0.5,1.0/
C
C 1 FORMAT(' ',59A1)
C 4 FORMAT(' ',11X,'ДИАГРАММА РАСПРЕДЕЛЕНИЯ ДОЛЕЙ В ВЫБОРКЕ'
C * '//',8X,'К О Д ВЫБОРКИ : ',4A8 /)
C 5 FORMAT(' ',16X,'ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ВЫБОРКИ : '
C 1 '//',11X,'ОБЪЕМ ВЫБОРКИ ',13X,'-',',I8
C 2 '//',11X,'ЧИСЛО ГРАДАЦИЙ',13X,'-',',I8 )
C 6 FORMAT(' ',11X,'ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ - ',F8.5
C 1 '//',11X,'КОЭФФИЦИЕНТ ВАРИИРОВАНИЯ - ',F8.5
C 2 '//',11X,'КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ - ',F8.5
C 3 '//',11X,'ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ (D) - ',F8.5
C 4 '//',11X,'МОДА',23X,'-',',F8.5
C 5 '//',11X,'КРИТЕРИЙ ХИ-КВАДРАТ : '
C 6 '//',11X,'СПРАВЕДЛИВА ГИПОТЕЗА Н',I1 )
C 7 FORMAT(' ',2X,' NN УРОВНИ ',4X,'ДОЛЯ-D'
C 1 ',5X,'ДОЛЯ',5X,'ДОЛЯ+D',2X,'ПЕРЕКРЫТИЕ'
C 2 '//',7X,'ГРАДАЦИЙ' )
C 8 FORMAT(' ',3X,I2,G11.4,1X,3F10.5,3X,9A1 )

```

```

9 FORMAT(' ',A8,5X,'K=',I2)
10 FORMAT(' ',11X,'МЕДИАНА',20X,'-',F8.5 )
C
C ЕСЛИ ЧИСЛО ГРАДАЦИИ < 2 ИЛИ > 9, ТО РАБОТА ПРЕКРАЩАЕТСЯ
C
      IF (K.LT.2) GO TO 999
      IF (K.GT.9) GO TO 999
C
C
      KLU=KEY(1)
      IF (KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11  KLU=IABS(KLU)
      DO 13 J=1,8
          I=9-J
          KEY(I)=MOD(KLU,2)
13  KLU=KLU/2
      GOTO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15  DO 17 I=1,8
      KEY(I)=1
      IF (KLU.LT.I) KEY(I)=0
17  CONTINUE
19  CONTINUE
C
      IF ( KEY(5).EQ.0 .AND. KEY(6).EQ.0 ) GO TO 999
      WRITE (NFWR,9) NAME,K
C ОПРЕДЕЛЕНИЕ КОЛИЧЕСТВА ПЕЧАТАЕМЫХ СТРОК ДИАГРАММЫ
      SD = 41
      IF (MD .GT. 0) SD = VSD+MD
C
      DO 55 I=1,59
55  OUT(I)=BLANK
C
C
C ШИРИНА СТОЛБИКА - L (БЕЗ ОГРАНИЧИТЕЛЕЙ)
      L=CONST-(K/2)*2
C ШИРИНА ДИАГРАММЫ - LG
      LG=K*(L+1)+1
C ПОЗИЦИЯ 1-ГО ЛЕВОГО ОГРАНИЧИТЕЛЯ - M
      M=(COL-LG)/2+1
C ПОЗИЦИЯ ПОСЛЕДНЕГО ПРАВОГО ОГРАНИЧИТЕЛЯ - LP
      LP=M+LG-1
C ШАГ ОГРАНИЧИТЕЛЕЙ - J
      J = L+1
C ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ
      D = T(8)
C НАХОЖДЕНИЕ МИНИМАЛЬНОЙ ДОЛИ
      Y = P(1)
      DO 65 I=2,K
          IF (P(I).LT. Y) Y=P(I)
65  CONTINUE
C ВЫЧИСЛЕНИЕ РАЗМЕРОВ ПО ЗНАЧЕНИЮ
      Z = T(4)+D
      R = Y -D
      IF (R.GE. ZERO) R=ZERO
      Z=Z-R
C ВЫЧИСЛЕНИЕ МАСШТАБА
      Y = Z/FLOAT(MD)
C ЗАПОЛНЕНИЕ ТАБЛИЦ
      DO 75 I=1,K
          Z = P(I)
          DP(I) = Z+D
          DM(I) = Z-D
          LINE(1,I) = SD-INT((Z+R)/Y+HALF)
          LINE(2,I) = SD-INT((Z+D+R)/Y+HALF)
          LINE(3,I) = SD-INT((Z-D+R)/Y+HALF)
          LINE(4,I) = 0

```

```

75  CONTINUE
    IF ( KEY(5).EQ.0 ) GOTO 416 ]
C  ПОЗИЦИЯ НУЛЕВОЙ ЛИНИИ
    NULL = SD-INT(R/Y+HALF)
C  ЦИКЛ ПО НОМЕРУ СТРОКИ
    DO 395 IS=VSD,SD
C  ЦИКЛ ВНУТРИ СТРОКИ (ОПРЕДЕЛЕНИЕ ЗОН)
    DO 165 I=1,K
        NK=LINE(4,I)+1
        GO TO (105,115,125,135,145,165),NK
105     IF (LINE(2,I).NE.IS) GO TO 165
        LINE(4,I)=1
115     IF (LINE(1,I).NE.IS) GO TO 165
        LINE(4,I)=2
        GO TO 165
125     LINE(4,I)=3
135     IF (LINE(3,I).GE.IS) GO TO 165
        LINE(4,I)=4
145     IF (NULL.NE.IS) GO TO 165
        LINE(4,I)=5
165  CONTINUE
C  ФОРМИРОВАНИЕ СТРОКИ
    DO 205 I=M,LР
205     OUT(I)=BLANK
        DO 295 I=1,K
            LK = M+(I-1)*J+1
            KJ = LINE(4,I)+1
C  ФОРМИРОВАНИЕ ЗОН
            GO TO (295,225,235,255,275),KJ
C  ФОРМИРОВАНИЕ ОГРАНИЧИТЕЛЕЙ
215     IP = LK-1
        OUT(IP) =VRG
        OUT(IP+J)=VRG
            IF (KJ.EQ.5) GO TO 295
C  ФОРМИРОВАНИЕ ДОВЕРИТЕЛЬНОГО ИНТЕРВАЛА
225     IP = LK-1+J/2
        OUT(IP) =MINA
        GO TO 295
235     KP = LK+J-2
        DO 245 IP=LK,KP
245     OUT(IP)=MINUS
        GO TO 215
C  ЗАПОЛНЕНИЕ СТОЛБИКА НОМЕРОМ ДОЛИ
255     KP = LK+J-2
        DO 265 IP=LK,KP
265     OUT(IP)=GRAD(I)
        GO TO 215
C  ФОРМИРОВАНИЕ НУЛЕВОЙ ЛИНИИ
275     KP=LK+J-1
        DO 285 IP=LK,KP
285     OUT(IP)=MINUS
C  СБРОС ИНДЕКСА ЗОНЫ
        LINE(4,I)=0
295  CONTINUE
C  НАХОДИМСЯ ЛИ МЫ НА НУЛЕВОЙ ЛИНИИ ?
        IF (IS.NE.NULL) GO TO 355
C  ФОРМИРОВАНИЕ НУЛЕВОЙ ЛИНИИ СПРАВА И СЛЕВА ОТ "СТОЛБИКОВ"
        KJ = (COL-LG)/4
        LK = M+LG
        KP = LK+KJ-1
        DO 315 IP=LK,KP
315     OUT(IP)=MINUS
        LK = M-KJ
        KP = M
        DO 335 IP=LK,KP
335     OUT(IP)=MINUS

```

```

C Вывод строки диаграммы
355 CONTINUE
    WRITE (NFWR,1) OUT
    IF (IS.LT.NULL) GO TO 395
    DO 365 I=1,59
365     OUT(I)=BLANK
395 CONTINUE
C Восстановление свободной строки
    DO 405 I=1,59
405     OUT(I)=BLANK
C 1-А СВОБОДНАЯ
    WRITE (NFWR,1) OUT
C ПОДПИСЬ ПОД ДИАГРАММОЙ
    WRITE (NFWR,4) KSG
C 1-А СВОБОДНАЯ
    WRITE (NFWR,1) OUT
C Вывод числовых характеристик
    WRITE (NFWR,5) N, K
    WRITE (NFWR,6) T(1), T(2), T(3), T(8), T(4), IND
    IF (IST.EQ.0) GO TO 410
    WRITE (NFWR,10) T(7)
    GO TO 415
410 WRITE (NFWR,1) OUT
415 CONTINUE
C 1-А СВОБОДНАЯ
    IF ( KEY(6).EQ.0 ) GOTO 999
    WRITE (NFWR,1) OUT
C Вывод шапки таблицы
416 WRITE (NFWR,7)
C Вывод таблицы
    DO 425 I=1,K
C ФОРМИРОВАНИЕ СТРОКИ УКАЗАТЕЛЕЙ ПЕРЕКРЫТИЯ ГРАНИЦ ДОЛЕЙ
    DO 417 J=1,9
417     PER(J)=BLANK
    DO 421 J=1,K
        IF (I.EQ. J) GO TO 420
        IF (P(I).LE. P(J)) GO TO 418
        IF (DM(I).LE. DP(J)) GO TO 419
        GO TO 421
418     IF (DP(I).LT. DM(J)) GO TO 421
419     PER(J) = GRAD(J)
        GO TO 421
420     PER(J) = MINA
421 CONTINUE
    Z = FLOAT(C(I))
425 WRITE (NFWR,8) I,Z,DM(I),P(I),DP(I),PER
C 1-А СВОБОДНАЯ
    WRITE (NFWR,1) OUT
C
999 RETURN
CV1 ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
END
02-DEC-87

```

1.4. Вычисление статистических характеристик двух выборок одновременно и формирование матрицы мультиномиальных частот. Матрица мультиномиальных частот необходима для вычисления корреляционной связи двух выборок, измеренных в номинальных шкалах. Если выборки представляют собой результат измерения с числами градаций r и c , то соответственно в результате суммирования в таблицу сопряженности « $r \times c$ » вносятся частоты n_{ij} — число объектов, одновременно соответствующих i -й градации по первому признаку ($i = 1, 2, \dots, r$; $j = 1, 2, \dots, c$). Объемы обеих выборок должны быть, естественно, равны.

Для построения матрицы мультиномиальных частот необходимо одновременно обрабатывать две выборки. Эту задачу выполняет подпрограмма *SP120*. Программная реализация задачи эквивалентна, в

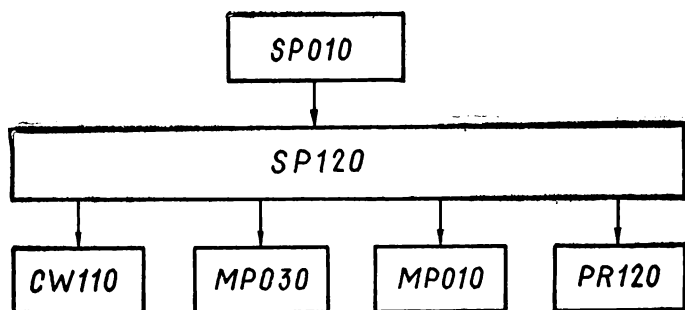


Рис. 52

основном, задаче обработки одной выборки. Вместо *SP110* подключается *SP120*; подпрограмма *DIAGR* не используется, вывод результатов работы осуществляет *PR120*. На рис. 52 представлена программная реализация обработки двух выборок одновременно.

1.5. Пример. В качестве примера проведем статистическую обработку двух выборок объемом по 100 элементов каждая при одинаковом числе градаций: $r = c = 4$. Элементы выборок приведены в задании (см. массивы данных после ключевых слов *FORX—01* и *FORX—02*) по 30 чисел в строке. Уровни градаций 1, 2, 3, 4 соответствуют числовым кодам элементов выборок. Необходимо задать также параметры:

IRP = 21 — код функции обработки двух выборок одновременно;

Q = 0,05 — уровень значимости риска (5 %);

MD = 20 — число, не равное нулю.

Далее следует текст задания на обработку, распечатка результатов выполнения задания и тексты.


```

C; SP120 - ОБРАБОТКА ДВУХ ВЫБОРОК, ПОСТРОЕНИЕ ТАБЛИЦЫ ЧАСТОТ
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫЧИСЛЕНИЯ СТАТИСТИЧЕСКИХ
C ХАРАКТЕРИСТИК ДВУХ ВЫБОРОЧНЫХ СОВОКУПНОСТЕЙ,
C; ПОСТРОЕНИЯ ТАБЛИЦЫ МУЛЬТИНОМИНАЛЬНЫХ ЧАСТОТ
C
CR ЛИТЕРАТУРА ПО АЛГОРИТМУ : Г.ШВАРЦ, ВЫБОРОЧНЫМ МЕТОД,
C; М., СТАТИСТИКА, 1978.
C
C; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
C П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧМФФ,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE SP120(KEY,NFWR,KSQN,X,NX,N,LS,K1,K2,C,MD,Q
* ,TM,P,Y,IND,IER,IERW)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
CP ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C
C KEY I*4 * * (8) МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 * * 1 НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<NFWR<=99/
C KSQN R*8 * * (4,2) К О Д Ы ВЫБОРОК
C X I*4 * * (NX) МАССИВ ЗНАЧЕНИЙ ВЫБОРОЧНОЙ СОВОКУПНОСТИ
C / НЕ РАВНЫХ ВСЕ ТОЖДЕСТВЕННО КОНСТАНТЕ /
C NX I*4 * * 1 СУММАРНАЯ ДЛИНА ВЫБОРОК /NX>=4/
C N I*4 * * (2) МАССИВ ДЛИН ВЫБОРОК
C LS I*4 * * 1 СУММАРНОЕ КОЛИЧЕСТВО ГРАДАЦИЙ /L>=4/
C K1 I*4 * * 1 КОЛИЧЕСТВО ГРАДАЦИЙ ПЕРВОЙ ВЫБОРКИ
C K2 I*4 * * 1 КОЛИЧЕСТВО ГРАДАЦИЙ ВТОРОЙ ВЫБОРКИ
C C I*4 * * (LS) МАССИВ ЗНАЧЕНИЙ УРОВНЕЙ (ДЛЯ ГРАДАЦИЙ)
C / НЕ РАВНЫХ ВСЕ ТОЖДЕСТВЕННО КОНСТАНТЕ /
C MD I*4 * * 1 МАКСИМАЛЬНАЯ ВЫСОТА ДИАГРАММЫ, ПОЗИЦИИ
C Q R*4 * * 1 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА /.01<Q<=.5/
C TM R*4 * * (K1,K2) ТАБЛИЦА МУЛЬТИНОМИНАЛЬНЫХ ЧАСТОТ
C P R*4 * * (LS) МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ /СУММА =1.0+-EPS/
C Y R*4 * * (8,2) МАССИВ ВЫЧИСЛЕННЫХ СТАТ. ХАРАКТЕРИСТИК;
C Y(1,I)=S - ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ I-И ВЫБОРКИ;
C Y(2,I)=V - КОЭФФИЦИЕНТ ВАРИИРОВАНИЯ - I - ;
C Y(3,I)=SP - КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ - I - ;
C Y(4,I)=ZMD - МОДА - I - ;
C Y(5,I)=HI2 - ЗНАЧЕНИЕ КРИТЕРИЯ ХИ-КВАДРАТ - I - ;
C Y(6,I)=THI - КРИТИЧЕСКОЕ ЗНАЧЕНИЕ ХИ-КВАДРАТ - I - ;
C Y(7,I)=0 - РЕЗЕРВИРОВАНО;
C Y(8,I)=D - ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ - I - .
C IND I*4 * * (2) ИНДЕКСЫ ПРОВЕРКИ ПО ХИ-КВАДРАТ;
C IND=0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
C IND=1 - НУЛЕВАЯ ГИПОТЕЗА ОТВЕРГАЕТСЯ.
C IER I*4 * * (2) ВЫРАБОТАННЫЕ КОДЫ ОШИБОК
C IERW I*4 * * (8,2) КОДЫ ОШИБОК, ПРИНЯТЫЕ ИЗ CИ110
C
DIMENSION KEY(8),X(NX),C(LS),P(LS),Y(8,2),KSQN(4,2),N(2)
* ,TM(K1,K2),IND(2),IER(2),IERW(8,2)
C
INTEGER X,C
REAL*8 NAME,KSQN
DATA NAME /' SP120 ',ZERO/0.0/,ONE/1.0/,EPS/0.001/
IND(1) = 11
IND(2) = 11
IER(1) = 0
IER(2) = 0
C
C ЕСЛИ KEY(1) > 0, ПРОВЕРКА И ПЕЧАТЬ ИСХОДНЫХ ДАННЫХ
N1 = N(1)
N2 = N(2)
NMAX = MAX0(N1,N2)

```

```

      NMIN = MIN0(N1,N2)
      KLC=KEY(1)
      KLP=KEY(2)
      IF (KEY(1).EQ.0) GO TO 140
      CALL CW110(KEY,NFWR,KSQN(1,1),X(1),N1,K1,C(1),Q
      ,IERW(1,1),MWC1)
*
CL; CW110 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ SP110,SP120;
      KEY(1) = KLC
      CALL CW110(KEY,NFWR,KSQN(1,2),X(N1+1),N2,K2,C(K1+1),Q
      ,IERW(1,2),MWC2)
*
      IF (LS .NE. (K1+K2)) GO TO 60
      IF (NX .EQ. (N1+N2)) GO TO 100
60 IER(1) = 4
   IER(2) = 4
   GO TO 999
C АНАЛИЗ КОДОВ ОШИБОК : "KF" И "MWC"
100 KF = IERW(1,1)
   IF (MWC1.LT.3) GO TO 120
C ЕСЛИ MWC1 >= 3, TO IER(1)=MWC1
   IER(1) = MWC1
   IF (MWC1.GE.3) GO TO 999
C
120 KF = IERW(1,2)
   IF (MWC2.LT.3) GO TO 140
C ЕСЛИ MWC2 >= 3, TO IER(2)=MWC2
   IER(2) = MWC2
   IF (MWC2.GE.3) GO TO 999
C УСТАНОВЛЕНЫ НАЧАЛЬНЫЕ ЗНАЧЕНИЯ КОДОВ ОШИБОК IER(1) И IER(2)
140 CONTINUE
   DO 200 I=1,LS
200   P(I) = ZERO
   DO 210 I=1,K1
   DO 210 J=1,K2
210   TM(I,J) = ZERO
   K = K1+1
   L1= K1
   L2= K2
   DO 600 I=1,NMAX
C ОПРЕДЕЛЕНИЕ ПРИНАДЛЕЖНОСТИ ЭЛЕМЕНТОВ 1-Й ВЫБОРКИ L1-Й ГРАДАЦИИ
   M = 1
   IF (I .GT. N1) GO TO 400
   J = X(I)
250   IF ( J .LE. C(L1) ) GO TO 270
   L1 = L1+1
   M = 0
   GO TO 250
270   IF ( M.EQ.0 .OR. L1.EQ.1 ) GO TO 300
   L1 = L1-1
   GO TO 250
300   P(L1)=P(L1)+ONE
C ОПРЕДЕЛЕНИЕ ПРИНАДЛЕЖНОСТИ ЭЛЕМЕНТОВ 2-Й ВЫБОРКИ L2-Й ГРАДАЦИИ
400   M = 1
   IF (I .GT. N2) GO TO 550
   J = X(N1+I)
450   IF ( J .LE. C(K1+L2) ) GO TO 470
   L2 = L2+1
   M = 0
   GO TO 450
470   IF ( M.EQ.0 .OR. L2.EQ.1 ) GO TO 500
   L2 = L2-1
   GO TO 450
500   P(K1+L2)=P(K1+L2)+ONE
C ФОРМИРОВАНИЕ ТАБЛИЦЫ МУЛЬТИНОМИНАЛЬНЫХ ЧАСТОТ
550   IF (I .LE. NMIN) GO TO 590
   GO TO 600
, 590   TM(L1,L2)=TM(L1,L2)+ONE

```

```

600 CONTINUE
    IF(N1 .EQ. N2) GO TO 610
    IER(1) = 2
610 D=FLOAT(N1)
C  ВЫЧИСЛЕНИЕ ДОЛЕЙ И ДРУГИХ ХАРАКТЕРИСТИК 1-Й ВЫБОРКИ
    R=ONE
    M = K1
    Z = ZERO
    DO 620 I=1,K1
        IF (P(I).EQ.ZERO) M=M-I
        P(I)=P(I)/D
        IF (P(I) .GT. Z) Z=P(I)
620 R = R*P(I)
    Y(4,1)=Z
    IF(M .EQ. K1) GO TO 625
    IER(1)=3
    GO TO 650
625 Z=ONE/FLOAT(K1)
    Y(1,1)=R**Z
    Y(2,1)=Y(1,1)/Z
    Y(3,1)=Y(1,1)/SQRT(D)
    Y(7,1)=ZERO
C  ПРОВЕРКА ГИПОТЕЗЫ О РАВЕНСТВЕ ЗНАЧЕНИЙ ЧАСТОТ ( ИЛИ ДОЛЕЙ )
C  ПЕРВОЙ ВЫБОРКИ ПО КРИТЕРИЮ ХИ-КВАДРАТ
C
    R = ZERO
    DO 630 I=1,K1
630 R=R+(P(I)-Z)**2
    Y(5,1)=R/Z*D
    CALL MP030( Q,K1-1,EP8,THI,IERTHI)
CL; MP030 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ ХИ-КВАДРАТ;
    IF (IERTHI.EQ.0) GO TO 640
    IER(1)=4+IERTHI
    GO TO 810
640 Y(6,1)=THI
    IND(1)= 0
    IF (Y(5,1) .GE. Y(6,1)) IND(1)= 1
C  ВЫЧИСЛЕНИЕ ДОЛЕЙ И ДРУГИХ ХАРАКТЕРИСТИК 2-Й ВЫБОРКИ
650 D=FLOAT(N2)
    R=ONE
    M = K2
    Z = ZERO
    DO 660 I=1,K2
        K=K1+I
        IF (P(K).EQ.ZERO) M=M-1
        P(K)=P(K)/D
        IF (P(K) .GT. Z) Z=P(K).
660 R = R*P(K)
    Y(4,2)=Z
    IF(M .EQ. K2) GO TO 665
    IER(2)=3
    GO TO 810
665 Z=ONE/FLOAT(K2)
    Y(1,2)=R**Z
    Y(2,2)=Y(1,2)/Z
    Y(3,2)=Y(1,2)/SQRT(D)
    Y(7,2)=ZERO
C  ПРОВЕРКА ГИПОТЕЗЫ О РАВЕНСТВЕ ЗНАЧЕНИЙ ЧАСТОТ ( ИЛИ ДОЛЕЙ )
C  ВТОРОЙ ВЫБОРКИ ПО КРИТЕРИЮ ХИ-КВАДРАТ
C
    R = ZERO
    DO 670 I=1,K2
670 R=R+(P(K1+I)-Z)**2
    Y(5,2)=R/Z*D
    IER(2)=0
    CALL MP030( Q,K2-1,EP8,THI,IERTHI)

```

```

      IF (IERTHI.EQ.0) GO TO 680
      IER(2)=4+IERTHI
      GO TO 810
680  Y(6,2)=THI
      IND(2)= 0
      IF (Y(5,2) .GE. Y(6,2)) IND(2)= 1
C    ВЫЧИСЛЕНИЕ ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ
      CALL MP010(Q,N1-1,EPS,TQV,IERT)
CL; MP010 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ Т-РАСПРЕДЕЛЕНИЯ
      IF (IERT.EQ. 0) GO TO 700
      IER(1)=4+IERT
      GO TO 810
700  Y(8,1)=Y(3,1)*TQV
C
      CALL MP010(Q,N2-1,EPS,TQV,IERT)
      IF (IERT.EQ. 0) GO TO 710
      IER(2)=4+IERT
      GO TO 810
710  Y(8,1)=Y(3,1)*TQV
      Y(8,2)=Y(3,2)*TQV
C
810  IF (KF.NE. 3) GOTO 830
      IER(1) = -IER(1)
      IER(2) = -IER(2)
      GO TO 999
830  IF (KLP.EQ.0) GOTO 999
      KEY(1)=KLP
C    CALL DIAGR2(KEY,NFWR,KSQN,N,LS,K1,K2,C,P,Y,MD,TM,IND,IER)
CL; DIAGR2 - ПЕЧАТЬ ДИАГРАММЫ РАСПРЕДЕЛЕНИЯ ДОЛЕЙ;
      KEY(1)=KLP
      CALL PR120(KEY,NFWR,KSQN,N,LS,K1,K2,P,Y,TM,IND,IER)
CL; PR120 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP120.
999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0      2-FEB-88
      END

```

```

CC; PR120 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP120
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫВОДА В ФАЙЛ 'NFWR'
C; РЕЗУЛЬТАТОВ РАБОТЫ П/П SP120
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE PR120(KEY,NFWR,KSQN,N,LS,K1,K2,P,Y,TM,IND,IER)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЯ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<NFWR<=99/
C K SQN R*4 (4,2) 4A8 7 К О Д Ы ВЫБОРОК
C LS I*4 1 -- - СУММАРНОЕ КОЛИЧЕСТВО ГРАДАЦИЙ
C N I*4 (2) 15 7 МАССИВ ДЛИН ВЫБОРОК
C K1 I*4 1 13 7 КОЛ-ВО ГРАДАЦИЙ ПЕРВОЙ ВЫБОРКИ
C K2 I*4 1 13 7 КОЛ-ВО ГРАДАЦИЙ ВТОРОЙ ВЫБОРКИ
C P R*4 (LS) G10.3 8 МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ
C Y R*4 (8,2) G12.3 3 МАССИВ ВЫЧИСЛЕННЫХ СТАТИСТИЧЕСКИХ
C ХАРАКТЕРИСТИК (СМ. ОПИСАНИЕ SP110)
C TM R*4 (K1,K2) F8.0 8 ТАБЛИЦА МУЛЬТИНОМИНАЛЬНЫХ ЧАСТОТ
C IND I*4 (2) 11 2 ИНДЕКС ПРОВЕРКИ ПО ХИ-КВАДРАТ:
C IND=0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
C IND=1 - НУЛЕВАЯ ГИПОТЕЗА ОТКЛОНЯЕТСЯ.
C IER I*4 (2) 13 2 КОД ОШИБКИ, ВЫРАБОТАННЫЙ SP120
C
C DIMENSION KEY(8), P(LS), IND(2), IER(2), N(2)
C , Y(8,2), K SQN(4,2), TM(K1,K2)
C REAL*8 TNK(3), TS1(4), TS2(4), TS3(4), NAME, K SQN.
C DATA TNK /' STAT.OB',' РАБОТКА ',' ВЫБОРОК '/
C , TS1 /' ',' ВЫЧИС',' ЛЕНИЕ ',' /
C , TS2 /' СТАТИС',' ТИЧЕСКИХ',' ХАРАКТЕ',' РИСТИК '/
C , TS3 /' ДВУХ В',' ВЫБОРОЧНЫ',' Х СОВОКУ',' ПНОСТЕЙ '/
C DATA NAME /' PR120 '/
C
C IF (NFWR.LE.0.OR.NFWR.GE.100) GO TO 999
C
C KLU=KEY(1)
C IF (KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 11 KLU=IABS(KLU)
C DO 13 J=1,8
C I=9-J
C KEY(I)=MOD(KLU,2)
C 13 KLU=KLU/2
C GO TO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 15 DO 17 I=1,8
C KEY(I)=1
C IF (KLU.LT.I) KEY(I)=0
C 17 CONTINUE
C 19 CONTINUE
C
C IF (KEY(1).EQ.0) GO TO 210
C WRITE (NFWR,120) TS1,TNK,TS2,TS3
C 120 FORMAT(/ 1H ,26(1H*)
C * /2H *,24(1H ),3H* ,4A8
C * /2H *,3A8 ,3H* ,4A8
C * /2H *,24(1H ),3H* ,4A8
C * / 1H ,26(1H*) )

```

```

210 IF (KEY(2).EQ.0) GOTO 310
    WRITE (NFWR,230) NAME, IND(1), IND(2), IER(1), IER(2)
230 FORMAT(/' ',A8,' - ВЫЧИСЛЕННЫЕ ХАРАКТЕРИСТИКИ: '
*      ' ',A8,' ВЫБОРКА 1 ВЫБОРКА 2 '
*      ' ',8X,' СПРАВЕДЛИВА ГИПОТЕЗА',14X,' Н',I1,10X,' Н',I4
*      ' ',8X,' КОД ОШИБКИ',22X,I3,9X,I3)

C
310 IF (KEY(3).EQ.0) GOTO 710
    WRITE (NFWR,340) ((Y(I,1),Y(I,2)),I=1,6)
340 FORMAT(/' ',8X,' ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ ~',2G12.3
*      ' ',8X,' КОЭФФИЦИЕНТ ВАРИИРОВАНИЯ ~',2G12.3
*      ' ',8X,' КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ ~',2G12.3
*      ' ',8X,' МОДА ~',2G12.3
*      ' ',8X,' ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ H12 ~',2G12.3
*      ' ',8X,' КРИТИЧЕСКОЕ ЗНАЧЕНИЕ H12 ~',2G12.3)

C
710 IF (KEY(7).EQ.0) GOTO 810
    WRITE (NFWR,750) (KSN(I,1),I=1,4),N(1),K1
    , (KSN(I,2),I=1,4),N(2),K2
750 FORMAT(/' ',17X,' К О Д Ы В Ы Б О Р О К ',9X,' ОБЪЕМЫ ГРАДАЦИЙ '
*      ' ',8X,' 1-',4A8,2X,' N1-',I5,1X,' K1-',I3
*      ' ',8X,' 2-',4A8,2X,' N2-',I5,1X,' K2-',I3)

C
810 IF (KEY(8).EQ.0) GO TO 999
    WRITE (NFWR,860)
860 FORMAT(/' ',8X
*      ' ',A8,' МАССИВ ЗНАЧЕНИЙ ВЫЧИСЛЕННЫХ ДОЛЕЙ 1-Й ВЫБОРКИ : '
    WRITE (NFWR,865) (P(I),I=1,K1)
865 FORMAT(' ',8X,5G10.3)
    J = K1 + 1
    WRITE (NFWR,870)
870 FORMAT(/' ',8X
*      ' ',A8,' МАССИВ ЗНАЧЕНИЙ ВЫЧИСЛЕННЫХ ДОЛЕЙ 2-Й ВЫБОРКИ : '
    WRITE (NFWR,875) (P(I),I=J,LS)
875 FORMAT(' ',8X,5G10.3)

C
    WRITE (NFWR,880)
880 FORMAT(/' ',11X,' ТАБЛИЦА МУЛЬТИНОМИНАЛЬНЫХ ЧАСТОТ : '
    DO 885 I=1,K1
885    WRITE (NFWR,890) I, (TM(I,J),J=1,K2)
890    FORMAT(' ',8X,I2,' : ',6F8.0/(' ',11X,6F8.0))

C
999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
END

```

§ 2. Оценка различия двух независимых выборок

Оценка различия двух независимых выборок осуществляется при числе градаций $k \geq 2$ путем проверки нулевой гипотезы H_0 с использованием как параметрических критериев (t -критериев Стьюдента и интервального оценивания), так и непараметрических (величины χ^2).

2.1. t -критерий Стьюдента. Для оценки различия между одноименными долями двух выборок вычисляется критерий Стьюдента:

$$t = \frac{p_1 - p_2}{S_d}.$$

Здесь p_1 — доля градации качественного признака первой выборки; p_2 — доля той же градации качественного признака второй выборки; S_d — ошибка разности двух выборочных долей,

$$S_d = \sqrt{S_{p_1}^2 + S_{p_2}^2},$$

где S_{p_1}, S_{p_2} — стандартная ошибка доли соответствующей выборки.

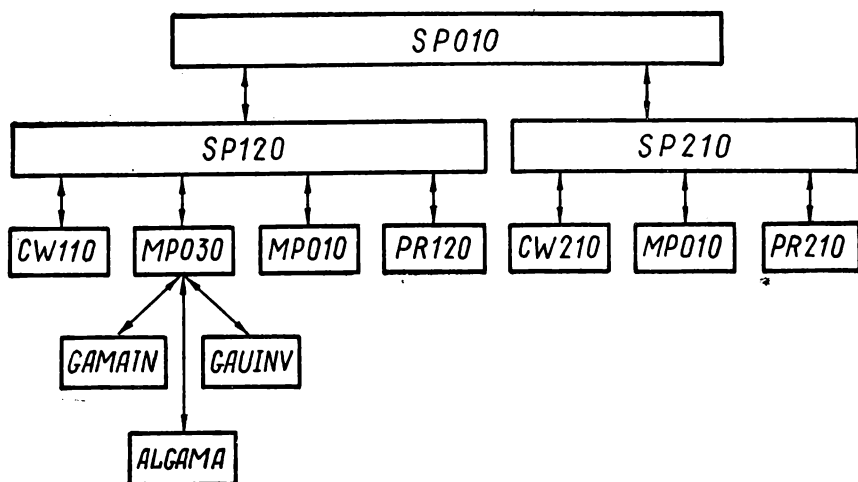


Рис. 53

Осуществляется проверка нуль-гипотезы H_0 о равенстве долей. $H_0 : p_1 = p_2$ отвергается при $t \geq t_{q,v}$; H_0 принимается при $t < t_{q,v}$, где $t_{q,v}$ — критическое значение t -критерия Стьюдента при числе степеней свободы $v = N_1 + N_2 - 2$ и уровне значимости риска q (обычно q равно 5 или 10 %).

2.2. Программная реализация оценки различия двух независимых выборок по t -критерию Стьюдента может быть представлена в виде блок-схемы на рис. 53.

Входной информацией для $SP21\theta$ являются не сами выборки, а доли градаций качественного признака, поэтому необходима предварительная обработка выборок подпрограммой $SP120$. Ниже приводятся тексты подпрограмм $SP21\theta$, $CW21\theta$, $PR21\theta$.

**ТЕСТОВОЕ ЗАДАНИЕ НА ОЦЕНКУ РАЗЛИЧИЯ 2-Х НЕЗАВИСИМЫХ
ВЫБОРОЧНЫХ СОВОКУПНОСТЕЙ И ВЫЧИСЛЕНИЕ
КОЭФФИЦИЕНТА КОРРЕЛЯЦИОННОЙ СВЯЗИ**

```

----- ОЦЕНКА РАЗЛИЧИЯ ДВУХ НЕЗАВИСИМЫХ ВЫБОРОК : 120
**** МАКЕТ ПОДГОТОВКИ ДАННЫХ ДЛЯ IR=002: ****SP210****
IRP -002 (I3)- КОД ФУНКЦИИ 220
NS -02 /ПО УМОЛЧАНИЮ/ (I2)- КОЛИЧЕСТВО ВЫБОРОК 412
K -02 (I2)- КОЛ-ВО ГРАДАЦИЙ ВЫБОРКИ 1
M -02 (I2)- КОЛ-ВО ГРАДАЦИЙ ВЫБОРКИ 2
(I2)-НОМЕР; (4A8)- К О Д ВЫБОРКИ; (I4)-ДЛИНА
XX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXX
KSNQ-01 ТЕСТОВАЯ /2 ГРАДАЦИЙ/ ВЫБОРКА 2 0100
KSNQ-02 ТЕСТОВАЯ /2 ГРАДАЦИЙ/ ВЫБОРКА 3 0100
(I2, I3,I2)C 4A8 J:НОМЕР ВЫБ.,КОД,КАНАЛ,ИМЯ ФАЙЛА
FILX-01 000 08-INF:DATA2.DAT
C СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) J
FORX-01 (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2
(I2, I3,I2)C 4A8 J:НОМЕР ВЫБ.,КОД,КАНАЛ,ИМЯ ФАЙЛА
FILX-02 000 08-INF:DATA3.DAT
C СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) J
FORX-02 (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2
(I2, I3,I2)C 4A8 J:НОМЕР ВЫБ.,КОД,КАНАЛ,ИМЯ ФАЙЛА
FILC-01 000 08-INF:DATA2.DAT
( СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) )
FORC-01 (4X,9I2)
1 2 3 4 5 6 7 8 9
(I2, I3,I2)C 4A8 J:НОМЕР ВЫБ.,КОД,КАНАЛ,ИМЯ ФАЙЛА
FILC-02 000 08-INF:DATA3.DAT
( СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) )
FORC-02 (4X,9I2)
1 2 3 4 5 6 7 8 9
Q -0.05 - (G14.7)- УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
NFWR-02 (I2) - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
(I1,I2)-ТИП КЛЮЧЕЙ ПЕЧАТИ: 0-ПОРАЗРЯДНЫЙ; И ИХ КОЛИЧЕСТВО
KEYC-0 8 1-ИНТЕРВАЛЬНЫЙ.
ИНДЕКСЫ/ИНТЕРВАЛЬНОЕ ЗНАЧЕНИЕ КЛЮЧА ПЕЧАТИ/ПОРАЗРЯДНОЕ/
12345678 - РАЗРЯДЫ КЛЮЧА ПЕЧАТИ :
KEY - 1 2 1/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SP120 :
KEY - 2 2 1/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP120 .
KEY - 1 1 2/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SP210 :
KEY - 2 1 2/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP210 .
KEY - 1 2 2/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SP220 :
KEY - 2 2 2/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP220 .
KEY - 1 1 4/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SP412 :
KEY - 2 1 4/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP412 .
MD -40 (I2)- МАКСИМАЛЬНАЯ ВЫСОТА ДИАГРАММЫ
(I1)-РАСПРЕДЕЛЕНИЕ ПАМЯТИ:=1-ВЫДАЕТСЯ ВСЕГДА;
END -1 КОНЕЦ ДАННЫХ: =0-ВЫДАЕТСЯ ТОЛЬКО АВАРИЙНО.

```


РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ ЗАДАНИЯ (РАСПЕЧАТКА РЕШЕНИЯ):

```

JRP - 2
K - 2
M - 2
K8QN NN XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX NNNN
K8QN- 1 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 2 100
K8QN- 2 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 3 100
FILX- 1 0 8-INF:DATA2.DAT
FORX- 1: (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2
FILX- 2 0 8-INF:DATA3.DAT
FORX- 2: (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2
FILX- 1 0 8-INF:DATA2.DAT
FORX- 1: (4X,9I2)
1 2
FILX- 2 0 8-INF:DATA3.DAT
FORX- 2: (4X,9I2)
1 2
Q 0.5000000E-01
NFWR 2
KEYC 0 8 8
KEY 1 2 1/8/11111111/
KEY 2 2 1/8/11111111/
KEY 1 1 2/8/11111111/
KEY 2 1 2/8/11111111/
KEY 1 2 2/8/11111111/
KEY 2 2 2/8/11111111/
KEY 1 1 4/8/11111111/
KEY 2 1 4/8/11111111/
MD 40
ОБЪЕМ РАБОЧЕЙ ПАМЯТИ - 4000 ЕДИНИЦ INTEGER
ТРЕБУЕМЫЙ ОБЪЕМ ПАМЯТИ - 481 ЕДИНИЦ INTEGER

ИНДЕКСЫ НАЧАЛА ПЕРЕМЕННЫХ В РАБОЧЕЙ ПАМЯТИ :

1 - 1 6 - 6 11 - 107 16 - 162
2 - 2 7 - 7 12 - 307 17 - 166
3 - 3 8 - 36 13 - 311 18 - 412
4 - 4 9 - 19 14 - 157 19 - 419
5 - 5 10 - 105 15 - 161 20 - 426

CW110 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0
УРОВЕНЬ ЗНАЧИМОСТИ РИСКА Q = 0.050
КОД ВЫБОРКИ : ОБЪЕМ :
ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 2 100
КОЛИЧЕСТВО ГРАДАЦИИ - K = 2
ЗНАЧЕНИЯ ЭЛ. ВЫБОРОЧНОЙ СОВОКУПНОСТИ
1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2

```

МАССИВ ЗНАЧЕНИЙ УРОВНЕЙ:

1 2

СW110 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА Q = 0.050

К О Д В Ы Б О Р К И :

ОБЪЕМ :

ТЕСТОВАЯ /2 ГРАДАЦИЙ/ В Ы Б О Р К А 3

100

КОЛИЧЕСТВО ГРАДАЦИЙ - K = 2

ЗНАЧЕНИЯ ЭЛ. В Ы Б О Р О Ч Н О Й С О В О К У П Н О С Т И

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |

МАССИВ ЗНАЧЕНИЙ УРОВНЕЙ:

1 2

* * * * * В Ы Ч И С Л Е Н И Е
 * * * * * С Т А Т . О Б Р А Б О Т К А В Ы Б О Р О К * * * * * С Т А Т И С Т И Ч Е С К И Х Х А Р А К Т Е Р И С Т И К
 * * * * * * * * * * Д В У Х В Ы Б О Р О Ч Н Ы Х С О В О К У П Н О С Т Е Й
 * * * * *

PR120 - В Ы Ч И С Л Е Н Н Ы Е Х А Р А К Т Е Р И С Т И К И : В Ы Б О Р К А 1 В Ы Б О Р К А 2

| СПРАВЕДЛИВА ГИПОТЕЗА | H1 | H1 |
|------------------------------|-----------|-----------|
| КОД ОШИБКИ | 0 | 0 |
| ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ - | 0.490 | 0.300 |
| КОЭФФИЦИЕНТ ВАРИИРОВАНИЯ - | 0.980 | 0.600 |
| КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ - | 0.490E-01 | 0.300E-01 |
| МОДА - | 0.600 | 0.900 |
| ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ H12 - | 4.00 | 64.0 |
| КРИТИЧЕСКОЕ ЗНАЧЕНИЕ H12 - | 3.84 | 3.84 |

К О Д Ы В Ы Б О Р О К

ОБЪЕМЫ ГРАДАЦИЙ

1-ТЕСТОВАЯ /2 ГРАДАЦИЙ/ В Ы Б О Р К А 2 N1- 100 K1- 2

2-ТЕСТОВАЯ /2 ГРАДАЦИЙ/ В Ы Б О Р К А 3 N2- 100 K2- 2

МАССИВ ЗНАЧЕНИЙ ВЫЧИСЛЕННЫХ ДОЛЕЙ 1-й В Ы Б О Р К И :
 0.600 0.400

МАССИВ ЗНАЧЕНИЙ ВЫЧИСЛЕННЫХ ДОЛЕЙ 2-й В Ы Б О Р К И :
 0.900 0.100

ТАБЛИЦА МУЛЬТИНОМИНАЛЬНЫХ ЧАСТОТ :

| | | |
|----|-----|-----|
| 1: | 60. | 0. |
| 2: | 30. | 10. |

CH210 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА $\alpha = 0.050$

КОДЫ ВЫБОРОК : ОБ"ЕМ
1 - ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА 2 N1 - 100
2 - ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА 3 N2 - 100
КОЛИЧЕСТВО ГРАДАЦИЙ - K = 2

МАССИВ ЗНАЧЕНИЯ ДОЛЕЙ 1-Я ВЫБОРКИ

0.60000 0.40000

МАССИВ ЗНАЧЕНИЯ ДОЛЕЙ 2-Я ВЫБОРКИ

0.90000 0.10000

* ОЦЕНКА РАЗЛИЧИЯ
* КРИТЕРИЙ СТУДЕНТА * МЕЖДУ ДОЛЯМИ 2-Х ВЫБОРОК
* ПО КРИТЕРИЮ СТУДЕНТА

PR210 - СПРАВЕДЛИВА ГИПОТЕЗА H1 ; КОД ОШИБКИ = 0

МАКСИМАЛЬНОЕ ЗНАЧЕНИЕ Т-КРИТЕРИЯ 5.22233

КРИТИЧЕСКОЕ ЗНАЧЕНИЕ Т-РАСПРЕДЕЛЕНИЯ 1.97984

* ГРАДАЦИЯ - 1 : T>=КРИТ. -СПРАВЕДЛИВА ГИПОТЕЗА H1

* ГРАДАЦИЯ - 2 : T>=КРИТ. -СПРАВЕДЛИВА ГИПОТЕЗА H1

КОДЫ ВЫБОРОК :

1 - ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА 2

2 - ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 3

МАССИВ ВЫЧИСЛЕННЫХ КОЭФФИЦИЕНТОВ Т-КРИТЕРИЯ

5.22233 5.22233

CH220 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА $\alpha = 0.050$

КОДЫ ВЫБОРОК : ОБ"ЕМ
1 - ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА 2 N1 - 100
2 - ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 3 N2 - 100
КОЛИЧЕСТВО ГРАДАЦИЙ - K = 2

МАССИВ ЗНАЧЕНИЯ ДОЛЕЙ 1-Я ВЫБОРКИ

0.60000 0.40000

МАССИВ ЗНАЧЕНИЯ ДОЛЕЙ 2-Я ВЫБОРКИ

0.90000 0.10000

* ОЦЕНКА РАЗЛИЧИЯ
* ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ * МЕЖДУ ДОЛЯМИ 2-Х ВЫБОРОК ПУТЕМ
* СОПОСТАВЛЕНИЯ ДОВ. ИНТЕРВАЛОВ

PR220 СПРАВЕДЛИВА ГИПОТЕЗА H1 ; КОД ОШИБКИ = 0

МАКСИМАЛЬНАЯ РАЗНОСТЬ ДОЛЕЙ 0.30000

СУММА ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ 0.16539

КРИТ. ЗНАЧ. Т-РАСПР. ДЛЯ 1-Я ВЫБ 2.0938

КРИТ. ЗНАЧ. Т-РАСПР. ДЛЯ 2-Я ВЫБ 2.0938

* ДОВЕРИТЕЛЬНЫЕ ГРАНИЦЫ 1-Х ДОЛЕЙ НЕ СОВПАДАЮТ

* ДОВЕРИТЕЛЬНЫЕ ГРАНИЦЫ 2-Х ДОЛЕЙ НЕ СОВПАДАЮТ

* ДОВЕРИТЕЛЬНЫЕ ГРАНИЦЫ ГЕН.ВЫБОРОК НЕ СОВПАДАЮТ

К О Д Ы В Ы Б О Р О К :

| | |
|-----------------------------------|---|
| 1 - ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА | 2 |
| 2 - ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА | 3 |

СН412 КОД МАКСИМАЛЬНОЙ ОШИБКИ =3

ИМЕНА ПАРАМЕТРОВ : NFWR, T, L, N, Q
 К О Д Ы О Ш И Б О К : 0 3 0 0 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА = 0.050

ТАБЛИЦА ДАННЫХ СФОРМИРОВАНА ИЗ ВЫБОРОК :

| | |
|-----------------------------------|---|
| 1 - ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА | 2 |
| 2 - ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА | 3 |

СУММА ДЛИН ВЫБОРОК = 200

< ТАБЛИЦА (2*2) ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ :

| | |
|---------|----------|
| 60.0000 | 0.000000 |
| 30.0000 | 10.0000 |

```

CC; SP210 - КРИТЕРИИ СТЬЮДЕНТА СРАВНЕНИЯ 2-Х ВЫБОРОК
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ СРАВНЕНИЯ 2-Х НЕЗАВИСИМЫХ
C; ВЫБОРОК ПО КРИТЕРИЮ СТЬЮДЕНТА
CR ЛИТЕРАТУРА ПО АЛГОРИТМУ :
C;
C; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧМФФ,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE SP210(KEY,NFWR,KSQN
* ,K,N,P1,P2,Q,T,TQV,IND,IER,IERW)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CР ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C
C KEY I*4 * * (8) МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 * 1 НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<NFWR<=99/
C K SQN R*8 * (4,2) МАССИВ К О Д О В ВЫБОРОК
C K I*4 * 1 КОЛИЧЕСТВО ГРАДАЦИЙ
C N I*4 * (2) МАССИВ ДЛИН ВЫБОРОК
C P1 R*4 * (K) МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ ПЕРВОЙ ВЫБОРКИ
C P2 R*4 * (K) МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ ВТОРОЙ ВЫБОРКИ
C Q R*4 * 1 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА /.005<Q<=.5/
C T R*4 * (K) МАССИВ ВЫЧИСЛЕННЫХ ЗНАЧЕНИЙ
C TQV R*4 * 1 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ Т-РАСПРЕДЕЛЕНИЯ
C IND I*4 * 1 ИНДЕКС ПРОВЕРКИ ПО Т-КРИТЕРИЮ :
C IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
C IND = 1 - НУЛЕВАЯ ГИПОТЕЗА ОТКЛОНЯЕТСЯ.
C IER I*4 * 1 ВЫРАБОТАННЫЙ КОД ОШИБКИ
C; IERW I*4 * (8) МАССИВ КОДОВ ОШИБОК, ПРИНЯТЫХ ИЗ CW210
C
C DIMENSION KEY(8),P1(K),P2(K),T(K),N(2),K SQN(4,2),IERW(8)
C REAL*8 NAME,K SQN
C DATA NAME /' SP210 '/,ONE/1.0/,EPS/0.001/
C
C IND = 11
C IER = 0
C ЕСЛИ KEY(1) > 0, ПРОВЕРКА И ПЕЧАТЬ ИСХОДНЫХ ДАННЫХ
C KLP=KEY(2)
C IF (KEY(1).EQ.0) GO TO 120
C CALL CW210(KEY,NFWR,K SQN,K,N,P1,P2,Q,IERW,MWC)
CL; CW210 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP120;
C
C АНАЛИЗ КОДОВ ОШИБОК : "KF" И "MWC"
C
C KF = IERW(1)
C IF (MWC.LT.3) GO TO 120
C ЕСЛИ MWC>=3, ТО IER=MWC
C IER = MWC
C IF (MWC.GE.3) GO TO 999
C
C УСТАНОВЛЕНЫ НАЧАЛЬНЫЕ ЗНАЧЕНИЯ КОДА ОШИБКИ IER;
C
C 120 N1 = N(1)
C N2 = N(2)
C S1 = ONE
C S2 = ONE
C D = ONE/FLOAT(K)
C DO 150, I=1,K
C S1 = S1 * P1(I)
C S2 = S2 * P2(I)
C 150 S1 = S1 ** D
C S2 = S2 ** D

```

```

      SP1= S1/SQRT (FLOAT (N1))
      SP2= S2/SQRT (FLOAT (N2))
C
      S = SQRT (SP1*SP1+SP2*SP2)
      DO 200 I=1,K
200    T(I)=ABS(P1(I)-P2(I))/S
      NU=N1+N2-2
C
      CALL MP010(Q,NU,EPS,TQV,IERT)
CL; MP010 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ Т-РАСПРЕДЕЛЕНИЯ,
      IF (IERT .EQ. 0) GOTO 300
      IER = 4+IERT
      GOTO 810
300 IND=0
      DO 320 I=1,K
      TIN=T(I)
      IF (TIN.GE.TQV) GO TO 310
      GO TO 320
310 IND=1
320 CONTINUE
C
      810 IF (KF .NE. 3) GO TO 830
      IER = -IER
      GO TO 999
      830 IF (KLP.EQ.0) GO TO 999
      KEY(1)=KLP
      CALL PR210(KEY,NFWR,KSQN,K,T,TQV,IND,IER)
CL; PR210 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP210.
C
      999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ 2.0
      END
      2-FEB-88

```

```

C; SW210 - КОНТРОЛЬ И/ИЛИ ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP210
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ КОНТРОЛЯ ВХОДНЫХ ДАННЫХ ДЛЯ SP210
C И ВЫДАЧИ В ФАЙЛ С НОМЕРОМ 'NFWR' РЕЗУЛЬТАТА ПРОВЕРКИ
C; И ВХОДНЫХ ДАННЫХ
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP210
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE CW210(KEY,NFWR,KSQN,K,N,P1,P2,Q,IERW,MWC)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
C K SQN R*8(4,2) 4A8 7 МАССИВ К О Д О В ВЫБОРОК
C K I*4 1 16 7 КОЛИЧЕСТВО ГРАДАЦИЙ
C N I*4 (2) 15 7 МАССИВ ДЛИН ВЫБОРОК
C P1 R*4 (K) F11.5 8 МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ ПЕРВОЙ ВЫБОРКИ
C P2 R*4 (K) F11.5 8 МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ ВТОРОЙ ВЫБОРКИ
C Q R*4 1 F8.3 3 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
C IERW I*4 (8) 612 [2] ВЫРАБОТАННЫЙ КОД ОШИБКИ : 6 РАЗРЯДОВ
C; MWC I*4 1 12 2 КОД МАКСИМАЛЬНОЙ ОШИБКИ
C
C DIMENSION KEY(8),P1(K),P2(K),K SQN(4,2),IERW(8),N(2)
C REAL*8 NAME,K SQN
C DATA NAME / ' CW210 ', CONST /0.001/
C
C KLU=KEY(1)
C IF (KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 11 KLU=IABS(KLU)
C DO 13 J=1,8
C I=9-J
C KEY(I)=MOD(KLU,2)
C 13 KLU=KLU/2
C BOTO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 15 DO 17 I=1,8
C KEY(I)=1
C IF (KLU.LT.I) KEY(I)=0
C 17 CONTINUE
C 19 CONTINUE
C
C КОНТРОЛЬ ВХОДНЫХ ДАННЫХ
C
C *- ПРОВЕРКА 'NFWR' ( 1-й РАЗРЯД )
C НАЧАЛЬНАЯ УСТАНОВКА : ОШИБОК НЕТ
C KF = 0
C ЕСЛИ 'NFWR' БОЛЬШЕ 99 ИЛИ МЕНЬШЕ 1, ЭТО ГРУБАЯ ОШИБКА
C IF (NFWR.GT.99) KF = 3
C IF (NFWR.LT.1) KF = 3
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
C IERW(1)=KF
C
C *- ПРОВЕРКА 'K' ( 2-й РАЗРЯД )
C НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'K' : 2 <= K <= 99
C KK = 0
C ЕСЛИ K >= 100 ЭТО ЛОГИЧЕСКАЯ ОШИБКА

```

```

      IF (K.GE.100) KK = 1
C  НЕПРАВИЛЬНОЕ ЗНАЧЕНИЕ 'K' : K < 2
      IF (K.LT.2) KK = 3
C  ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(2)=KK
C
C-*--      ПРОВЕРКА 'N' ( 3-И РАЗРЯД ) -
C  KN1=0 ОЗНАЧАЕТ, ЧТО 'N(1)'- НОРМАЛЬНОЕ, Т.Е. БОЛЬШЕ 9
      KN1= 0
C  ЕСЛИ 'N1' МЕНЬШЕ ЧЕМ 10, ЭТО ЛОГИЧЕСКАЯ ОШИБКА
      IF (N(1).LT.10) KN1= 1
C  ЕСЛИ 'N1' МЕНЬШЕ ЧЕМ 6, ЭТО ГРУБАЯ ОШИБКА
      IF (N(1).LT. 2) KN1 = 3
C  ПО ВЕРХНЕМУ ПРЕДЕЛУ ДЛЯ 'N1' ПРОВЕРКИ НЕТ
C
C  KN2=0 ОЗНАЧАЕТ, ЧТО 'N(2)'- НОРМАЛЬНОЕ, Т.Е. БОЛЬШЕ 9
      KN2 = 0
C  ЕСЛИ 'N(2)' МЕНЬШЕ ЧЕМ 10, ЭТО ЛОГИЧЕСКАЯ ОШИБКА
      IF (N(2).LT.10) KN2 = 1
C  ЕСЛИ 'N(2)' МЕНЬШЕ ЧЕМ 2, ЭТО ГРУБАЯ ОШИБКА
      IF (N(2).LT. 2) KN2 = 3
C  ПО ВЕРХНЕМУ ПРЕДЕЛУ ДЛЯ 'N2' ПРОВЕРКИ НЕТ
C  ФОРМИРОВАНИЕ КОДА ОШИБКИ
      KN = MAX0(KN1,KN2)
      IERW(3)=KN
C
C-*--      ПРОВЕРКА 'P1'-МАССИВА ( 4-И РАЗРЯД ) -
      KPM1= 11
C  ЕСЛИ 'K' - ПЛОХОЕ, ТО ПРОВЕРКА НЕВОЗМОЖНА
      IF (KK.GT.1) GO TO 50
C  НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'P1'
      KPM1 = 0
      S=0.0
      DO 45 I=1,K
        R=P1(I)
C  ОЧЕНЬ МАЛОЕ ЗНАЧЕНИЕ 'P1' ЭЛЕМЕНТА
      IF ( R.LT.CONST) KPM1 = 1
C  НЕПРАВИЛЬНОЕ ЗНАЧЕНИЕ 'P1' ЭЛЕМЕНТА
      IF ( R.GE.1.0.OR.R.EQ.0.0) KPM1 = 3
C  СУММИРОВАНИЕ ДОЛЕЙ 'P1'
      45 S= S + R
C  НЕПРАВИЛЬНАЯ СУММА ДОЛЕЙ 'P1'
      IF (S.GT.1.+CONST.OR.S.LT.1.-CONST) KPM1 = 3
C  ФОРМИРОВАНИЕ КОДА ОШИБКИ
      50 IERW(4)=KPM1
C
C-*--      ПРОВЕРКА 'P2'-МАССИВА ( 5-И РАЗРЯД )
      KPM2=11
C  ЕСЛИ 'K' - ПЛОХОЕ, ТО ПРОВЕРКА НЕВОЗМОЖНА
      IF (KK.GT.1) GO TO 60
C  НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'P2'
      KPM2 = 0
      S=0.0
      DO 55 I=1,K
        R=P2(I)
C  ОЧЕНЬ МАЛОЕ ЗНАЧЕНИЕ 'P2' ЭЛЕМЕНТА
      IF ( R.LT.CONST) KPM2 = 1
C  НЕПРАВИЛЬНОЕ ЗНАЧЕНИЕ 'P2' ЭЛЕМЕНТА
      IF ( R.GE.1.0.OR.R.EQ.0.0) KPM2 = 3
C  СУММИРОВАНИЕ ДОЛЕЙ 'P2'
      55 S= S + R
C  НЕПРАВИЛЬНАЯ СУММА ДОЛЕЙ 'P2'
      IF (S.GT.1.+CONST.OR.S.LT.1.-CONST) KPM2 = 3
C  ФОРМИРОВАНИЕ КОДА ОШИБКИ
      60 IERW(5)=KPM2
C

```



```

C*-   ПРОВЕРКА 'Q' ( 6-Й РАЗРЯД )
C   НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'Q' : 0.025 <= Q <= 0.2
      KQ = 0
C   ЕСЛИ 'Q' БОЛЬШЕ ЧЕМ 0.2 ИЛИ МЕНЬШЕ ЧЕМ 0.025,
C       ТО СЧИТАЕМ ЧТО ДОПУЩЕНА ЛОГИЧЕСКАЯ ОШИБКА
      IF (Q.LT.0.025) KQ = 1
      IF (Q.GT.0.2)   KQ = 1
C   ВВОДИМ ОГРАНИЧЕНИЕ : 0.005 <= Q <= 0.5
      IF (Q.LT.0.005) KQ = 3
      IF (Q.GT.0.5)   KQ = 3
C   ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(6)=KQ
      IERW(8)=6

C   ОПРЕДЕЛЕНИЕ МАКСИМАЛЬНОГО КОДА ОШИБКИ
C
      MWC = MAX0 (KK,KN,KPM1,KPM2,KQ)

C   ШКАЛА КОДИРОВКИ ОШИБОЧНЫХ ПАРАМЕТРОВ :
C
C   -----
C   РАЗРЯД КОДА ОШИБКИ  I  1  I  2  I  3  I  4  I  5  I  6
C   -----
C   ПАРАМЕТР В ОБРАЩЕНИИ I NFWR I  K  I  N  I  P1 I  P2 I  Q
C   -----
C   ВОЗМОЖНЫЕ ЗНАЧЕНИЯ  I  0  I  0  I  0  I  0  I  0  I  0
C   I  -  I  1  I  1  I  1  I  1  I  1  I  -
C   I  -  I  -  I  -  I  -  I  -  I  2
C   I  3  I  3  I  3  I  3  I  3  I  -
C   I  -  I  -  I  -  I  1  I  1  I  -
C
C   ПЕЧАТЬ ВХОДНЫХ ДАННЫХ
C
      IF (KF.GT.0) GO TO 999
      IF (KEY(2).EQ.0) GO TO 225

C
      WRITE (NFWR,220) NAME, MWC
220  FORMAT('/' ',8X,'КОД МАКСИМАЛЬНОЙ ОШИБКИ '=' ,I2)
C
225  IF (MWC.LT.1) GO TO 310
      J=IERW(8)
      WRITE (NFWR,227) (IERW(I),I=1,J)
227  FORMAT('/' ',8X,'ИМЕНА ПАРАМЕТРОВ : '
      *      , ' NFWR, K, N, P1, P2, Q'
      *      /' ',8X,'К О Д Ы ОШИБОК : ',6(3X,I2)/)
310  IF (KEY(3).EQ.0) GOTO 710
      WRITE (NFWR,330) Q
330  FORMAT('/' ',8X,'УРОВЕНЬ ЗНАЧИМОСТИ РИСКА Q =' ,F8.3)
C
710  IF (KEY(7).EQ.0) GO TO 810
      WRITE (NFWR,740) (KSN(I,1),I=1,4),N(1)
      *      , (KSN(I,2),I=1,4),N(2),K
740  FORMAT(' ' ',8X,12X,'К О Д Ы ВЫБОРОК : ',6X,
      *      2X,' ОБЪЕМ : '
      *      /' ',8X,'1 - ',4A8,2X,'N1 - ',I5
      *      /' ',8X,'2 - ',4A8,2X,'N2 - ',I5
      *      /' ',8X,'КОЛИЧЕСТВО ГРАДАЦИИ - K =' ,I6)
C
810  IF (KEY(8).LT.0) GO TO 999
C
      IF (KK.GT.1) GO TO 999
      WRITE (NFWR,850) P1
850  FORMAT('/' ',8X,'МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ 1-Й ВЫБОРКИ' /
      *      , (' ',12X,4F11.5))
C
      WRITE (NFWR,860) P2
860  FORMAT(' ' ',8X,'МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ 2-Й ВЫБОРКИ' /
      *      , (' ',12X,4F11.5))
C
999  RETURN
CV;  ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
      END

```

```

CC; PR210 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP210
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫВОДА В ФАЙЛ 'NFWR'
C; РЕЗУЛЬТАТОВ РАБОТЫ П/П SP210
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP210
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММО,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE PR210(KEY,NFWR,KSQN,K,T,TQV,IND,IER)
C
ОПИСАНИЕ ПАРАМЕТРОВ :
C
-----
C
CP ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЯ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА/1<=NFWR<=99/
C KSQN R*8 (4,2) 4A8 7 МАССИВ К О Д О В ВЫБОРК
C K I*4 1 -- - КОЛИЧЕСТВО ГРАДАЦИИ
C T R*4 (K) F12.5 8 МАССИВ ВЫЧИСЛЕННЫХ ЗНАЧЕНИЙ
C TQV R*4 1 F12.5 3 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ Т-РАСПРЕДЕЛЕНИЯ
C IND I*4 1 I1 2 ИНДЕКС ОТВЕТА ПО КРИТЕРИЮ СТУДЕНТА
C; IER I*4 1 I3 2 КОД ОШИБКИ, ВЫРАБОТАННЫЙ SP210
C
DIMENSION KEY(8),T(K),Ksqn(4,2)
REAL*4 T1(3),T5(7),P(13)
REAL*8 NAME,Ksqn
REAL*8 TNK(3),TS1(4),TS2(4),TS3(4)
C
EQUIVALENCE ( P(1),T1(1))
* , ( P(4),T2)
* , ( P(5),T3)
* , ( P(6),T4)
* , ( P(7),T5(1))
C
DATA P/'* GF','АДАЦ','ИЯ -',' : T','>=KP'
* , '<KP','ИТ. ','-СПР','АВЕД','ЛИВА'
* , 'ГИП','ОТЕЗ','А Н'
DATA TNK /' КРИТЕ','РИИ СТУ','ДЕНТА
* ,TS1 /' ,ОЦЕНКА , 'РАЗЛИЧИЯ',
* ,TS2 /' МЕЖД','У ДОЛЯМИ', '2-X ВЫБ','ОРОК
* ,TS3 /' ПО',' КРИТЕРИ','Ю СТУДЕН','ТА
DATA NAME /' PR210
P
KLU=KEY(1)
IF (KLU) 11,999,15
C - ПОРАЗЯРДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
13 KLU=KLU/2
GO TO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF (KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
IF (KEY(1).EQ.0) GO TO 210
C

```

```

      IF (NFWR.LE.0.OR.NFWR.GE.100) GO TO 999
C
      WRITE (NFWR,120) TS1,TNK,TS2,TS3
120 FORMAT(/ 1H ,26(1H*))
      *      /2H *,24(1H ),3H*      ,4A8
      *      /2H *,3A8      ,3H*      ,4A8
      *      /2H *,24(1H ),3H*      ,4A8
      *      / 1H ,26(1H*) )
210 IF (KEY(2).EQ.0) GOTO 310
C
      WRITE (NFWR,230) NAME, IND, IER
230 FORMAT(/' ',A8,' - СПРАВЕДЛИВА ГИПОТЕЗА Н'.I1
      *      ', ' ;      КОД ОШИБКИ =' ,I3)
C
      310 IF (KEY(3).EQ.0) GOTO 510
      TMAX=T(1)
      DO 340 I=2,K
      IF (T(I).GT.TMAX) TMAX=T(I)
340 CONTINUE
C
      WRITE (NFWR,350) TMAX,TQV
350 FORMAT(/' ', 8X
      *      , 'МАКСИМАЛЬНОЕ ЗНАЧЕНИЕ Т-КРИТЕРИЯ      ',F12.5
      *      /' ', 8X
      *      , 'КРИТИЧЕСКОЕ ЗНАЧЕНИЕ Т-РАСПРЕДЕЛЕНИЯ ',F12.5)
510 IF (KEY(5).EQ.0) GO TO 710
C
      DO 560 I=1,K
      IF (T(I).GE.TQV) GO TO 555
      IN = 0
      WRITE (NFWR,900) T1,I,T2,T4,T5,IN
      GO TO 560
555 CONTINUE
      IN = 1
      WRITE (NFWR,900) T1,I,T2,T3,T5,IN
560 CONTINUE
C
      710 IF (KEY(7).EQ.0) GOTO 810
      WRITE (NFWR,770) (KSN(I,1),I=1,4),(KSN(I,2),I=1,4)
770 FORMAT(/' ',8X,12X,'К О Д Ы      ВЫБОРОК      : '
      *      /' ',8X,'1 - ',4A8
      *      /' ',8X,'2 - ',4A8)
810 IF (KEY(8).EQ.0) GOTO 810
      WRITE (NFWR,880) T
880 FORMAT(/' ',8X
      *      , 'МАССИВ ВЫЧИСЛЕННЫХ КОЭФФИЦИЕНТОВ Т-КРИТЕРИЯ'
      *      , (/ ' ',8X,4F12.5))
C
      900 FORMAT(' ',8X,3A4,I2,9A4,I1)
C
      999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
      END

```

2.3. Критерий интервального оценивания. Критерий интервального оценивания может быть использован, если выполняется условие

$$pN \geq 5,$$

где p — выборочная доля; N — объем выборки.

Доверительные границы для доли генеральной совокупности определяются значениями

$$p \pm t_{q,v} S_p,$$

где $t_{q,v}$ — q -квантиль распределения Стьюдента при v степенях свободы; S_p — стандартная ошибка доли.

Если вычисленная доля генеральной совокупности значительно отклоняется от значения 0,5 в любую сторону (например, $< 0,3$ или

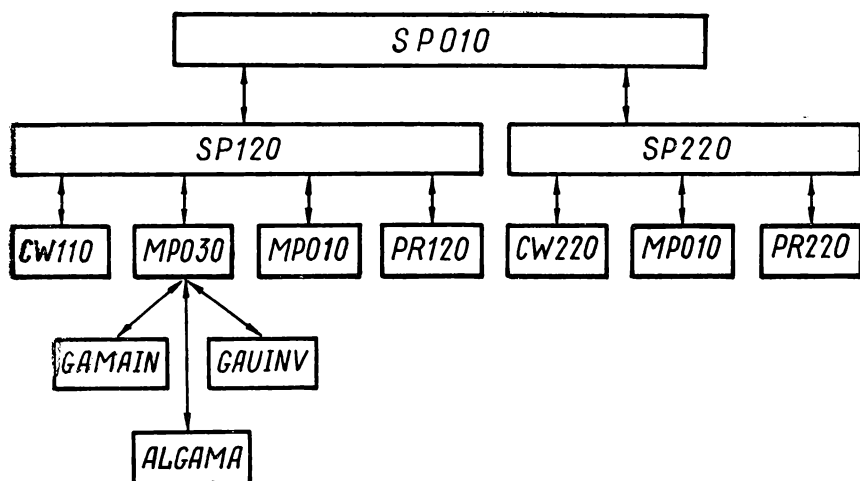


Рис. 54

$> 0,7$), то желательно вычисление несимметричных доверительных интервалов по формуле

$$p_{\text{в}}, p_{\text{н}} = \frac{1}{N + Z_q^2} \left(n_1 + \frac{Z_q^2}{2} \pm \sqrt{\frac{n_1(N - n_1)}{N} + \frac{Z_q^2}{4}} \right),$$

где $p_{\text{в}}, p_{\text{н}}$ — верхняя и нижняя доверительные границы, которые не равно удалены от p ; N — объем выборки; n_1 — наблюдаемая частота $n_1 = Np$; Z_q^2 — q -квантиль нормального распределения для заданного уровня значимости риска q (q обычно равно 5 или 10 %).

Осуществляется проверка нулевой гипотезы о равенстве долей, $H_0: p_1 = p_2$ отвергается с заданным уровнем значимости q , когда доверительные границы не перекрываются.

2.4. Программная реализация. Блок-схема программной реализации аналогична таковой п. 2. 2, с той разницей, что вместо программ SP210, CW210, PR210 используются подпрограммы SP220, CW220, PR220 соответственно (рис. 54).

```

С; SP220 - СОПОСТАВЛЕНИЕ ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ
С
СF П/П ПРЕДНАЗНАЧЕНА ДЛЯ СРАВНЕНИЯ 2-Х НЕЗАВИСИМЫХ {
С; ВЫБОРОК ПО КРИТЕРИЮ ИНТЕРВАЛЬНОГО ОЦЕНИВАНИЯ:
С
СR ЛИТЕРАТУРА ПО АЛГОРИТМУ :
С;
СТ; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
С
СА П/П ПРОВЕРЕНА : ~ НЕЧАЕВ А.Н., КАФЕДРА ЧММО,
С; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
С
SUBROUTINE SP220(KEY,NFWR,KSQN
* ,K,N,P1,P2,Q,R,D,L,TQV1,TQV2,IND,IER,IERW)
С
С ОПИСАНИЕ ПАРАМЕТРОВ :
С
С -----
С
СР ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
С
С KEY I*4 * * (8) МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
С NFWR I*4 * 1 НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
С KSQN R*8 * (4,2) МАССИВ К О Д О В ВЫБОРОК
С K I*4 * 1 КОЛИЧЕСТВО ГРАДАЦИЙ
С N I*4 * (2) МАССИВ ДЛИН ВЫБОРОК
С P1 R*4 * (K) МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ 1-Й ВЫБОРКИ
С P2 R*4 * (K) МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ 2-Й ВЫБОРКИ
С Q R*4 * 1 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
С R R*4 * 1 МАХ. РАЗНОСТЬ СООТВ. ДОЛЕЙ 2-Х ВЫБОРОК
С D R*4 * 1 СУММА ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ
С L I*4 * (K) МАССИВ ИНДЕКСОВ ПО ГРАДАЦИЯМ
С TQV1 R*4 * 1 КРИТИЧЕСКОЕ ЗНАЧ. Т-РАСПР. для 1-й ВЫБ.
С TQV2 R*4 * 1 КРИТИЧЕСКОЕ ЗНАЧ. Т-РАСПР. для 2-й ВЫБ.
С IND I*4 * 1 ИНДЕКС ПРОВЕРКИ ПО ИНТЕРВ. ОЦЕНИВАНИЮ :
С IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
С IND = 1 - НУЛЕВАЯ ГИПОТЕЗА ОТКЛОНЯЕТСЯ.
С IER I*4 * 1 ВЫРАБОТАННЫЙ КОД ОШИБКИ
С IERW I*4 * (8) МАССИВ КОДОВ ОШИБОК, ПРИНЯТЫМ ИЗ SW220
С
DIMENSION KEY(8),P1(K),P2(K),L(K)
* ,KSQN(4,2),IERW(8),N(2)
REAL*8 NAME,KSQN
DATA NAME /' SP220 '/,ZERO/0.0/,ONE/1.0/
С
IND = 11
IER = 0
KLP = KEY(2)
IF (KEY(1).EQ.0) GO TO 120
CL; SW220 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP220;
CALL SW220(KEY,NFWR,KSQN,K,N,P1,P2,Q,IERW,MWC)
С
С АНАЛИЗ КОДОВ ОШИБОК : IERW,MWC
С
KF = IERW(1)
IF (MWC.LT.3) GO TO 120
С ЕСЛИ MWC>=3, TO IER=MWC
IER = MWC
IF (MWC.GE.3) GO TO 999
С
С УСТАНОВЛЕНЫ НАЧАЛЬНЫЕ ЗНАЧЕНИЯ КОДА ОШИБКИ IER;
С
120 N1=N(1)
N2=N(2)
S1=ONE
S2=ONE
D=ONE/FLOAT(K)

```

```

      DO 150 I=1,K
        S1=S1*P1(I)
150    S2=S2*P2(I)
      S1=S1**D
      S2=S2**D
      SP1=S1/SQRT(FLOAT(N1))
      SP2=S2/SQRT(FLOAT(N2))
C
      CALL MP010(Q,N1-1,0.001,TQV1,IERT)
      IF (IERT.EQ. 0) GOTO 160
      IER = 4+IERT
      GOTO 810
160 CALL MP010(Q,N2-1,0.001,TQV2,IERT)
CL; MP010 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ Т-РАСПРЕДЕЛЕНИЯ
      IF (IERT.EQ. 0) GOTO 170
      IER = 4+IERT
      GOTO 810
170 D1 = TQV1*SP1
      D2 = TQV2*SP2
      D = D1+D2
200 CONTINUE
C
      IND=0
      R=ZERO
      DO 220 I=1,K
        D1=ABS(P1(I)-P2(I))
        IF (D1.GT.R) R=D1
        IF (D1.GT.D) GO TO 210
      L(I)=0
      GO TO 220
210 IND=1
      L(I)=1
220 CONTINUE
C
810 IF (KF.NE.3) GO TO 830
      IER = -IER
      GO TO 999
830 IF (KLP.EQ.0) GO TO 999
      KEY(1)=KLP
      CALL PR220(KEY,NFWR,KSQN,K,L,R,D,TQV1,TQV2,IND,IER)
CL; PR220 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP220
C
      999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
      END
2-FEB-88

```

```

СС; СW220 -- КОНТРОЛЬ И/ИЛИ ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP220
С;
СF П/П ПРЕДНАЗНАЧЕНА ДЛЯ КОНТРОЛЯ ВХОДНЫХ ДАННЫХ ДЛЯ SP220
С И ВЫДАЧИ В ФАЙЛ С НОМЕРОМ 'NFWR' РЕЗУЛЬТАТА ПРОВЕРКИ
С; И ВХОДНЫХ ДАННЫХ
С;
СR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP220
С;
СТ; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
С;
СА П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ,
С; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
С;
SUBROUTINE СW220(KEY,NFWR,KSQN,K,N,P1,P2,Q,IERW,MWC)
С;
С ОПИСАНИЕ ПАРАМЕТРОВ :
С -----
С
СР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
С ПЕЧАТИ
С KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
С NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
С KSQN R*8(4,2) 4A8 7 МАССИВ К О Д О В ВЫБОРОК
С K I*4 1 IS 7 КОЛИЧЕСТВО ГРАДАЦИИ
С N I*4 (2) IS 7 МАССИВ ДЛИН ВЫБОРОК
С P1 R*4 (K) F11.5 8 МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ ПЕРВОЙ ВЫБОРКИ
С P2 R*4 (K) F11.5 8 МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ ВТОРОЙ ВЫБОРКИ
С Q R*4 1 F8.3 3 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
С IERW I*4 (8) 6I2 [2] ВЫРАБОТАННЫЙ КОД ОШИБКИ : 6 РАЗРЯДОВ
С; MWC I*4 1 I2 2 КОД МАКСИМАЛЬНОЙ ОШИБКИ
С
DIMENSION KEY(8),P1(K),P2(K),KSQN(4,2),IERW(8),N(2)
REAL*8 NAME,KSQN
DATA NAME /' СW220 '/,CONST /0.001/
С
KLU=KEY(1)
IF(KLU) 11,999,15
С - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
13 KLU=KLU/2
GO TO 19
С - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF(KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
С
С КОНТРОЛЬ ВХОДНЫХ ДАННЫХ
С
С-*-- ПРОВЕРКА 'NFWR' ( 1-И РАЗРЯД )
С НАЧАЛЬНАЯ УСТАНОВКА : ОШИБОК НЕТ
KF = 0
С ЕСЛИ 'NFWR' БОЛЬШЕ 99 ИЛИ МЕНЬШЕ 1, ЭТО ГРУБАЯ ОШИБКА
IF (NFWR.GT.99) KF = 3
IF (NFWR.LT.1) KF = 3
С ФОРМИРОВАНИЕ КОДА ОШИБКИ
IERW(1) = KF
С
С-*-- ПРОВЕРКА 'K' ( 2-И РАЗРЯД )
С НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'K' : 2 <= K <= 99
KK = 0
С ЕСЛИ K >= 100 ЭТО ЛОГИЧЕСКАЯ ОШИБКА

```

```

      IF (K.GE.100) KK = 1
C   НЕПРАВИЛЬНОЕ ЗНАЧЕНИЕ 'K' : K < 2
      IF (K.LT.2) KK = 3
C   ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(2) = KK
C
C-*--      ПРОВЕРКА 'N' ( 3-И РАЗРЯД )
C   KN1=0 ОЗНАЧАЕТ, ЧТО 'N(1)' - НОРМАЛЬНОЕ, Т.Е. БОЛЬШЕ 9
      KN1 = 0
C   ЕСЛИ N(1)' МЕНЬШЕ ЧЕМ 10, ЭТО ЛОГИЧЕСКАЯ ОШИБКА
      IF (N(1).LT.10) KN1 = 1
C   ЕСЛИ N(1)' МЕНЬШЕ ЧЕМ 2, ЭТО ГРУБАЯ ОШИБКА
      IF (N(1).LT. 2) KN1 = 3
C   Ю ВЕРХНЕМУ ПРЕДЕЛУ ДЛЯ 'N1' ПРОВЕРКИ НЕТ
C
C   K.2 =0 ОЗНАЧАЕТ, ЧТО 'N(2)' - НОРМАЛЬНОЕ, Т.Е. БОЛЬШЕ 9
      KN2 = 0
C   ЕСЛИ 'N(2)' МЕНЬШЕ ЧЕМ 10, ЭТО ЛОГИЧЕСКАЯ ОШИБКА
      IF (N(2).LT.10) KN2 = 1
C   ЕСЛИ 'N(2)' МЕНЬШЕ ЧЕМ 2, ЭТО ГРУБАЯ ОШИБКА
      IF (N(2).LT. 2) KN2 = 3
C   ПО ВЕРХНЕМУ ПРЕДЕЛУ ДЛЯ 'N2' ПРОВЕРКИ НЕТ
C   ФОРМИРОВАНИЕ КОДА ОШИБКИ
      KN = MAX0(KN1,KN2)
      IERW(3) = KN
C
C-*--      ПРОВЕРКА 'P1'-МАССИВА ( 4-И РАЗРЯД )
      KPM1=11
C   ЕСЛИ 'K' - ПЛОХОЕ, ТО ПРОВЕРКА НЕВОЗМОЖНА
      IF (KK.GT.1) GO TO 50
C   НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'P1'
      KPM1 = 0
      S=0.0
      DO 45 I=1,K
        R=P1(I)
C   ОЧЕНЬ МАЛОЕ ЗНАЧЕНИЕ 'P1' ЭЛЕМЕНТА
      IF ( R.LT.CONST) KPM1 = 1
C   НЕПРАВИЛЬНОЕ ЗНАЧЕНИЕ 'P1' ЭЛЕМЕНТА
      IF ( R.GE.1.0.OR.R.EQ.0.0) KPM1 = 3
C   СУММИРОВАНИЕ ДОЛЕЙ 'P1'
      45 S= S + R
C   НЕПРАВИЛЬНАЯ СУММА ДОЛЕЙ 'P1'
      IF (S.GT.1.+CONST.OR.S.LT.1.-CONST) KPM1 = 3
C   ФОРМИРОВАНИЕ КОДА ОШИБКИ
      50 IERW(4) = KPM1
C
C-*--      ПРОВЕРКА 'P2'-МАССИВА ( 5-И РАЗРЯД )
      KPM2=11
C   ЕСЛИ 'K' - ПЛОХОЕ, ТО ПРОВЕРКА НЕВОЗМОЖНА
      IF (KK.GT.1) GO TO 60
C   НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'P2'
      KPM2 = 0
      S=0.0
      DO 55 I=1,K
        R=P2(I)
C   ОЧЕНЬ МАЛОЕ ЗНАЧЕНИЕ 'P2' ЭЛЕМЕНТА
      IF ( R.LT.CONST) KPM2 = 1
C   НЕПРАВИЛЬНОЕ ЗНАЧЕНИЕ 'P2' ЭЛЕМЕНТА
      IF ( R.GE.1.0.OR.R.EQ.0.0) KPM2 = 3
C   СУММИРОВАНИЕ ДОЛЕЙ 'P2'
      55 S= S + R
C   НЕПРАВИЛЬНАЯ СУММА ДОЛЕЙ 'P2'
      IF (S.GT.1.+CONST.OR.S.LT.1.-CONST) KPM2 = 3
C   ФОРМИРОВАНИЕ КОДА ОШИБКИ
      60 IERW(5) = KPM2
C

```



```

C--      ПРОВЕРКА 'Q' ( 6-И РАЗРЯД )
C      НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'Q' : 0.025 <= Q <= 0.2
      KQ = 0
C      ЕСЛИ 'Q' БОЛЬШЕ ЧЕМ 0.2 ИЛИ МЕНЬШЕ ЧЕМ 0.025,
      ТО СЧИТАЕМ ЧТО ДОПУЩЕНА ЛОГИЧЕСКАЯ ОШИБКА
      IF (Q.LT.0.025) KQ = 1
      IF (Q.GT.0.2) KQ = 1
C      ВВОДИМ ОГРАНИЧЕНИЕ : 0.005 <= Q <= 0.5
      IF (Q.LT.0.005) KQ = 3
      IF (Q.GT.0.5) KQ = 3
C      ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(6) = KQ
      IERW(8) = 6

C
C      ОПРЕДЕЛЕНИЕ МАКСИМАЛЬНОГО КОДА ОШИБКИ
C
      MWC = MAX0 (KK,KN,KPM1,KPM2,KQ)

C
C      ШКАЛА КОДИРОВКИ ОШИБОЧНЫХ ПАРАМЕТРОВ :
C
CE      РАЗРЯД КОДА ОШИБКИ  I  1  I  2  I  3  I  4  I  5  I  6
C
C      ПАРАМЕТР В ОБРАЩЕНИИ INFWR I  K  I  N  I  P1 I  P2 I  Q
C
C      ВОЗМОЖНЫЕ ЗНАЧЕНИЯ  I  0  I  0  I  0  I  0  I  0  I  0
C      I  -  I  1  I  1  I  1  I  1  I  1  I  1
C      I  -  I  -  I  -  I  -  I  -  I  2
C      I  3  I  3  I  3  I  3  I  3  I  3  I  -
C      I  -  I  -  I  -  I  11 I  11 I  -
C
C      ПЕЧАТЬ ВХОДНЫХ ДАННЫХ
C
      IF (KF.GT.Q) GO TO 999
      IF (KEY(2).EQ.0) GO TO 225

C
      WRITE (NFWR,220) NAME, MWC
220  FORMAT('/' ,A8,4X,'КОД МАКСИМАЛЬНОЙ ОШИБКИ =' ,I2)
C
225  IF (MWC.LT.1) GO TO 310
      J = IERW(8)
      WRITE (NFWR,227) (IERW(I),I=1,J)
227  FORMAT('/' ,8X,'ИМЕНА ПАРАМЕТРОВ : '
*      , ' NFWR, K, N, P1, P2, Q'
*      /' ,8X,'К О Д Ы ОШИБОК :',6(3X,I2)/)
310  IF (KEY(3).EQ.0) GO TO 710
      WRITE (NFWR,330) Q
330  FORMAT('/' ,8X,'УРОВЕНЬ ЗНАЧИМОСТИ РИСКА Q=' ,F8.3)
C
710  IF (KEY(7).EQ.0) GO TO 810
      WRITE (NFWR,740) (KSN(I,1),I=1,4),N(1)
*      , (KSN(I,2),I=1,4),N(2),K
740  FORMAT('/' ,8X,12X,'К О Д Ы ВЫБОРОК :',6X,
*      2X,' ОБЪЕМ : '
*      /' ,8X,'1 - ',4A8,2X,'N1 - ',I5
*      /' ,8X,'2 - ',4A8,2X,'N2 - ',I5
*      /' ,8X,'КОЛИЧЕСТВО ГРАДАЦИЙ - K =',I5)
C
810  IF (KEY(8).EQ.0) GO TO 999
      IF (KK.GT.1) GO TO 999
      WRITE (NFWR,850) P1
850  FORMAT('/' ,8X,'МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ 1-И ВЫБОРКИ' /
*      ,(' ',12X,4F11.5))
C
      WRITE (NFWR,860) P2
860  FORMAT('/' ,8X,'МАССИВ ЗНАЧЕНИЙ ДОЛЕЙ 2-И ВЫБОРКИ' /
*      ,(' ',12X,4F11.5))

C
999  RETURN
CV;  ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
      END

```

```

CC; PR220 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP220
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫВОДА В ФАЙЛ 'NFWR'
C; РЕЗУЛЬТАТОВ РАБОТЫ П/П SP220
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP220
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE PR220(KEY,NFWR,KSQN,K,L,R,D,TQV1,TQV2,IND,IER)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА/1<=NFWR<=99/
C K SQN R*8(4,2) 4A8 7 МАССИВ К О Д О В ВЫБОРОК
C K I*4 1 -- - КОЛИЧЕСТВО ГРАДАЦИИ
C L I*4 (K) A 5 МАССИВ ВЫЧ. ИНДЕКС. ПО СООТВ. ДОЛЯМ
C R R*4 1 G20.5 3 МАКСИМАЛЬНАЯ РАЗНОСТЬ СООТВ. ДОЛЕЙ
C D R*4 1 G20.5 3 СУММА ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ
C TQV1 R*4 1 G20.5 3 КРИТ. ЗНАЧ. Т-РАСПРЕД. ДЛЯ 1-И ВЫБ.
C TQV2 R*4 1 G20.5 3 КРИТ. ЗНАЧ. Т-РАСПРЕД. ДЛЯ 2-И ВЫБ.
C IND I*4 1 I1 2 ИНДЕКС ПРОВЕРКИ ПО ИНТЕРВ. ОЦЕНКИ.
C IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
C IND = 1 - НУЛЕВАЯ ГИПОТЕЗА ОТВЕРГАЕТСЯ.
C IER I*4 1 I3 2 КОД ОШИБКИ, ВЫРАБОТАННЫЙ SP220
C
C DIMENSION KEY(8),L(K),K SQN(4,2)
C
C REAL*8 NAME,KSQN
C REAL*4 T1(6),T2(2),T3(3),T5(3),P(15)
C REAL*8 TNK(3),TS1(4),TS2(4),TS3(4)
C
C EQUIVALENCE ( P(1), T1(1))
C * , ( P(7), T2(1))
C * , ( P(9), T3(1))
C * , (P(12), T4)
C * , (P(13), T5(1))
C
C DATA P / '* ДО', 'ВЕРИ', 'ТЕЛЬ', 'НЫЕ', 'ГРАН'
C * , 'ИЦЫ', '-Х Д', 'ОЛЕЙ', 'СОВ', 'ПАДА'
C * , 'ЮТ', 'НЕ', 'ГЕН', 'ВЫБ', 'ОРОК' /
C DATA TNK / 'ИНТЕРВАЛ', 'ЬНОЕ ОЦ', 'ЕНИВАНИЕ' /
C * , TS1 / 'ОЦЕНКА', 'РАЗЛИЧИЯ', '
C * , TS2 / 'МЕЖДУ Д', 'ОЛЯМИ 2-', 'Х ВЫБОРО', 'К ПУТЕМ' /
C * , TS3 / 'СОПОСТА', 'ВЛЕНИЯ', 'ДОВ. ИНТ', 'ЕРВАЛОВ' /
C DATA NAME / ' PR220' /
C
C KLU=KEY(1)
C IF(KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 11 KLU=IABS(KLU)
C DO 13 J=1,8
C I=9-J
C KEY(I)=MOD(KLU,2)
C 13 KLU=KLU/2
C GO TO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 15 DO 17 I=1,8

```

```

        KEY(I)=1
        IF (KLU.LT.I) KEY(I)=0
17      CONTINUE
19      CONTINUE
        IF (KEY(1).EQ.0) GO TO 210
C
        IF (NFWR.LE.0.OR.NFWR.GE.100) GO TO 999
C
        WRITE (NFWR,120) TS1,TNK,TS2,TS3
120     FORMAT(/ 1H ,26(1H*))
        *           /2H *,24(1H ),3H* ,4A8
        *           /2H *,3A8 ,3H* ,4A8
        *           /2H *,24(1H ),3H* ,4A8
        *           / 1H ,26(1H*) )
210     IF (KEY(2).EQ.0) GO TO 310
C
        WRITE (NFWR,230) NAME, IND, IER
230     FORMAT(/' ',A8,'СПРАВЕДЛИВА ГИПОТЕЗА Н',I1
        *           ' ; КОД ОШИБКИ = ',I3)
C
310     IF (KEY(3).EQ.0) GO TO 510
        WRITE (NFWR,340) R,D,TQV1,TQV2
340     FORMAT(/' ',8X, 'МАКСИМАЛЬНАЯ РАЗНОСТЬ ДОЛЕЙ ',G20.5
        *           /' ',8X, 'СУММА ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ',G20.5
        *           /' ',8X,'КРИТ. ЗНАЧ. Т-РАСПР. ДЛЯ 1-И ВЫБ',G20.5
        *           /' ',8X,'КРИТ. ЗНАЧ. Т-РАСПР. ДЛЯ 2-И ВЫБ',G20.5)
C
510     IF (KEY(5).EQ.0) GO TO 710
C
        DO 550 I=1,K
        IF (L(I).EQ.1) GO TO 545
        WRITE (NFWR,900) T1,I,T2,T3
        GO TO 550
545     CONTINUE
        WRITE (NFWR,900) T1,I,T2,T4,T3
550     CONTINUE
        IF (IND.EQ.1) GO TO 555
        WRITE (NFWR,910) T1,T5,T3
        GO TO 560
555     CONTINUE
        WRITE (NFWR,910) T1,T5,T4,T3
560     CONTINUE
C
710     IF (KEY(7).EQ.0) GO TO 999
        WRITE (NFWR,770) (KSN(I,1),I=1,4),(KSN(I,2),I=1,4)
770     FORMAT(/' ',8X,12X,'К О Д Ы В Ы Б О Р О К : '
        *           /' ',8X,'1 - ',4A8
        *           /' ',8X,'2 - ',4A8)
C
900     FORMAT(' ',8X,6A4,I4,6A4)
910     FORMAT(' ',8X,13A4)
C
999     RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
END

```

§ 3. Оценка различия « m » независимых выборок

Используется метод дисперсионного однофакторного анализа при числе градаций $k = 2$.

3.1. Дисперсионный однофакторный анализ. Предварительно вычисляются суммы квадратов отклонений Q и числа степеней свободы ν общая

$$Q_{\text{общ}} = \sum_{i=1}^m n_i - \frac{\left(\sum_{i=1}^m n_i\right)^2}{\sum_{i=1}^m N_i}; \quad \nu_{\text{общ}} = \sum_{i=1}^m N_i - 1;$$

межвыборочная

$$Q_{\text{меж}} = \sum_{i=1}^m n_i p_i - \frac{\left(\sum_{i=1}^m n_i\right)^2}{\sum_{i=1}^m N_i}; \quad \nu_{\text{меж}} = m - 1;$$

внутривыборочная

$$Q_{\text{вн}} = Q_{\text{общ}} - Q_{\text{меж}}; \quad \nu_{\text{вн}} = \nu_{\text{общ}} - \nu_{\text{меж}} = \sum_{i=1}^m N_i - m,$$

где n_i — число объектов в i -й выборке, обладающих данным признаком; N_i — объем i -й выборки; $p_i = n_i/N_i$ — доля признака; m — число выборок (вариантов).

Дисперсия

$$D = Q/\nu.$$

Если выборки взяты из одной генеральной совокупности (нет различий), то все три суммы дисперсии будут одинаковыми. Для оценки существенности различия между выборками сравниваются $D_{\text{меж}}$ и $D_{\text{вн}}$ по критерию Фишера, для чего вычисляется величина F ,

$$F = D_{\text{меж}}/D_{\text{вн}},$$

и сравнивается с критическим значением $F_{q,\nu}$.

H_0 не отвергается при $F \leq F_{q,\nu}$ (отвергается при $F \geq F_{q,\nu}$). Критическое значение $F_{q,\nu}$ вычисляется при уровне значимости риска q и степенях свободы $\nu_{\text{вн}}$ и $\nu_{\text{меж}}$.

Если $F \geq F_{q,\nu}$, то производится оценка попарных значений между отдельными двумя выборками по критерию Стьюдента $t = \frac{p_1 - p_2}{S_d}$, где

S_d — ошибка разности сравниваемых долей, $S_d = S \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$; S — показатель изменчивости (при дисперсионном анализе полагают внутривыборочные дисперсии одинаковыми), $S = \sqrt{D_{\text{вн}}}$; p_1, p_2 — соответствующие доли двух вариантов опыта (выборок); N_1, N_2 — объемы выборок.

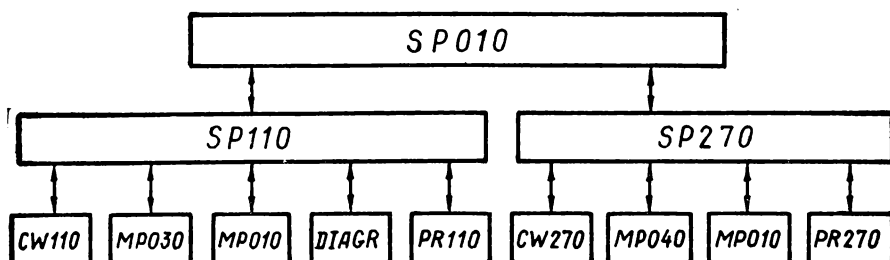


Рис. 55

$H_0 p_1 = p_2$ отвергается при $t > t_{q,v}$, $v = N_1 + N_2 - 2$ и уровне значимости риска q .

3.2. Программная реализация. Блок-схема программной реализации (+ $SP27\theta$ — функциональный модуль) представлена на рис. 55.

Подпрограмма $MP\theta4\theta$ вычисляет критические значения распределения F , подпрограммы $CW27\theta$ и $PR27\theta$ осуществляют проверку и печать входных данных и соответственно печать результатов работы. Назначение остальных подпрограмм приведено в п. 1.2.

Тексты подпрограмм $SP27\theta$, $CW27\theta$, $PR27\theta$ приводятся после примеров.

3.3. Пример. В качестве примера проведем обработку методом одноклассового дисперсионного анализа групп из шести модельных выборок, каждая из которых имеет по две градации качественного признака, значения которых, а также элементы выборок кодированы числами 1 и 2.

При подготовке «потока» (исходного файла задания) для обработки комплексом программ «СОКИ-Н» необходимо задать следующие параметры: $IRP = 003$, $NS = 06$, $K = 2$, а также длины и коды выборок, уровни градаций, уровень значимости риска, ключи печати, высоту диаграммы.

В приведенном ниже задании исходных данных сами выборки, а также форматы ввода вынесены во «вторичный файл», что задано кодом «110» в строке с ключом « $FILX$ ». Далее в этой строке задан номер канала ввода — «8» — и расширенное имя файла с данными « $MX3 : NDATX6DAT$ ».

В данном случае все выборки имеют один и тот же формат записи данных, что совсем не обязательно. Выборки помещены в один и тот же файл, что также не обязательно.

Ниже приведен текст задания на обработку, распечатка файла данных и результат выполнения задания (распечатка). Следует обратить внимание на то, что подпрограмма $SP11\theta$ статистической обработки выборки вызывается шесть раз по числу выборок для вычисления долей градаций исследуемого признака.

Если же доли градаций известны, то можно провести дисперсионный анализ сразу, без предварительной обработки выборок, что иллюстрируется примером второго задания на обработку тех же шести выборок, по выходным данным — долям двух градаций — качественного признака.

Таблица долей двух градаций для шести выборок задана в строке, следующей за строкой с ключевым словом «FORZ». Задание длин выборок необходимо и в этом случае. Режим обработки по входным данным — долям градаций определяется заданием $IRP = 103$.

```

ЗАДАНИЕ НА ОЦЕНКУ РАЗЛИЧИЯ "М" НЕЗАВИСИМЫХ ВЫБОРОЧНЫХ
СОВОКУПНОСТЕЙ МЕТОДОМ ДИСПЕРСИОННОГО ОДНОФАКТОР-
НОГО АНАЛИЗА. ДАННЫЕ - ВЫБОРКИ В ФАЙЛЕ

----- ОЦЕНКА РАЗЛИЧИЯ "М" НЕЗАВИСИМЫХ ВЫБОРОК : 110
**** МАКЕТ ПОДГОТОВКИ ДАННЫХ ДЛЯ IR=003: ****SP270****
IRP -003 (I3)- КОД ФУНКЦИИ
NS -06 (I2)- КОЛИЧЕСТВО ВЫБОРОК
K -02 (I2)- КОЛИЧЕСТВО ГРАДАЦИИ
M -01 (I2)- НОМЕР ИССЛЕДУЕМОЙ ГРАДАЦИИ
      (I2)-НОМЕР; (4A8)- К О Д ВЫБОРКИ; (I4)-ДЛИНА
      XX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXX
KSGN-01 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 11 0080
KSGN-02 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 12 0085
KSGN-03 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 13 0078
KSGN-04 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 14 0087
KSGN-05 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 15 0070
KSGN-06 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 16 0082
      (I2, I3,I2)[ 4A8 ]:НОМЕР ВЫБ.,КОД,КАНАЛ,ИМЯ ФАЙЛА
FILX-01 110 08-INF:NDATX6.DAT
FILX-02 110 08-INF:NDATX6.DAT
FILX-03 110 08-INF:NDATX6.DAT
FILX-04 110 08-INF:NDATX6.DAT
FILX-05 110 08-INF:NDATX6.DAT
FILX-06 110 08-INF:NDATX6.DAT
      (I2, I3,I2)[ 4A8 ]:НОМЕР ВЫБ.,КОД,КАНАЛ,ИМЯ ФАЙЛА
FILC-01 000 08-INF:DATA6.DAT
      [ СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) ]
FORC-01 (4X,9I2)
      1 2
Q - 0.05 <- (G14.7)- УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
NFWR-02 (I2) - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
      (I1,I2)-ТИП КЛЮЧЕЙ ПЕЧАТИ: 0-ПОРАЗРЯДНЫЕ; И ИХ КОЛИЧЕСТВО
KEYC-0 04 1-ИНТЕРВАЛЬНЫЕ.
      ИНДЕКСЫ/ИНТЕРВАЛЬНОЕ ЗНАЧЕНИЕ КЛЮЧА ПЕЧАТИ/ПОРАЗРЯДНОЕ/
      12345678 - РАЗРЯДЫ КЛЮЧА ПЕЧАТИ :
KEY - 1 1 1/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SP110 ;
KEY - 2 1 1/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP110 ;
KEY - 1 7 2/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP270 ;
KEY - 2 7 2/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP270 .
MD -20 (I2)- МАКСИМАЛЬНАЯ ВЫСОТА ДИАГРАММЫ
      (I1)-РАСПРЕДЕЛЕНИЕ ПАМЯТИ:=1-ВЫДАЕТСЯ ВСЕГДА;
END -1 КОНЕЦ ДАННЫХ: =0-ВЫДАЕТСЯ ТОЛЬКО АВАРИЙНО.

```

ФАЙЛ ДАННЫХ К ЗАДАНИЮ, СОДЕРЖАЩИЙ ФОРМАТЫ ВВОДА, ЗНАЧЕНИЯ
ВЫБОРОК И УРОВНИ ГРАДАЦИИ

```

[ СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) ]
FORX-01 (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
FORX-02 (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
FORX-03 (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
FORX-04 (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
FORX-05 (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
FORX-06 (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
[ СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) ]
FORC-01 (4X,9I2)
1 2

```

РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ ЗАДАНИЯ (РАСПЕЧАТКА РЕШЕНИЯ):

```

IRP - 3
NS - 6
K - 2
M - 1
KSNQ NN XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX NNNN
KSNQ- 1 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 11 80
KSNQ- 2 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 12 85
KSNQ- 3 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 13 78
KSNQ- 4 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 14 87
KSNQ- 5 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 15 70
KSNQ- 6 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 16 82
FILX- 1 110 8-INF:NDATX6.DAT
FORX- 1: (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
FILX- 2 110 8-INF:NDATX6.DAT
FORX- 2: (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
FILX- 3 110 8-INF:NDATX6.DAT
FORX- 3: (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
FILX- 4 110 8-INF:NDATX6.DAT
FORX- 4: (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
FILX- 5 110 8-INF:NDATX6.DAT
FORX- 5: (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 2 2 2 2 2
FILX- 6 110 8-INF:NDATX6.DAT
FORX- 6: (4X,30I2)
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
FILC- 1 0 8-INF:DATA6.DAT
FORC- 1: (4X,9I2)
1 2
Q 0.5000000E-01
NFWR 2
KEYC 0 4 4
KEY 1 1 1/8/11111111/
KEY 2 1 1/8/11111111/
KEY 1 7 2/8/11111111/
KEY 2 7 2/8/11111111/
MD 20
ОБЪЕМ РАБОЧЕЙ ПАМЯТИ - 4000 ЕДИНИЦ INTEGER
ТРЕБУЕМЫЙ ОБЪЕМ ПАМЯТИ - 1102 ЕДИНИЦ INTEGER

```

ИНДЕКСЫ НАЧАЛА ПЕРЕМЕННЫХ В РАБОЧЕЙ ПАМЯТИ :

| | | | | | | | |
|-----|---|------|-----|------|-----|------|------|
| 1 - | 1 | 6 - | 6 | 11 - | 175 | 16 - | 362 |
| 2 - | 2 | 7 - | 7 | 12 - | 657 | 17 - | 374 |
| 3 - | 3 | 8 - | 36 | 13 - | 659 | 18 - | 1040 |
| 4 - | 4 | 9 - | 19 | 14 - | 360 | 19 - | 1047 |
| 5 - | 5 | 10 - | 169 | 15 - | 361 | 20 - | 1054 |

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ВЫБОРКИ :

ОБЪЕМ ВЫБОРКИ - 85
 ЧИСЛО ГРАДАЦИЙ - 2
 ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ - 0.47000
 КОЭФФИЦИЕНТ ВАРЬИРОВАНИЯ - 0.94000
 КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ - 0.05098
 ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ (D) - 0.10738
 МОДА - 0.67059
 КРИТЕРИЙ ХИ-КВАДРАТ : СПРАВЕДЛИВА ГИПОТЕЗА H1

| NN | УРОВНИ ГРАДАЦИЙ | ДОЛЯ-D | ДОЛЯ | ДОЛЯ+D | ПЕРЕКРЫТИЕ |
|----|-----------------|---------|---------|---------|------------|
| 1 | 1.000 | 0.56321 | 0.67059 | 0.77797 | * |
| 2 | 2.000 | 0.22203 | 0.32941 | 0.43679 | * |

 * * * * * ВЫЧИСЛЕНИЕ
 * СТАТ.ОБРАБОТКА ВЫБОРКИ * СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК
 * * * * * ОДНОЙ ВЫБОРОЧНОЙ СОВОКУПНОСТИ

PR110 - СПРАВЕДЛИВА ГИПОТЕЗА H1 ; КОД ОШИБКИ = 0

ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ = 0.47000
 КОЭФФИЦИЕНТ ВАРЬИРОВАНИЯ = 0.94000
 КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ = 0.50979E-01
 МОДА = 0.67059
 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ H12 = 9.8941
 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ H12 = 3.8415

К О Д ВЫБОРКИ :
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 12

МАССИВ ЗНАЧЕНИЯ ВЫЧИСЛЕННЫХ ДОЛЕЙ:
 0.67059 0.32941

SW110 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА α = 0.050

К О Д ВЫБОРКИ :

ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 13

КОЛИЧЕСТВО ГРАДАЦИЙ - K = 2

ЗНАЧЕНИЯ ЭЛ. ВЫБОРОЧНОЙ СОВОКУПНОСТИ

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |

ОБЪЕМ :
 78

МАССИВ ЗНАЧЕНИЯ УРОВНЕЙ:

1 2

K= 2

[illegible]

К О Д В Ы Б О Р К И : Т Е С Т О В А Я / 2 Г Р А Д А Ц И И / В Ы Б О Р К А 13

| | | |
|---|---|---------|
| ОБЪЕМ ВЫБОРКИ | - | 78 |
| ЧИСЛО ГРАДАЦИЙ | - | 2 |
| ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ | - | 0.49852 |
| КОЭФФИЦИЕНТ ВАРИИРОВАНИЯ | - | 0.99704 |
| КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ | - | 0.05643 |
| ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ (D) | - | 0.11927 |
| КРИТЕРИЙ | - | 0.53846 |
| КРИТЕРИЙ ХИ-КВАДРАТ : СПРАВЕДЛИВА ГИПОТЕЗА Н0 | | |

| №№ | УРОВНИ ГРАДАЦИЯ | ДОЛЯ-D | ДОЛЯ | ДОЛЯ+D | ПЕРЕКРЫТИЕ |
|----|--------------------|---------|---------|---------|------------|
| 1 | 1.000 | 0.41919 | 0.53846 | 0.65773 | #2 |
| 2 | 2.000 | 0.34227 | 0.46154 | 0.58081 | 1* |

```
*****
*                                     *
*  СТАТ.ОБРАБОТКА ВЫБОРКИ          *
*                                     *
*****
                                     ВЫЧИСЛЕНИЕ
СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК
ОДНОЙ ВЫБОРОЧНОЙ СОВОКУПНОСТИ
```

PR110 - СПРАВЕДЛИВА ГИПОТЕЗА Н0 ; КОД ОШИБКИ = 0

| | | |
|----------------------------|---|-------------|
| ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ | = | 0.49852 |
| КОЭФФИЦИЕНТ ВАРЬИРОВАНИЯ | = | 0.99704 |
| КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ | = | 0.56446E-01 |
| МОДА | = | 0.53846 |
| ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ H12 | = | 0.46154 |
| КРИТИЧЕСКОЕ ЗНАЧЕНИЕ H12 | = | 3.8415 |

К О Д В Ы Б О Р К И :
ТЕСТОВАЯ /2 ГРАДАЦИИ/ В Ы Б О Р К А 13

МАССИВ ЗНАЧЕНИЙ ВЫЧИСЛЕННЫХ ДОЛЕЙ: 0.53846 0.46154

SW110 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА $Q = 0.050$

К О Д В Ы Б О Р К И :
ТЕСТОВАЯ /2 ГРАДАЦИИ/ В Ы Б О Р К А 14
КОЛИЧЕСТВО ГРАДАЦИЙ - К = 2

06"EM 87

ЗНАЧЕНИЯ ЭЛ. ВЫБОРОЧНОЙ СОВОКУПНОСТИ

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | | |

МАССИВ ЗНАЧЕНИЙ УРОВНЕЙ:

1 2

DIAGR K= 2

[illegible]

ДИАГРАММА РАСПРЕДЕЛЕНИЯ ДОЛЕЙ В ВЫБОРКЕ

К О Д В Ы Б О Р К И : Т Е С Т О В А Я / 2 Г Р А Д А Ц И И / В Ы Б О Р К А 14

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ВЫБОРКИ :

| | | |
|----------------------------|-----------------|---------|
| ОБЪЕМ ВЫБОРКИ | - | 87 |
| ЧИСЛО ГРАДИАН | - | 2 |
| ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ | - | 0.35651 |
| КОЭФФИЦИЕНТ ВАРИИРОВАНИЯ | - | 0.71301 |
| КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ | - | 0.03822 |
| ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ (D) | - | 0.08044 |
| МОДА | - | 0.85057 |
| КРИТЕРИЙ ХИ-КВАДРАТ | СПРАВЕДЛИВА ГИП | |

| № | УРОВНИ ГРАДАЦИИ | ДОЛЯ-D | ДОЛЯ | ДОЛЯ+D | ПЕРЕКРЫТИЕ |
|---|--------------------|---------|---------|---------|------------|
| 1 | 1.000 | 0.77013 | 0.85057 | 0.93102 | 8 |
| 2 | 2.000 | 0.06898 | 0.14943 | 0.22987 | 8 |

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ВЫБОРКИ :

ОБЪЕМ ВЫБОРКИ - 70
 ЧИСЛО ГРАДАЦИЙ - 2
 ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ - 0.23212
 КОЭФФИЦИЕНТ ВАРЬИРОВАНИЯ - 0.46423
 КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ - 0.02774
 ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ (D) - 0.05885
 МОДА - 0.94286
 КРИТЕРИЙ ХИ-КВАДРАТ : СПРАВЕДЛИВА ГИПОТЕЗА H1

| NN | УРОВНИ ГРАДАЦИИ | ДОЛЯ-D | ДОЛЯ | ДОЛЯ+D | ПЕРЕКРЫТИЕ |
|----|-----------------|----------|---------|---------|------------|
| 1 | 1.000 | 0.88401 | 0.94286 | 1.00170 | * |
| 2 | 2.000 | -0.00170 | 0.05714 | 0.11599 | * |

 * ВYЧИСЛЕНИЕ
 * СТАТ.ОБРАБОТКА ВЫБОРКИ * СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК
 * * ОДНОЙ ВЫБОРОЧНОЙ СОВОКУПНОСТИ

PR110 - СПРАВЕДЛИВА ГИПОТЕЗА H1 ; КОД ОШИБКИ = 0

ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ = 0.23212
 КОЭФФИЦИЕНТ ВАРЬИРОВАНИЯ = 0.46423
 КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ = 0.27743E-01
 МОДА = 0.94286
 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ H12 = 54.914
 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ H12 = 3.8415

К О Д ВЫБОРКИ :
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 15

МАССИВ ЗНАЧЕНИЙ ВЫЧИСЛЕННЫХ ДОЛЕЙ:
 0.94286 0.05714

SN110 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА α = 0.050
 К О Д ВЫБОРКИ : ОБЪЕМ :
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 16 82
 КОЛИЧЕСТВО ГРАДАЦИЙ - K = 2

ЗНАЧЕНИЯ ЭЛ. ВЫБОРОЧНОЙ СОВОКУПНОСТИ

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |

МАССИВ ЗНАЧЕНИЙ УРОВНЕЙ:
 1 2

K= 2

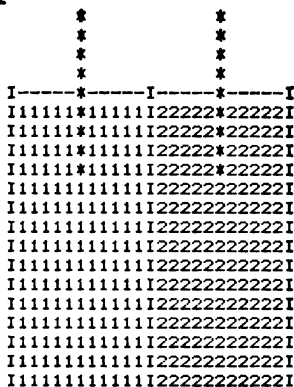


ДИАГРАММА РАСПРЕДЕЛЕНИЯ ДОЛЕЙ В ВЫБОРКЕ

К О Д В Ы Б О Р К И : Т Е С Т О В А Я / 2 Г Р А Д А Ц И И / В Ы Б О Р К А 1 6

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ВЫБОРКИ

| | | |
|---|---|---------|
| ОБЪЕМ ВЫБОРКИ | - | 82 |
| ЧИСЛО ГРАДАЦИЙ | - | 2 |
| ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ | - | 0.50000 |
| КОЭФФИЦИЕНТ ВАРИИРОВАНИЯ | - | 1.00000 |
| КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ | - | 0.05522 |
| ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ (D) | - | 0.11646 |
| МОДА | - | 0.50000 |
| КРИТЕРИЙ ХИ-КВАДРАТ : СПРАВЕДЛИВА ГИПОТЕЗА Н0 | | |

| №№ | УРОВНИ ГРАДАЦИЯ | ДОЛЯ-D | ДОЛЯ | ДОЛЯ+D | ПЕРЕКРЫТИЕ |
|----|--------------------|---------|---------|---------|------------|
| 1 | 1.000 | 0.38354 | 0.50000 | 0.61646 | *2 |
| 2 | 2.000 | 0.38354 | 0.50000 | 0.61646 | 1* |

```

*****
*                                     *
*      СТАТ.ОБРАБОТКА ВЫБОРКИ      *
*                                     *
*****

```

PR110 - СПРАВЕДЛИВА ГИПОТЕЗА Н0 ; КОД ОШИБКИ = 0

| | | |
|---------------------------|---|-------------|
| ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ | = | 0.50000 |
| КОЭФФИЦИЕНТ ВАРИИРОВАНИЯ | = | 1.0000 |
| КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛ | = | 0.55216E-01 |
| МОДА | = | 0.50000 |
| ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ H12 | = | 0.00000 |
| КРИТИЧЕСКОЕ ЗНАЧЕНИЕ H12 | = | 3.8413 |

К О Д В Ы Б О Р К И :
ТЕСТОВАЯ /2 ГРАДАЦИИ/ В Ы Б О Р К А 16

МАССИВ ЗНАЧЕНИЙ ВЫЧИСЛЕННЫХ ДОЛЕЙ:
0.50000 0.50000

СН270 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА = 0.050

ТАБЛИЦА СОСТАВЛЕНА ИЗ ДОЛЕЙ 1-й ГРАДАЦИИ 6 ВЫБОРОК:

| NN | КОД | ВЫБОРКИ | ДЛИНА | ДОЛЯ |
|----|-------------------------|------------|-------|---------|
| 1 | - ТЕСТОВАЯ /2 ГРАДАЦИИ/ | ВЫБОРКА 11 | 80 | 0.80000 |
| 2 | - ТЕСТОВАЯ /2 ГРАДАЦИИ/ | ВЫБОРКА 12 | 85 | 0.67059 |
| 3 | - ТЕСТОВАЯ /2 ГРАДАЦИИ/ | ВЫБОРКА 13 | 78 | 0.53846 |
| 4 | - ТЕСТОВАЯ /2 ГРАДАЦИИ/ | ВЫБОРКА 14 | 87 | 0.85057 |
| 5 | - ТЕСТОВАЯ /2 ГРАДАЦИИ/ | ВЫБОРКА 15 | 70 | 0.94286 |
| 6 | - ТЕСТОВАЯ /2 ГРАДАЦИИ/ | ВЫБОРКА 16 | 82 | 0.50000 |

 * ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ
 * ДИСПЕРСИОННЫЙ АНАЛИЗ * АНАЛИЗ
 * ГРУППЫ "К" НЕЗАВИСИМЫХ ВЫБОРОК

 PR270 - СПРАВЕДЛИВА ГИПОТЕЗА H1 КОД ОШИБКИ = 0

ТАБЛИЦА СОСТАВЛЕНА ИЗ ДОЛЕЙ 1-й ГРАДАЦИИ 6 ВЫБОРОК:

ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ КРИТЕРИЯ ФИШЕРА 13.516
 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ F-РАСПРЕДЕЛЕНИЯ = 2.2319

РЕЗУЛЬТАТЫ ПОПАРНОГО СРАВНЕНИЯ ПО Т-КРИТЕРИЮ:

| ВЫБОРКИ | ИНДЕКС-Т | ВЫЧ.ЗНАЧ. | КРИТ.ЗНАЧ. | ИНДЕКС-D |
|---------|----------|-----------|------------|----------|
| 1 2 | 0 | 1.95533 | 1.98061 | 1 |
| 1 3 | 1 | 3.86842 | 1.98077 | 0 |
| 1 4 | 0 | 0.76845 | 1.98057 | 1 |
| 1 5 | 1 | 2.05440 | 1.98094 | 1 |
| 1 6 | 1 | 4.49314 | 1.98068 | 0 |
| 2 3 | 1 | 1.98330 | 1.98066 | 1 |
| 2 4 | 1 | 2.77766 | 1.98046 | 1 |
| 2 5 | 1 | 3.97031 | 1.98083 | 0 |
| 2 6 | 1 | 2.59383 | 1.98057 | 1 |
| 3 4 | 1 | 4.71098 | 1.98061 | 0 |
| 3 5 | 1 | 5.78104 | 1.98099 | 0 |
| 3 6 | 0 | 0.57234 | 1.98072 | 1 |
| 4 5 | 0 | 1.35273 | 1.98079 | 1 |
| 4 6 | 1 | 5.36090 | 1.98053 | 0 |
| 5 6 | 1 | 6.40518 | 1.98090 | 0 |

РЕЗУЛЬТАТЫ ПОПАРНОГО СРАВНЕНИЯ ПО КРИТЕРИЮ D:

| ВЫБОРКИ | ИНДЕКС-D | СУММА | РАЗНОСТЬ | КРИТИЧЕСКИЕ ЗНАЧЕНИЯ |
|---------|----------|---------|----------|----------------------|
| 1 2 | 1 | 0.20179 | 0.12941 | 2.11100 2.10640 |
| 1 3 | 0 | 0.21368 | 0.26154 | 2.11100 2.11300 |
| 1 4 | 1 | 0.17485 | 0.05057 | 2.11100 2.10460 |
| 1 5 | 1 | 0.15325 | 0.14286 | 2.11100 2.12110 |
| 1 6 | 0 | 0.21086 | 0.30000 | 2.11100 2.10910 |
| 2 3 | 1 | 0.22665 | 0.13213 | 2.10640 2.11300 |
| 2 4 | 1 | 0.18782 | 0.17999 | 2.10640 2.10460 |
| 2 5 | 0 | 0.16623 | 0.27227 | 2.10640 2.12110 |
| 2 6 | 1 | 0.22384 | 0.17059 | 2.10640 2.10910 |
| 3 4 | 0 | 0.19971 | 0.31211 | 2.11300 2.10460 |
| 3 5 | 0 | 0.17812 | 0.40440 | 2.11300 2.12110 |
| 3 6 | 1 | 0.23573 | 0.03846 | 2.11300 2.10910 |
| 4 5 | 1 | 0.13929 | 0.09228 | 2.10460 2.12110 |
| 4 6 | 0 | 0.19690 | 0.35057 | 2.10460 2.10910 |
| 5 6 | 0 | 0.17530 | 0.44286 | 2.12110 2.10910 |

**ЗАДАНИЕ НА ОЦЕНКУ РАЗЛИЧИЯ "М" НЕЗАВИСИМЫХ ВЫБОРОЧНЫХ
СОВОКУПНОСТЕЙ - ДИСПЕРСИОННЫЙ ОДНОФАКТОРНЫЙ АНАЛИЗ
ДАННЫЕ - ДОЛИ ДВУХ ГРАДАЦИЙ КАЧЕСТВЕННОГО ПРИЗНАКА**

```

----- ОЦЕНКА РАЗЛИЧИЯ "М" НЕЗАВИСИМЫХ ВЫБОРОК :
**** МАКЕТ ПОДГОТОВКИ ДАННЫХ ДЛЯ IR=103: ****SP270****
IRP -103 (I3)- КОД ВЫПОЛНЯЕМОЙ ФУНКЦИИ
NS -06 (I2)- КОЛИЧЕСТВО ВЫБОРОК
K -02 (I2)- КОЛИЧЕСТВО ГРАДАЦИЙ
M -01 (I2)- НОМЕР ИССЛЕДУЕМОЙ ГРАДАЦИИ
(I2)-НОМЕР; (4A8)- К О Д ВЫБОРКИ; (I4)-ДЛИНА
XX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXX
KSNQ-01 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 31 0080
KSNQ-02 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 32 0085
KSNQ-03 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 33 0078
KSNQ-04 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 34 0087
KSNQ-05 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 35 0070
KSNQ-06 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 36 0082
(I2, I3,I2)[ 4A8 ]:НОМЕР ВЫБ.,КОД,КАНАЛ,ИМЯ ФАЙЛА
FILZ-01 000 08-INF:DATAPZ.DAT
[ СПИСОК СПЕЦИФИКАЦИЙ ОПЕРАТОРА F O R M A T (14A4) ]
FORZ-01 (4X,2F8.5)
0.80000 0.20000
0.67000 0.33000
0.54000 0.46000
0.85000 0.15000
0.94000 0.06000
0.50000 0.50000
Q - 0.05 <- (G14.7)- УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
NFWR-02 (I2) - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
(I1,I2)-ТИП КЛЮЧЕЙ ПЕЧАТИ: 0-ПОРАЗРЯДНЫЕ; И ИХ КОЛИЧЕСТВО
KEYC-0 02 1-ИНТЕРВАЛЬНЫЕ.
ИНДЕКСЫ/ИНТЕРВАЛЬНОЕ ЗНАЧЕНИЕ КЛЮЧА ПЕЧАТИ/ПОРАЗРЯДНОЕ/
12345678 - РАЗРЯДЫ КЛЮЧА ПЕЧАТИ :
KEY - 1 7 2/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SP270 ;
KEY - 2 7 2/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP270 .
MD -20 (I2)- МАКСИМАЛЬНАЯ ВЫСОТА ДИАГРАММЫ
(I1)-РАСПРЕДЕЛЕНИЕ ПАМЯТИ:=1-ВЫДАЕТСЯ ВСЕГДА;
END -1 КОНЕЦ ДАННЫХ: =0-ВЫДАЕТСЯ ТОЛЬКО АВАРИЙНО.

```

РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ ЗАДАНИЯ (РАСПЕЧАТКА РЕШЕНИЯ)

```

IRP -103
NS - 6
K - 2
M - 1
KSN NN XXXXXXXXXXXXXXXXXXXXXXXXXXXX NNNN
KSN- 1 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 31 80
KSN- 2 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 32 85
KSN- 3 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 33 78
KSN- 4 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 34 87
KSN- 5 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 35 70
KSN- 6 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 36 82
FILZ- 1 0 8-INF:DATAPZ.DAT
FORZ- 1: (4X,2F8.5)
0.80000 0.20000
0.67000 0.33000
0.54000 0.46000
0.85000 0.15000
0.94000 0.06000
0.50000 0.50000
Q 0.5000000E-01
NFWR 2
MEYC 0 2 2
KEY 1 7 2/8/11111111/
KEY 2 7 2/8/11111111/
MD 20
ОБЪЕМ РАБОЧЕЙ ПАМЯТИ - 4000 ЕДИНИЦ INTEGER
ТРЕБУЕМЫЙ ОБЪЕМ ПАМЯТИ - 642 ЕДИНИЦ INTEGER

```

ИНДЕКСЫ НАЧАЛА ПЕРЕМЕННЫХ В РАБОЧЕЙ ПАМЯТИ :

| | | | | | | | |
|-----|---|------|-----|------|-----|------|-----|
| 1 - | 1 | 6 - | 6 | 11 - | 175 | 16 - | 132 |
| 2 - | 2 | 7 - | 7 | 12 - | 176 | 17 - | 144 |
| 3 - | 3 | 8 - | 36 | 13 - | 177 | 18 - | 580 |
| 4 - | 4 | 9 - | 19 | 14 - | 119 | 19 - | 587 |
| 5 - | 5 | 10 - | 169 | 15 - | 120 | 20 - | 594 |

СН270 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА = 0.050

ТАБЛИЦА СОСТАВЛЕНА ИЗ ДЛЛЕЙ 1-Я ГРАДАЦИИ 6 ВЫБОРК:

| NN - | К О Д | В Ы Б О Р К И | ДЛИНА | ДОЛЯ |
|------|-----------------------|---------------|-------|---------|
| 1 - | ТЕСТОВАЯ /2 ГРАДАЦИИ/ | ВЫБОРКА 31 | 80 | 0.80000 |
| 2 - | ТЕСТОВАЯ /2 ГРАДАЦИИ/ | ВЫБОРКА 32 | 85 | 0.67000 |
| 3 - | ТЕСТОВАЯ /2 ГРАДАЦИИ/ | ВЫБОРКА 33 | 78 | 0.54000 |
| 4 - | ТЕСТОВАЯ /2 ГРАДАЦИИ/ | ВЫБОРКА 34 | 87 | 0.85000 |
| 5 - | ТЕСТОВАЯ /2 ГРАДАЦИИ/ | ВЫБОРКА 35 | 70 | 0.94000 |
| 6 - | ТЕСТОВАЯ /2 ГРАДАЦИИ/ | ВЫБОРКА 36 | 82 | 0.50000 |

 * ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ
 * ДИСПЕРСИОННЫЙ АНАЛИЗ * А Н А Л И З
 * ГРУППЫ "К" НЕЗАВИСИМЫХ ВЫБОРОК

PR270 - СПРАВЕДЛИВА ГИПОТЕЗА H1 КОД ОШИБКИ = 0

ТАБЛИЦА СОСТАВЛЕНА ИЗ ДОЛЕЙ 1-И ГРАДАЦИИ 6 ВЫБОРОК

К О Д Ы В Ы Б О Р О К

ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 31
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 32
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 33
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 34
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 35
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 36

ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ КРИТЕРИЯ ФИШЕРА 13.324
 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ F-РАСПРЕДЕЛЕНИЯ = 2.2319

РЕЗУЛЬТАТЫ ПОПАРНОГО СРАВНЕНИЯ ПО Т-КРИТЕРИЮ:

| ВЫБОРКИ | ИНДЕКС-Т | ВЫЧ.ЗНАЧ. | КРИТ.ЗНАЧ. | ИНДЕКС-D | |
|---------|----------|-----------|------------|----------|---|
| 1 | 2 | 0 | 1.96173 | 1.98061 | 1 |
| 1 | 3 | 1 | 3.84078 | 1.98077 | 0 |
| 1 | 4 | 0 | 0.75875 | 1.98057 | 1 |
| 1 | 5 | 1 | 2.01075 | 1.98094 | 1 |
| 1 | 6 | 1 | 4.48744 | 1.98068 | 0 |
| 2 | 3 | 0 | 1.94890 | 1.98066 | 1 |
| 2 | 4 | 1 | 2.77434 | 1.98046 | 1 |
| 2 | 5 | 1 | 3.93223 | 1.98083 | 0 |
| 2 | 6 | 1 | 2.58161 | 1.98057 | 1 |
| 3 | 4 | 1 | 4.67315 | 1.98061 | 0 |
| 3 | 5 | 1 | 5.71095 | 1.98099 | 0 |
| 3 | 6 | 0 | 0.59448 | 1.98072 | 1 |
| 4 | 5 | 0 | 1.31760 | 1.98079 | 1 |
| 4 | 6 | 1 | 5.34532 | 1.98053 | 0 |
| 5 | 6 | 1 | 6.35579 | 1.98090 | 0 |

РЕЗУЛЬТАТЫ ПОПАРНОГО СРАВНЕНИЯ ПО КРИТЕРИЮ D:

| ВЫБОРКИ | ИНДЕКС-D | СУММА | РАЗНОСТЬ | КРИТИЧЕСКИЕ ЗНАЧЕНИЯ |
|---------|----------|-------|----------|-------------------------|
| 1 | 2 | 1 | 0.20184 | 0.13000 2.11100 2.10640 |
| 1 | 3 | 0 | 0.21365 | 0.26000 2.11100 2.11300 |
| 1 | 4 | 1 | 0.17498 | 0.05000 2.11100 2.10460 |
| 1 | 5 | 1 | 0.15461 | 0.14000 2.11100 2.12110 |
| 1 | 6 | 0 | 0.21086 | 0.30000 2.11100 2.10910 |
| 2 | 3 | 1 | 0.22667 | 0.13000 2.10640 2.11300 |
| 2 | 4 | 1 | 0.18800 | 0.18000 2.10640 2.10460 |
| 2 | 5 | 0 | 0.16764 | 0.27000 2.10640 2.12110 |
| 2 | 6 | 1 | 0.22389 | 0.17000 2.10640 2.10910 |
| 3 | 4 | 0 | 0.19981 | 0.31000 2.11300 2.10460 |
| 3 | 5 | 0 | 0.17945 | 0.40000 2.11300 2.12110 |
| 3 | 6 | 1 | 0.23570 | 0.04000 2.11300 2.10910 |
| 4 | 5 | 1 | 0.14078 | 0.09000 2.10460 2.12110 |
| 4 | 6 | 0 | 0.19702 | 0.35000 2.10460 2.10910 |
| 5 | 6 | 0 | 0.17666 | 0.44000 2.12110 2.10910 |

```

C; SP270 - ДИСПЕРСИОННЫЙ ОДНОФАКТОРНЫЙ АНАЛИЗ
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ПРОВЕДЕНИЯ ОДНОФАКТОРНОГО
C; ДИСПЕРСИОННОГО АНАЛИЗА
C
CR ЛИТЕРАТУРА ПО АЛГОРИТМУ : АФИФИ А., ЭЙЗЕН С.,
C; СТАТИСТИЧЕСКИЙ АНАЛИЗ, М., 1982, СТР.237-239
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММО,
C; ) Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE SP270(KEY,NFWR,KSQN,K,M,N,P,L,Q
* ,F,TF,IND,ST,TS,TIN,NST,INDST,IER,IERW)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CP ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C
C KEY I*4 * * (8) МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 * * 1 НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<NFWR<=99/
C K SQN R*8 * * (4*M) МАССИВ К О Д О В ВЫБОРОК
C K I*4 * * 1 НОМЕР ИССЛЕДУЕМОЙ ГРАДАЦИИ
C M I*4 * * 1 КОЛИЧЕСТВО ВАРИАНТОВ ОПЫТА /2<=M/
C N I*4 * * (M) МАССИВ ДЛИН ВЫБОРОК /2<=N(I)/
C P R*4 * * (M) МАССИВ ДОЛЕЙ К-И ГРАДАЦИИ ИЗ "М" ВЫБОРОК
C L I*4 * * 1 ! ДОЛЖНО БЫТЬ: L=M*(M-1)/2
C Q R*4 * * 1 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА /.005<=Q<=.5/
C F R*4 * * 1 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ КРИТЕРИЯ ФИШЕРА
C TF R*4 * * 1 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ F-РАСПРЕДЕЛЕНИЯ
C IND I*4 * * 1 ИНДЕКС ОТВЕТА ПО КРИТЕРИЮ ФИШЕРА :
C IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ (НЕТ РАЗЛИЧИЯ);
C IND = 1 - НЕЛЕВАЯ ГИПОТЕЗА ОТКЛОНЯЕТСЯ.
C ST R*4 * * (L,3) ВЫЧИСЛЕННЫЕ ЗНАЧЕНИЯ КРИТЕРИЯ СТЬЮДЕНТА
C TS R*4 * * (L) КРИТИЧЕСКИЕ ЗНАЧ. Т-РАСПР. (ДЛЯ Т-КРИТ.)
C TIN R*4 * * (L,2) КРИТИЧЕСКИЕ ЗНАЧ. Т-РАСПР. (ДЛЯ ИНТ.ОЦ.)
C NST I*4 * * (L,2) УКАЗАТЕЛИ ПАР НОМЕРОВ ВЫБОРОК
C INDST I*4 * * (L,2) ИНДЕКСЫ ОТВЕТА ПО КРИТЕРИЮ СТЬЮДЕНТА
C IER I*4 * * 1 ВЫРАБОТАННЫЙ КОД ОШИБКИ
C IERW I*4 * * (8) МАССИВ КОДОВ ОШИБОК ПРИНЯТЫХ ИЗ SW270
C
C DIMENSION KEY(8),K SQN(4,M),N(M),P(M),IERW(8)
* ,ST(L,3),TS(L),TIN(L,2),NST(L,2),INDST(L,2)
C REAL*8 NAME,K SQN
C
C DATA NAME/' SP270 ',ZERO/0.0/,ONE/1.0/
C
C IND = 11
C IER = 0
C
C KLP=KEY(2)
C IF (KEY(1).EQ.0) GO TO 120
C CALL SW270(KEY,NFWR,K SQN,K,M,N,P,L,Q,IERW,MWC)
C; SW270 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP270;
C
C АНАЛИЗ КОДОВ ОШИБОК : "KF" И "MWC"
C
C KF = IERW(1)
C IF (MWC.LT.3) GO TO 120
C ЕСЛИ MWC>=3, ТО IER=MWC
C IER = MWC
C IF (MWC .GE. 3) GO TO 999
C 120 IND= 0
C ВЫЧИСЛЕНИЕ ОБЩЕЙ ДИСПЕРСИИ

```

```

S=ZERO
R=ZERO
DO 150 I=1,M
Z=FLOAT(N(I))
R=R+P(I)*Z
S=S+Z
150 CONTINUE
QD=R-R*R/S
D=QD/(S-ONE)
C ВЫЧИСЛЕНИЕ ДИСПЕРСИИ МЕЖДУ ВЫБОРКАМИ
Z=ZERO
DO 250 I=1,M
Z=Z+P(I)**2*FLOAT(N(I))
250 CONTINUE
QM=Z-R*R/S
DM=QM/(M-1)
C ВЫЧИСЛЕНИЕ ВНУТРИВЫБОРОЧНОЙ ДИСПЕРСИИ
DV=(QD-QM)/(S-FLOAT(M-2))
C ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТА F
F=DM/DV
C ВЫЧИСЛЕНИЕ ТАБЛИЧНОГО ЗНАЧЕНИЯ РАСПРЕДЕЛЕНИЯ ФИШЕРА
CALL MP040(Q,M-1,S-ONE,0.001,TF,IERF)
CL; MP040 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ РАСПР. ФИШЕРА
IF (IERF.EQ.0) GO TO 300
IER = 4+IERF
GO TO 810

C
C НЕОБХОДИМА ЛИ ПОПАРНАЯ ПРОВЕРКА ПО Т-КРИТЕРИЮ
300 IF (F.LE.TF) GO TO 810
C ОРГАНИЗАЦИЯ ПОПАРНОЙ ПРОВЕРКИ
IL=0
MM=M-1
DO 350 I=1,MM
N1 = N(I)
IP=I+1
DO 350 J=IP,M
N2 = N(J)
IL=IL+1
C ЗАПОМНИЛИ НОМЕРА ВЫБОРОК
NST(IL,1)=I
NST(IL,2)=J
C ----- Т-КРИТЕРИЙ -----
C ВЫЧИСЛЕНИЕ SD - ОШИБКИ РАЗНОСТИ СРАВНИВАЕМЫХ ДОЛЕЙ
SD=SQRT(DV)*SQRT(ONE/FLOAT(N1)+ONE/FLOAT(N2))
C Т-КРИТЕРИЙ
ST(IL,1)=ABS(P(I)-P(J))/SD
C ВЫЧИСЛЕНИЕ ТАБЛИЧНОГО ЗНАЧЕНИЯ РАСПРЕДЕЛЕНИЯ СТЮДЕНТА
CALL MP010(Q, N1+N2-2, 0.001, TZ, IERT)
CL; MP010 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ Т-РАСПРЕДЕЛЕНИЯ
IF (IERT.EQ.0) GOTO 310
IER = 4+IERT
GOTO 810
C ЗАПОМНИЛИ TZ
310 TS(IL)=TZ
C ПРОВЕРКА НУЛЕВОЙ ГИПОТЕЗЫ
ID=0
IF (ST(IL,1).GT.TS(IL)) ID=1
C ЗАПОМНИЛИ ИНДЕКС
INDST(IL,1)=ID
C ----- ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ -----
C ПОЛАГАЕМ, ЧТО ИМЕЕМ ТОЛЬКО 2 ГРАДАЦИИ
SP1=SQRT(P(I)*(ONE-P(I)))/SQRT(FLOAT(N1))
SP2=SQRT(P(J)*(ONE-P(J)))/SQRT(FLOAT(N2))
C
CALL MP010(Q, N1-1, 0.001, TQV1, IERT)
CL; MP010 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ Т-РАСПРЕДЕЛЕНИЯ

```

```

      IF(IERT.EQ. 0) GOTO 320
      IER = 4+IERT
      GOTO 810
320 CALL MP010(Q, N2-1, 0.001, TQV2, IERT)
CL; MP010 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ T-РАСПРЕДЕЛЕНИЯ
      IF(IERT.EQ. 0) GOTO 330
      IER = 4+IERT
      GOTO 810
C
330 TIN(IL,1)=TQV1
      TIN(IL,2)=TQV2
C
      ST(IL,2)=TQV1*SP1+TQV2*SP2
      ST(IL,3)=ABS(P(I)-P(J))
      ID = 0
      IF (ST(IL,2).GT.ST(IL,3)) ID = 1
      INDST(IL,2) = ID
350      CONTINUE
C
400 IND = 1
C
810 IF (KF.NE. 3) GO TO 830
      IER=-IER
      GO TO 999
830 IF (KLP.EQ.0) GO TO 999
      KEY(1)=KLP
      CALL PR270(KEY,NFWR,KSQN,K,M,L
      * ,F,TF,IND,ST,TS,TIN,NST,INDST,IER)
CL; PR270 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP270
999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ 2.0
      END
02/02/88

```

```

C; SW270 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP270
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ КОНТРОЛЯ ВХОДНЫХ ДАННЫХ ДЛЯ SP270
C И ВЫДАЧИ В ФАЙЛ С НОМЕРОМ "NFWR" РЕЗУЛЬТАТА ПРОВЕРКИ
C; И ВХОДНЫХ ДАННЫХ
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP270
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧМФФ
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE SW270(KEY,NFWR,KSQN,K,M,N,P,L,Q,IERW,MWC)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
C KSQN R*8 (4,M) 4A8 7 МАССИВ КОДОВ ВЫБОРОК
C K I*4 1 I3 7 НОМЕР ИССЛЕДУЕМОЙ ГРАДАЦИИ /1<=K<=99/
C M I*4 1 I3 7 КОЛИЧЕСТВО ВАРИАНТОВ ОПЫТА (ВЫБОРОК)
C N I*4 (M) I5 7 МАССИВ ДЛИН ВЫБОРОК /N(I)>=2/
C P R*4 (M) F8.3 7 МАССИВ ДОЛЕЙ К-И ГРАДАЦИИ ВЫБОРОК
C L I*4 1 I8 [8]ДОЛЖНО БЫТЬ: L=M*(M-1)/2
C Q R*4 1 F8.3 3 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА/.005<=Q<=.5/
C IERW I*4 (8) 7I2 [2]ВЫРАБОТАННЫЙ КОД ОШИБКИ : 7 РАЗРЯДОВ
C;MWC I*4 1 12 2 КОД МАКСИМАЛЬНОЙ ОШИБКИ
C
C
C DIMENSION KEY(8),IERW(8),KSQN(4,M),N(M),P(M)
C REAL*8 NAME,KSQN
C DATA NAME /' SW270 '/,CONST/0.001/,EPS/0.001/
C
C KLU=KEY(1)
C IF(KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 11 KLU=IABS(KLU)
C DO 13 J=1,8
C I=9-J
C KEY(I)=MOD(KLU,2)
C 13 KLU=KLU/2
C GOTO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 15 DO 17 I=1,8
C KEY(I)=1
C IF(KLU.LT.I) KEY(I)=0
C 17 CONTINUE
C 19 CONTINUE
C КОНТРОЛЬ ВХОДНЫХ ДАННЫХ
C
C -*- ПРОВЕРКА 'NFWR' ( 1-И РАЗРЯД )
C НАЧАЛЬНАЯ УСТАНОВКА : ОШИБОК НЕТ
C KF = 0
C ЕСЛИ 'NFWR' БОЛЬШЕ 99 ИЛИ МЕНЬШЕ 1, ЭТО ГРУБАЯ ОШИБКА
C IF (NFWR.GT.99) KF = 3
C IF (NFWR.LT.1) KF = 3
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
C IERW(1) = KF
C
C
C -*- ПРОВЕРКА 'K' ( 2-И РАЗРЯД )
C НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'K' : 1 <= K <= 99
C KK = 0

```



```

C ЕСЛИ K >= 100 ЧТО ЛОГИЧЕСКАЯ ОШИБКА
  IF (K.GE.100) KK = 1
C НЕПРАВИЛЬНОЕ ЗНАЧЕНИЕ 'K' : K < 1
  IF (K.LT.1) KK = 3
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
  IERW(2) = KK
C
C-#-      ПРОВЕРКА 'M' ( 3-Й РАЗРЯД )
C НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'M' : 3 <= M <= 9
  KM = 0
C ЕСЛИ M >= 100 ЭТО ЛОГИЧЕСКАЯ ОШИБКА
  IF (M.GE.100) KM = 1
C НЕПРАВИЛЬНОЕ ЗНАЧЕНИЕ 'M' : M < 3
  IF (M.LT.3) KM = 3
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
  IERW(3) = KM
C
C-#-      ПРОВЕРКА 'N' ( 4-Й РАЗРЯД )
  KNM = 11
  IF (KM.GT. 1) GO TO 40
C KNM=0 ОЗНАЧАЕТ, ЧТО 'N'- НОРМАЛЬНОЕ, Т.Е. БОЛЬШЕ 9
  KNM = 0
  KPM = 0
  KPN = 0
  DO 35 I=1,M
C ЕСЛИ 'N(I)' МЕНЬШЕ ЧЕМ 10, ТО ЭТО ЛОГИЧЕСКАЯ ОШИБКА
  IF (N(I).LT.10) KPM = 1
C ЕСЛИ 'N(I)' МЕНЬШЕ ЧЕМ 2, ЭТО ГРУБАЯ ОШИБКА
  IF (N(I).LT. 2) KPN = 3
35 CONTINUE
C ПО ВЕРХНЕМУ ПРЕДЕЛУ ДЛЯ 'N' ПРОВЕРКИ НЕТ
  KNM = MAX0 ( KPM, KPN)
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
40 IERW(4) = KNM
C
C-#-      ПРОВЕРКА 'P' ( 5-Й РАЗРЯД )
  KP = 11
C ЕСЛИ 'M'- НЕПРАВИЛЬНОЕ, ПРОВЕРКА НЕВОЗМОЖНА
  IF (KM.GT.1) GO TO 50
C НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'M'.
  KP = 0
C ПРОВЕРКА НА ОЧЕНЬ МАЛОЕ ЗНАЧЕНИЕ 'P':
  KPM = 0
  DO 44 I=1,M
    IF (P(I).LT.CONST) KPM = 1
44 CONTINUE
C ПРОВЕРКА НА НЕПРАВИЛЬНОЕ ЗНАЧЕНИЕ 'P' :
  KPN = 0
  DO 48 I=1,M
    IF (P(I).GT.1.0-CONST .OR. P(I).EQ.0.0) KPN = 3
48 CONTINUE
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
  KP = MAX0(KPM,KPN)
50 IERW(5) = KP
C
C-#-      ПРОВЕРКА 'L' ( 6-Й РАЗРЯД )
C НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'L' = M*(M-1)/2
  KL = 3
  IF (L.EQ.(M*(M-1)/2)) KL = 0
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
  IERW(6) = KL
C
C-#-      ПРОВЕРКА 'Q' ( 7-Й РАЗРЯД )
C НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'Q' : 0.025 <= Q <= 0.2
  Q0 = 0
C ЕСЛИ 'Q' БОЛЬШЕ ЧЕМ 0.2 ИЛИ МЕНЬШЕ ЧЕМ 0.025,

```

290

```

CC; PR270 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP270
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫВОДА В ФАЙЛ 'NFWR'
C; РЕЗУЛЬТАТОВ РАБОТЫ П/П SP270
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP270
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE PR270(KEY,NFWR,KSQN,K,M,L
* ,F,TF,IND,ST,TS,TIN,NST,INDST,IER)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
C KSQN R*8(4,M) 4A6 7 МАССИВ КОДОВ ВЫБОРОК
C K I*4 1 I3 7 НОМЕР ИССЛЕДУЕМОЙ ГРАДАЦИИ
C M I*4 1 I3 7 КОЛИЧЕСТВО ВАРИАНТОВ ОПЫТА
C L I*4 1 -- - L ДОЛЖНО БЫТЬ РАВНО M*(M-1)/2
C F R*4 1 G18.5 3 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ КРИТЕРИЯ ФИШЕРА
C TF R*4 1 G16.5 3 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ F-РАСПРЕДЕЛЕНИЯ
C IND I*4 1 I1 2 ИНДЕКС ПРОВЕРКИ ПО КРИТЕРИЮ ФИШЕРА:
C IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
C IND = 1 - НУЛЕВАЯ ГИПОТЕЗА ОТВЕРГАЕТСЯ.
C ST R*4(L,3) F10.5 8 ВЫЧИСЛ.ЗНАЧЕНИЯ Т-КРИТ. И ИНТ. ОЦЕН.
C TS R*4 (L) F10.5 8 КРИТИЧ. ЗНАЧЕНИЯ Т-КРИТЕРИЯ ДЛЯ ПАР
C TIN R*4(L,2) F10.5 8 КРИТИЧ. ЗНАЧЕНИЯ ИНТ. ОЦЕН. ДЛЯ ПАР
C NST I*4(L,2) I4 8 МАССИВ ПАР НОМЕРОВ ВЫБОРОК
C INDST I*4(L,2) I1 8 ИНДЕКСЫ ПРОВЕРКИ ПО КРИТ. СТУДЕНТА
C; IER I*4 1 I3 2 КОД ОШИБКИ, ВЫРАБОТАННЫЙ SP270
C
DIMENSION KEY(8),KSQN(4,M)
* ,ST(L,3),TS(L),TIN(L,2),NST(L,2),INDST(L,2)
REAL*8 NAME, KSQN, TNK(3), TS1(4), TS2(4), TS3(4)
DATA TNK /' ДИСПЕРС','ИОННЫЙ ',' АНАЛИЗ '/
* ,TS1 /' ОДНОФА','КТОРНЫЙ ',' ДИСПЕРС','ИОННЫЙ '//
* ,TS2 /' ',' А Н А ','Л И З ',' '//
* ,TS3 /' ГРУППЫ ','"К" НЕЗА','ВИСИМЫХ ','ВЫБОРОК '//
DATA NAME/' PR270 '/
C
KLU=KEY(1)
IF(KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
13 KLU=KLU/2
GO TO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF(KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
C
IF (KEY(1).EQ.0) GO TO 210
WRITE (NFWR,120) TS1,TNK,TS2,TS3

```


где a — число изменений благоприятных ответов на неблагоприятные; d — число изменений неблагоприятных ответов на благоприятные.

$H_0: p_a = p_d = 0,5$, т. е. в общей совокупности доля изменений благоприятных ответов на неблагоприятные равна доле изменений неблагоприятных ответов на благоприятные.

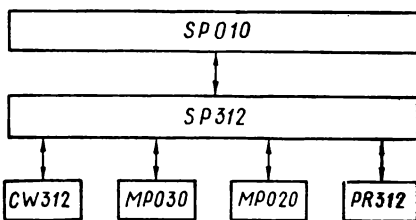


Рис. 56

Проверка нулевой гипотезы, когда сумма $a + d > 50$, осуществляется путем сопоставления наибольшего из значений a, d с критическим значением биномиального распределения $n_{кр}$.

H_0 отвергается, если $\max(a, d) > n_{кр}$.

Если $a + d > 50$, для проверки $H_0: p_a = p_d = 0,5$ применяется критерий χ^2 :

$$\chi^2 = \frac{(|a + d| - 1)}{a + d}.$$

H_0 отвергается, если $\chi^2 \geq \chi_{q,v}^2$, где $\chi_{q,v}^2$ вычисляется при $v = 1$ и заданном уровне значимости риска.

4.2. Программная реализация. Оценка различия двух связанных выборок может быть представлена блок-схемой на рис. 56.

Подпрограмма $SP312$ — функциональный модуль, определяющий оценку различия двух связанных выборок; подпрограмма $CW312$ осуществляет контроль входных данных и их печать, а подпрограмма $PR312$ печать результатов, подпрограмма $MP040$ — вычисление критических значений биномиального распределения. Тексты подпрограмм $SP312, CW312, PR312$ приводятся в порядке перечисления после примера.

4.3. Пример.

В задании на оценку различия двух связанных выборок необходимо указать $IRP = 004$ (точнее — $XX4$, где X — любая цифра от 0 до 9). Данные задаются непосредственно в виде таблицы экспериментальных данных.

Уровень значимости риска $Q = 0,05$.

ЗАДАНИЕ НА ОЦЕНКУ РАЗЛИЧИЯ ДВУХ СВЯЗАННЫХ ВЫБОРОЧНЫХ СОВОКУПНОСТЕЙ ПО КРИТЕРИЮ МАКНИМАРЫ (ХИ-КВАДРАТ)

```

----- ОЦЕНКА РАЗЛИЧИЯ ДВУХ СВЯЗАННЫХ ВЫБОРОК :
**** МАКЕТ ПОДГОТОВКИ ДАННЫХ ДЛЯ IR= 4: ****SP312****
IRP -004 (I3)- КОД ФУНКЦИИ
NS -02 /ПО УМОЛЧАНИЮ/ (I2)- КОЛИЧЕСТВО ВЫБОРОК
K -02 /ПО УМОЛЧАНИЮ/ (I2)- КОЛИЧЕСТВО ГРАДАЦИИ ВЫБОРКИ 1
M -02 /ПО УМОЛЧАНИЮ/ (I2)- КОЛИЧЕСТВО ГРАДАЦИИ ВЫБОРКИ 2
(I2)-НОМЕР; (4A8)- К О Д ВЫБОРКИ; (I4)-ДЛИНА
XX XXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXX
KSNQ-01 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 41 0002
KSNQ-02 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 42 0002
(I2, I3, I2)[ 4A8 ]:НОМЕР ВЫБ.,КОД,КАНАЛ,ИМЯ ФАЙЛА
FILT-01 000 08-INF:DATAT4.DAT
[ -СПИСОК СПЕЦИФИКАЦИЙ ОПЕРАТОРА F O R M A T (14A4) ]
FORT-01 (4X,4F6.1)
36.0 48.0 40.0 22.0
Q -0.05 - (G14.7)- УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
NFWR-02 (I2) - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
(I1, I2)-ТИП КЛЮЧЕЙ ПЕЧАТИ: 0-ПОРАЗРЯДНЫЕ; И ИХ КОЛИЧЕСТВО.
KEYC-0 02 1-ИНТЕРВАЛЬНЫЕ.
ИНДЕКСЫ/ИНТЕРВАЛЬНОЕ ЗНАЧЕНИЕ КЛЮЧА ПЕЧАТИ/ПОРАЗРЯДНОЕ/
12345678 - РАЗРЯДЫ КЛЮЧА ПЕЧАТИ :
KEY - 1 1 3/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SP312 :
KEY - 2 1 3/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP312 .
(I1)-РАСПРЕДЕЛЕНИЕ ПАМЯТИ:=1-ВЫДАЕТСЯ ВСЕГДА;
END -1 КОНЕЦ ДАННЫХ: =0-ВЫДАЕТСЯ ТОЛЬКО АВАРИЙНО.

```

РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ ЗАДАНИЯ (РАСПЕЧАТКА РЕШЕНИЯ)

```

IRP - 4
KSNQ NN XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX NNNN
KSNQ- 1 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 41 2
KSNQ- 2 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 42 2
FILT- 1 0 8-INF:DATAT4.DAT
FORT- 1: (4X,4F6.1)
      36.0 48.0 40.0 22.0
Q      0.5000000E-01
NFWR 2
KEYC 0 2 2
KEY 1 1 3/8/11111111/
KEY 2 1 3/8/11111111/
ОБЪЕМ РАБОЧЕЙ ПАМЯТИ - 4000 ЕДИНИЦ INTEGER
ТРЕБУЕМЫЙ ОБЪЕМ ПАМЯТИ - 133 ЕДИНИЦ INTEGER

```

ИНДЕКСЫ НАЧАЛА ПЕРЕМЕННЫХ В РАБОЧЕЙ ПАМЯТИ :

| | | | | | | | |
|-----|---|------|-----|------|-----|------|-----|
| 1 - | 1 | 6 - | 6 | 11 - | 107 | 16 - | 60 |
| 2 - | 2 | 7 - | 7 | 12 - | 108 | 17 - | 60 |
| 3 - | 3 | 8 - | 36 | 13 - | 109 | 18 - | 124 |
| 4 - | 4 | 9 - | 19 | 14 - | 55 | 19 - | 125 |
| 5 - | 5 | 10 - | 105 | 15 - | 59 | 20 - | 126 |

СW312 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА = 0.050

ТАБЛИЦА СФОРМИРОВАНА ИЗ ВЫБОРОК:
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 41
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 42

ТАБЛИЦА (2*2) ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ "Т":

| | |
|------|------|
| 36.0 | 40.0 |
| 48.0 | 22.0 |

```

*****
*                               * ОЦЕНКА РАЗЛИЧИЯ
* КРИТЕРИЙ МАКНИМАРА * ДВУХ СВЯЗАННЫХ ВЫБОРОК,
*                               * ИМЕЮЩИХ ТОЛЬКО ПО ДВЕ ГРАДАЦИИ
*****

```

PR312 - СПРАВЕДЛИВА ГИПОТЕЗА Н0 ; КОД ОШИБКИ = 0

| | |
|------------------------------------|--------|
| ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ КРИТЕРИЯ | 2.9138 |
| КРИТИЧЕСКОЕ ЗНАЧЕНИЕ РАСПРЕДЕЛЕНИЯ | 3.841 |

ТАБЛИЦА СФОРМИРОВАНА ИЗ ВЫБОРОК:
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 41
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 42

```

СВ; SP312 - ОЦЕНКА РАЗЛИЧИЯ ДВУХ СВЯЗАННЫХ ВЫБОРОК ПО ХИ-КВАДРАТ
С
СВ П/П ПРЕДНАЗНАЧЕНА ДЛЯ ОЦЕНКИ РАЗЛИЧИЯ ДВУХ СВЯЗАННЫХ
С ВЫБОРОК, ИМЕЮЩИХ ТОЛЬКО ПО ДВЕ ГРАДАЦИИ,
С; ПО КРИТЕРИЮ МАКНИМАРЫ (ХИ-КВАДРАТ)
С
СВ ЛИТЕРАТУРА ПО АЛГОРИТМУ : РУНИОН Р., СПРАВОЧНИК
С ПО НЕПАРАМЕТРИЧЕСКОЙ СТАТИСТИКЕ, М., ФИНАНСЫ И
С СТАТИСТИКА, 1982.
С;
СВ; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
С
СВ П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧИПТ,
С; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
С
SUBROUTINE SP312(KEY,NFWR,KSQN,L,T,Q,HIM,THI,IND,IER,IERW)
С
С ОПИСАНИЕ ПАРАМЕТРОВ
С -----
С
СВ ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
С
С KEY I*4 * (8) МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
С NFWR I*4 * 1 НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
С KSQN R*8 * (4,L) МАССИВ КОДОВ ВЫБОРОК
С L I*4 * 1 ДЛИНА "KSQN" L=1 ИЛИ L=2
С T R*4 * (2,2) ТАБЛИЦА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ
С Q R*4 * 1 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА / <=Q<=
С HIM R*4 * 1 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ КРИТЕРИЯ МАКНИМАРЫ
С THI R*4 * 1 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ РАСПРЕД. ХИ-КВАДРАТ
С IND I*4 * 1 ИНДЕКС ПРОВЕРКИ ПО ИНТЕРВ. ОЦЕНИВАНИЮ
С IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
С IND = 1 - НЕЛЕВАЯ ГИПОТЕЗА ОТКЛОНЯЕТСЯ.
С IER I*4 * 1 ВЫРАБОТАННЫЙ КОД ОШИБКИ
С IERW I*4 * (8) МАССИВ КОДОВ ОШИБОК, ПРИНЯТЫХ ИЗ СВ312
С
DIMENSION KEY(8),IERW(8),T(2,2),KSQN(4,L)
REAL*8 NAME, KSQN
DATA NAME /' SP312 '// EPS/0.001/, ONE/1.0/, FIFTY/50.0/
С
IND = 11
IER = 0
KLP=KEY(2)
IF (KEY(1).EQ.0) GO TO 120
CALL SW312(KEY,NFWR,KSQN,L,T,Q,IERW,MWC)
СВ СВ312 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ SP312;
С АНАЛИЗ КОДОВ ОШИБОК "KF" И "MWC"
KF = IERW(1)
IF (MWC.LT.3) GO TO 120
С ЕСЛИ MWC >= 3, ТО IER=MWC
IER = MWC
IF (MWC.GE.3) GO TO 999
С
120 S = T(1,1) + T(2,2)
С
IF (S.LE. FIFTY) GO TO 250
HIM = (ABS(T(1,1)-T(2,2))-ONE)**2/S
С ВЫЧИСЛЕНИЕ КРИТИЧЕСКОГО ЗНАЧЕНИЯ ХИ-КВАДРАТ
CALL MP030(Q, 1, EPS, THI, IERTHI)
СВ MP030 - ВЫЧ. КРИТИЧЕСКИХ ЗНАЧЕНИЙ РАСПРЕДЕЛЕНИЯ ХИ-КВАДРАТ;
IF (IERTHI.EQ. 0) GOTO 230
IER = 4+IERTHI
GOTO 810
230 IND = 0
IF (HIM.GE. THI) IND=1
GO TO 810

```



```

250 HIM = AMAX1(T(1,1),T(2,2))
C  ВЫЧИСЛЕНИЕ КРИТИЧЕСКОГО ЗНАЧЕНИЯ БИНОМИАЛЬНОГО РАСПРЕД .
CALL MP020(Q, INT(S+EPS), EPS, THI, IERBIN)
CL; MP020 - ВЫЧ. КРИТИЧ. ЗНАЧЕНИИ БИНОМИАЛЬНОГО РАСПРЕДЕЛЕНИЯ;
      IF (IERBIN .EQ. 0) GOTO 730
      IER = 4+IERBIN
      GOTO 810
730 IND = 0
      IF (HIM .GE. THI) IND=1
C
810 IF (KF .NE. 3) GO TO 830
      IER = -IER
      GO TO 999
830 IF (KLP.EQ.0) GO TO 999
      KEY(1)=KLP
      CALL PR312(KEY,NFWR,KSQN,L,HIM,THI,IND,IER)
CL; PR312 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP312.
999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
      END

```

```

C; SW312 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP312
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ КОНТРОЛЯ ВХОДНЫХ ДАННЫХ ДЛЯ SP312
C И ВЫДАЧИ В ФАЙЛ С НОМЕРОМ "NFWR" РЕЗУЛЬТАТА ПРОВЕРКИ
C И ВХОДНЫХ ДАННЫХ
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP312
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ
C Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE SW312(KEY,NFWR,KSQN,L,T,Q,IERW,MWC)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
C KSQN R*8 (4,L) 4A8 7 МАССИВ КОДОВ ВЫБОРОК
C L I*4 1 -- - ДЛИНА "KSQN" /L=1 ИЛИ L=2/
C T R*4 (2,2) F10.1 8 ТАБЛИЦА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ
C Q R*4 1 F8.3 3 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
C IERW I*4 (8) 4I2 [2] ВЫРАБОТАННЫЙ КОД ОШИБКИ : 4 РАЗРЯД
C;MWC I*4 1 I2 2 КОД МАКСИМАЛЬНОЙ ОШИБКИ
C
DIMENSION KEY(8),T(2,2),KSQN(4,L),IERW(8)
REAL*8 NAME, KSQN
DATA NAME /' SW312 '/, ZERO/0.0/
C
KLU=KEY(1)
IF(KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
13 KLU=KLU/2
GOTO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF(KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
C
C КОНТРОЛЬ ВХОДНЫХ ДАННЫХ
C
C-*- ПРОВЕРКА 'NFWR' ( 1-й РАЗРЯД )
C НАЧАЛЬНАЯ УСТАНОВКА : ОШИБОК НЕТ
KF = 0
C ЕСЛИ 'NFWR' БОЛЬШЕ 99 ИЛИ МЕНЬШЕ 1, ЭТО ГРУБАЯ ОШИБКА
IF (NFWR.GT.99) KF = 3
IF (NFWR.LT.1) KF = 3
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
IERW(1) = KF
C
C-*- ПРОВЕРКА 'T' ( 2-й РАЗРЯД )
C KT=0 ОЗНАЧАЕТ, ЧТО 'T'- НОРМАЛЬНАЯ ТАБЛИЦА,
C Т.Е. ЭЛЕМЕНТЫ 'T' БОЛЬШЕ НУЛЯ
KT = 0
IF (T(1,1).LT.ZERO .OR. T(2,2).LT.ZERO) KT = 2

```

```

      IF ((T(1,1)+T(2,2)).EQ.ZERO) KT = 2
C  ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(2) = KT
C
C--  ПРОВЕРКА 'L' ( 3-й РАЗРЯД )
      KL = 3
      IF (L.EQ.1 .OR. L.EQ.2) KL = 0
C  ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(3) = KL
C
C--  ПРОВЕРКА 'Q' ( 4-й РАЗРЯД )
C  НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'Q' : 0.025 <= Q <= 0.2
      KQ = 0
C  ЕСЛИ 'Q' БОЛЬШЕ ЧЕМ 0.2 ИЛИ МЕНЬШЕ ЧЕМ 0.025,
C  ТО СЧИТАЕМ ЧТО ДОПУЩЕНА ЛОГИЧЕСКАЯ ОШИБКА
      IF (Q.LT.0.025) KQ = 1
      IF (Q.GT.0.2) KQ = 1
C  ВВОДИМ ОГРАНИЧЕНИЕ : 0.005 <= Q <= 0.5
      IF (Q.LT.0.005) KQ = 3
      IF (Q.GT.0.5) KQ = 3
C  ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(4) = KQ
      IERW(8) = 4
C
C  ОПРЕДЕЛЕНИЕ МАКСИМАЛЬНОГО КОДА ОШИБКИ
      MWC = MAX0 ( KT, KL, KQ )
C
C  ШКАЛА КОДИРОВКИ ОШИБОЧНЫХ ПАРАМЕТРОВ :
C
CE  РАЗРЯД КОДА ОШИБКИ  I  1  I  2  I  3  I  4
C  -----
C  ПАРАМЕТР В ОБРАЩЕНИИ  I  NFWR  I  T  I  L  I  Q
C  -----
C  ВОЗМОЖНЫЕ ЗНАЧЕНИЯ  I  0  I  0  I  0  I  0
C  I  -  I  -  I  -  I  1
C  I  -  I  -  I  -  I  -
C  I  3  I  3  I  3  I  3
C
C  ПЕЧАТЬ ВХОДНЫХ ДАННЫХ
      IF (KF.GT.1) GO TO 999
      IF (KEY(2).EQ.0) GOTO 225
      WRITE (NFWR,220) NAME, MWC
220  FORMAT(' ',A8,4X,'КОД МАКСИМАЛЬНОЙ ОШИБКИ =',I2)
225  IF (MWC.LT.1) GOTO 310
      J=IERW(8)
      WRITE (NFWR,230) (IERW(I),I=1,J)
230  FORMAT(' ',8X,'ИМЕНА ПАРАМЕТРОВ : '
*      ', NFWR, T, L, Q'
*      '/ ',8X,'КОДЫ ОШИБОК : ',4(3X,I2)/)
C
310  IF (KEY(3).EQ.0) GOTO 710
      WRITE (NFWR,340) Q
340  FORMAT(' ',8X,'УРОВЕНЬ ЗНАЧИМОСТИ РИСКА =',F8.3)
C
710  IF ((KEY(7).EQ.0).OR. (L.EQ.1)) GOTO 810
      WRITE (NFWR,750) ((KSNQ(I,J),I=1,4),J=1,2)
750  FORMAT(' ',8X,'ТАБЛИЦА СФОРМИРОВАНА ИЗ ВЫБОРКИ : '
*      '/ ',8X,5X,4A8))
C
810  IF (KEY(8).EQ.0) GOTO 999
      WRITE (NFWR,870)
870  FORMAT(' ',8X
*      ' ТАБЛИЦА (2*2) ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ "Т": '/')
C
      WRITE (NFWR,880) ((T(I,J),J=1,2),I=1,2)
880  FORMAT(' ',12X,F10.1,5X,F10.1)
C
999  RETURN
CV;  ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0. 8-DEC-87
      END

```

```

CC; PR312 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP312
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫВОДА В ФАЙЛ 'NFWR'
C; РЕЗУЛЬТАТОВ РАБОТЫ П/П SP312
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP312
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧМФФ,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE PR312(KEY,NFWR,KSQN,L,HIM,THI,IND,IER)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CP ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
C KSQN R*8(4,L) 4A8 7 МАССИВ КОДОВ ВЫБОРОК
C L I*4 1 -- - ДЛИНА "KSQN" /L=1 ИЛИ L=2/
C HIM R*4 1 G20.5 3 ВЫЧИСЛЕН.ЗНАЧЕНИЕ КРИТЕРИЯ МАКНИМАРЫ
C THI R*4 1 G20.5 3 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ РАСПРЕДЕЛЕНИЯ
C IND I*4 1 I1 2 ИНДЕКС ОТВЕТА ПО КРИТЕРИЮ
C IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
C IND = 1 - НУЛЕВАЯ ГИПОТЕЗА ОТВЕРГАЕТСЯ.
C IER I*4 1 I3 2 КОД ОШИБКИ, ВЫРАБОТАННЫЙ SP312
C
DIMENSION KEY(8),KSQN(4,L)
REAL*8 NAME, KSQN
REAL*8 TNK(3),TS1(4),TS2(4),TS3(4)
DATA TNK // КРИТЕР', 'ИИ МАК', 'НИМАРА' //
* ,TS1 // 'ОЦЕНКА', 'РАЗЛИЧИЯ', //
* ,TS2 // 'ДВУХ', 'СВЯЗАНН', 'ЫХ ВЫБ', 'ОРОК', //
* ,TS3 // 'ИМЕЮЩИХ', 'ТОЛЬКО', 'ПО ДВЕ Г', 'РАДАЦИИ' //
DATA NAME // PR312 //
C
IF (NFWR.LE.0.OR.NFWR.GE.100) GO TO 999
KLU=KEY(1)
IF (KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
KLU=KLU/2
13 GOTO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF (KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
C
IF (KEY(1).EQ.0) GOTO 210
WRITE (NFWR,120) TS1,TNK,TS2,TS3
120 FORMAT (/ 1H ,26(1H*))
* /2H *,24(1H ),3H* ,4A8
* /2H *,3A8 ,3H* ,4A8
* /2H *,24(1H ),3H* ,4A8
* / 1H ,26(1H*) )
C
210 IF (KEY(2).EQ.0) GOTO 310

```

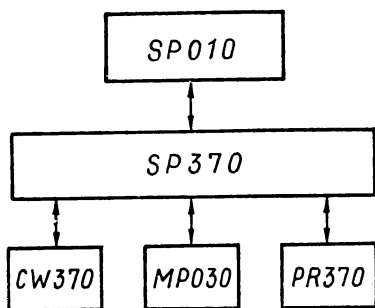



Рис. 57

MP030 — подпрограмма вычисления критических значений распределения χ^2 .

Тексты подпрограмм SP370, CW370, PR370 приведены ниже, после примера.

5.3. Пример. Использование критерия Кокрена при оценке различия m связанных выборок рассмотрим на численном примере. Четыре выборки — результаты измерений над 15 объектами при четырех условиях сведены в таблицу «15 × 4». Конкретные

численные значения указаны (по строкам) в задании на обработку программой «СОКИ-Н» после строки с ключевым словом «FORT» (см. ниже).

Необходимо задать также:

IRP = 005 — режим обработки,

NS = 04 — количество выборок — результатов опыта,

k = 15 — количество объектов,

m = 04 — количество условий опыта ($m = NS$),

Q = 0,05 — уровень значимости риска (5 %).

Следует отметить, что первая и четвертая строки таблицы данных нулевые, а все элементы последней равны единице. Подпрограмма контроля и печати входных данных присваивает код ошибки / параметру T — таблице данных — в том случае, если в исходной таблице

ЗАДАНИЕ — ТЕСТ НА ОЦЕНКУ РАЗЛИЧИЯ "М" СВЯЗАННЫХ ВЫБОРОЧНЫХ
СОВОКУПНОСТЕЙ ПО КРИТЕРИЮ КОКРЕНА

```

----- ОЦЕНКА РАЗЛИЧИЯ "М" СВЯЗАННЫХ ВЫБОРОК :
**** МАКЕТ ПОДГОТОВКИ ДАННЫХ ДЛЯ IR= 5: ****SP370****
IRP -005 (I3)- КОД ФУНКЦИИ
NS -04 (I2)- КОЛ-ВО ВЫБОРОК - РЕЗУЛЬТАТОВ ОПЫТА
K -15 (I2)- КОЛИЧЕСТВО ОБЪЕКТОВ
M -04 (I2)- КОЛИЧЕСТВО УСЛОВИЙ ОПЫТА
(I2)-НОМЕР; (4A8)- К О Д В Ы Б О Р К И; (I4)-ДЛИНА
XX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXX
KSNQ-01 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 51 0015
KSNQ-02 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 52 0015
KSNQ-02 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 53 0015
KSNQ-02 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 54 0015
(I2, I3, I2) ( 4A8 ) : НОМЕР ВЫБ., КОД, КАНАЛ, ИМЯ ФАЙЛА
FILTR-01 000 08-INF:DATAT5.DAT
[ СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) ]
FORT-01 (4X,15F4.0)
0 0 0 0 1 0 1 1 0 0 1 1 0 1 1
0 1 1 0 0 1 1 0 1 0 1 1 0 0 1
0 0 0 0 1 0 0 0 1 0 0 1 0 0 1
0 1 1 0 1 1 1 1 0 1 1 0 1 1 1
Q - 0.05 <- (G14.7)- УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
NFWR-02 (I2) - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
(I1, I2)-ТИП КЛЮЧЕЙ ПЕЧАТИ: 0-ПОРАЗРЯДНЫМ; И ИХ КОЛИЧЕСТВО
KEYS-0 02 1-ИНТЕРВАЛЬНЫМ.
ИНДЕКСЫ/ИНТЕРВАЛЬНОЕ ЗНАЧЕНИЕ КЛЮЧА ПЕЧАТИ/ПОРАЗРЯДНОЕ/
12345678 - РАЗРЯДЫ КЛЮЧА ПЕЧАТИ :
KEY - 1 7 3/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SP370 :
KEY - 2 7 3/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP370 .
(I1)-РАСПРЕДЕЛЕНИЕ ПАМЯТИ:=1-ВЫДАЕТСЯ ВСЕГДА;
END -1 КОНЕЦ ДАННЫХ: =0-ВЫДАЕТСЯ ТОЛЬКО АВАРИЙНО.
  
```

все элементы строки или столбца равны нулю или единице, что может быть следствием ошибки при подготовке информации или проведении эксперимента.

РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ ЗАДАНИЯ (РАСПЕЧАТКА РЕШЕНИЯ):

IRP - 5
NS - 4
K -15
M - 4
KSN NN XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX NNNN
KSN- 1 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 51 15
KSN- 2 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 52 15
KSN- 2 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 53 15
KSN- 2 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 54 15
FIL- 1 0 8-INF:DATAT5.DAT
FORT- 1:(4X,15F4.0)
0. 0. 0. 0. 1. 0. 1. 1. 0. 0. 1. 1. 0. 1. 1.
0. 1. 1. 0. 0. 1. 1. 0. 1. 0. 1. 1. 0. 0. 1.
0. 0. 0. 0. 1. 0. 0. 0. 1. 0. 0. 1. 0. 0. 1.
0. 1. 1. 0. 1. 1. 1. 1. 0. 1. 1. 0. 1. 1. 1.
Q 0.500000E-01
NFWR 2
KEYC 0 2 2
KEY 1 7 3/8/11111111/
KEY 2 7 3/8/11111111/
ОБЪЕМ РАБОЧЕЙ ПАМЯТИ - 4000 ЕДИНИЦ INTEGER
ТРЕБУЕМАЯ ОБЪЕМ ПАМЯТИ - 319 ЕДИНИЦ INTEGER

ИНДЕКСЫ НАЧАЛА ПЕРЕМЕННЫХ В РАБОЧЕЙ ПАМЯТИ :

| | | | | | | | |
|-----|---|------|-----|------|-----|------|-----|
| 1 - | 1 | 6 - | 6 | 11 - | 141 | 16 - | 133 |
| 2 - | 2 | 7 - | 7 | 12 - | 142 | 17 - | 134 |
| 3 - | 3 | 8 - | 36 | 13 - | 143 | 18 - | 310 |
| 4 - | 4 | 9 - | 19 | 14 - | 72 | 19 - | 311 |
| 5 - | 5 | 10 - | 137 | 15 - | 132 | 20 - | 312 |

СВ370 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 1

ИМЕНА ПАРАМЕТРОВ : NFWR, L, N, M, T, Q
К О Д Ы ОШИБОК : 0 0 0 0 1 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА = 0.050

ТАБЛИЦА СФОРМИРОВАНА ИЗ ВЫБОРОК:
ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 51
ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 52
ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 53
ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 54

ЗНАЧЕНИЯ ЭЛЕМЕНТОВ ТАБЛИЦЫ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ
 СТРОКА (СТРОК = 15 СТОЛБЦОВ = 4)

| | | | | |
|----|----|----|----|----|
| 1 | 0. | 0. | 0. | 0. |
| 2 | 0. | 1. | 0. | 1. |
| 3 | 0. | 1. | 0. | 1. |
| 4 | 0. | 0. | 0. | 0. |
| 5 | 1. | 0. | 1. | 1. |
| 6 | 0. | 1. | 0. | 1. |
| 7 | 1. | 1. | 0. | 1. |
| 8 | 1. | 0. | 0. | 1. |
| 9 | 0. | 1. | 1. | 0. |
| 10 | 0. | 0. | 0. | 1. |
| 11 | 1. | 1. | 0. | 1. |
| 12 | 1. | 1. | 1. | 0. |
| 13 | 0. | 0. | 0. | 1. |
| 14 | 1. | 0. | 0. | 1. |
| 15 | 1. | 1. | 1. | 1. |

* ОЦЕНКА РАЗЛИЧИЯ М СВЯЗ. ВЫБОРОК
 * Q - КРИТЕРИЙ КОКРЕНА * М УСЛОВИЙ ОПЫТА НАД N ОБЪЕКТАМИ
 * ДВУХВАРИАНТНЫЙ ХАРАКТЕР ОТВЕТОВ

PR370 - СПРАВЕДЛИВА ГИПОТЕЗА Н0 КОД ОШИБКИ = 0
 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ КРИТЕРИЯ КОКРЕНА = 7.1429
 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ РАСПРЕД. ХИ-КВАДРАТ = 7.8147

ТАБЛИЦА СФОРМИРОВАНА ИЗ ВЫБОРОК:
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 51
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 52
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 53
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 54


```

CC; SP370 - ОЦЕНКА РАЗЛИЧИЯ ПО КРИТЕРИЮ КОКРЕНА
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ОЦЕНКИ РАЗЛИЧИЯ "М" СВЯЗАННЫХ ВЫБОРОК
C Т. Е. РЕЗУЛЬТАТОВ "М" УСЛОВИЙ ОПЫТА НАД "N" ОБЪЕКТАМИ
C (ЕСЛИ ОТВЕТЫ НОСЯТ ДВУХВАРИАНТНЫЙ ХАРАКТЕР)
C ПО КРИТЕРИЮ КОКРЕНА
C
CR ЛИТЕРАТУРА ПО АЛГОРИТМУ : РУНИОН Р., СПРАВОЧНИК ПО
C НЕПАРАМЕТРИЧЕСКОЙ СТАТИСТИКЕ, М., "ФИНАНСЫ И
C СТАТИСТИКА", 1982
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММОФ,
C Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE SP370(KEY,NFWR,KSQN, L, N, M, T, Q
* , R, C, HIQ, THI, IND, IER, IERW)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CP ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C
C KEY I*4 * (8) МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 * 1 НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
C KSQN R*8 * (4,L) МАССИВ КОДОВ ВЫБОРОК
C L I*4 * 1 ДЛИНА "KSQN" /L=N ИЛИ L=1/
C N I*4 * 1 ЧИСЛО ОБЪЕКТОВ (СТРОК В "Т")
C M I*4 * 1 ЧИСЛО УСЛОВИЙ ОПЫТА (СТОЛБЦОВ В "Т")
C T R*4 * (N,M) ТАБЛИЦА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ
C Q R*4 * 1 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА /.005<=Q<=.5/
C R R*4 * (M) РАБОЧИЙ МАССИВ
C C R*4 * (N) РАБОЧИЙ МАССИВ
C HIQ R*4 * 1 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ КРИТЕРИЯ КОКРЕНА
C THI R*4 * 1 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ РАСПР. ХИ-КВАДРАТ
C IND I*4 * 1 ИНДЕКС ПРОВЕРКИ ПО КРИТЕРИЮ КОКРЕНА :
C IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ (НЕТ РАЗЛИЧИЯ);
C IND = 1 - НЕЛЕВАЯ ГИПОТЕЗА ОТКЛОНЯЕТСЯ.
C IER I*4 * 1 ВЫРАБОТАННЫЙ КОД ОШИБКИ
C IERW I*4 * (8) МАССИВ КОДОВ ОШИБОК, ПРИНЯТЫЙ ИЗ CW370
C
C DIMENSION KEY(8),IERW(8)
C * ,KSQN(4,L),T(N,M),R(M),C(N)
C REAL*8 NAME, KSQN
C DATA NAME/' SP370 '/, ZERO/0.0/, EPS/0.001/
C
C IND = 11
C IER = 0
C KLP=KEY(2)
C IF (KEY(1).EQ.0) GO TO 120
C CALL CW370(KEY,NFWR,KSQN,L,N,M,T,Q,IERW,MWC)
CL; CW370 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP370;
C АНАЛИЗ КОДОВ ОШИБОК "KF" И "MWC"
C KF = IERW(1)
C IF (MWC.LT.3) GO TO 120
C ЕСЛИ: MWC >= 3, TO IER=MWC
C IER = MWC
C IF (MWC.GE.3) GO TO 999
C ВЫЧИСЛЕНИЕ СУММ ЭЛЕМЕНТОВ ПО СТРОКАМ И СТОЛБЦАМ
C 120 DO 130 J=1,M
C Z = ZERO
C DO 125 I=1,N
C 125 Z = Z + T(I,J)
C R(J) - СУММА ЗНАЧЕНИЙ ЭЛЕМЕНТОВ J-ГО СТОЛБЦА
C 130 R(J) = Z

```

```

C      , DO 140 I=1,N
        Z = ZERO
        DO 135 J=1,M
C      C(I) - СУММА ЗНАЧЕНИЙ ЭЛЕМЕНТОВ I-й СТРОКИ
135      Z = Z + T(I,J)
140      C(I) = Z
C      Z - СУММА ЗНАЧЕНИЙ ЭЛЕМЕНТОВ ВСЕЙ ТАБЛИЦЫ
        Z = ZERO
        DO 150 J=1,M
150      Z = Z + R(J)
C      ВЫЧИСЛЕНИЕ ЗНАЧЕНИЯ КРИТЕРИЯ КОКРЕНА
        S = ZERO
        DO 230 J=1,M
230      S = S + R(J)*R(J)
        FM = FLOAT(M)
        G = (FM-1)*(FM*S-Z*Z)
        S = ZERO
        DO 250 I=1,N
250      S = S + C(I)*C(I)
        HIQ = G / (FM*Z-S)
C      ВЫЧИСЛЕНИЕ THI - КРИТИЧЕСКОГО ЗНАЧЕНИЯ HI2
        CALL MP030(Q, M-1, EPS, THI, IERTHI)
CL; MP030 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ РАСПР. ХИ-КВАДРАТ
        IF (IERTHI.EQ.0) GO TO 330
        IER = 4+IERTHI
        GOTO 810
C      ПРОВЕРКА СПРАВЕДЛИВОСТИ НУЛЕВОЙ ГИПОТЕЗЫ
330 IND=0
        IF (HIQ.GE.THI) IND=1
810      IF (KF.NE.3) GO TO 830
        IER = -IER
        GO TO 999
830      IF (KLP.EQ.0) GO TO 999
        KEY(1) = KLP
        CALL PR370(KEY,NFWR,KSQN,L,HIQ,THI,IND,IER)
CL; PR370 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP370.
999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0      02/02/88
      END

```

```

CS; CW370 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP370
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ КОНТРОЛЯ ВХОДНЫХ ДАННЫХ ДЛЯ SP370
C И ВЫДАЧИ В ФАЙЛ С НОМЕРОМ "NFWR" РЕЗУЛЬТАТА ПРОВЕРКИ
C; И ВХОДНЫХ ДАННЫХ
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP370
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE CW370(KEY,NFWR,KSQN,L,N,M,T,Q,IERW,MWC)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
C KSQN R*8 (4,L) 4A8 7 МАССИВ КОДОВ ВЫБОРОК
C L I*4 1 -- - ДЛИНА МАССИВА KSQN /L=M ИЛИ L=1/
C N I*4 1 I5 8 ЧИСЛО ОБЪЕКТОВ (СТРОК В "Т")
C M I*4 1 I5 8 ЧИСЛО УСЛОВИЙ (СТОЛБЦОВ В "Т")
C T R*4 (N,M) 6F8.0 8 ТАБЛИЦА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ
C Q R*4 1 F8.3 3 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
C IERW I*4 (8) 6I2 [2]ВЫРАБОТАННЫЙ КОД ОШИБКИ : 6 РАЗРЯДОВ
C;MWC I*4 1 I2 2 КОД МАКСИМАЛЬНОЙ ОШИБКИ
C
DIMENSION KEY(8),IERW(8),KSQN(4,L),T(N,M)
REAL *8 NAME, KSQN
DATA NAME/' CW370 '/, ZERO/0.0/, ONE/1.0/
C
KLU=KEY(1)
IF(KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
13 KLU=KLU/2
GO TO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF(KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
C
C КОНТРОЛЬ ВХОДНЫХ ДАННЫХ
C
C-* ПРОВЕРКА 'NFWR' ( 1-й РАЗРЯД )
C НАЧАЛЬНАЯ УСТАНОВКА : ОШИБОК НЕТ
KF = 0
C ЕСЛИ 'NFWR' БОЛЬШЕ 99 ИЛИ МЕНЬШЕ 1, ЭТО ГРУБАЯ ОШИБКА
IF (NFWR.GT.99) KF = 3
IF (NFWR.LT.1) KF = 3
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
IERW(1) = KF
C
C-* ПРОВЕРКА 'N' ( 3-й РАЗРЯД )
C KN=0 ОЗНАЧАЕТ, ЧТО 'N'- НОРМАЛЬНОЕ, Т.Е. >= 2
KN = 0

```

```

C  ЕСЛИ 'N' МЕНЬШЕ ЧЕМ 2, ЭТО ГРУБАЯ ОШИБКА
   IF (N.LT.2) KN = 3
C  ПО ВЕРХНЕМУ ПРЕДЕЛУ ДЛЯ 'N' ПРОВЕРКИ НЕТ
C  ФОРМИРОВАНИЕ КОДА ОШИБКИ
   IERW(3) = KN
C
C--  ПРОВЕРКА 'M' ( 4-й РАЗРЯД )
C  KM=0 ОЗНАЧАЕТ, ЧТО 'M' - НОРМАЛЬНОЕ, Т.Е. >= 2
   KM = 0
C  ЕСЛИ 'M' МЕНЬШЕ ЧЕМ 2, ЭТО ГРУБАЯ ОШИБКА
   IF (M.LT.2) KM = 3
C  ПО ВЕРХНЕМУ ПРЕДЕЛУ ДЛЯ 'M' ПРОВЕРКИ НЕТ
C  ФОРМИРОВАНИЕ КОДА ОШИБКИ
   IERW(4) = KM
C
C--  ПРОВЕРКА 'L' ( 2-й РАЗРЯД )
C  KL=0 ОЗНАЧАЕТ, ЧТО 'L' - НОРМАЛЬНОЕ, Т.Е. РАВНО 1
C  ИЛИ 'L' РАВНО 'M'
   KL = 3
C  ЕСЛИ 'L' РАВНО 'M': ТАБЛИЦА СФОРМИРОВАНА ИЗ ВЫБОРОК
   IF (L.EQ.M) KL = 0
C  ЕСЛИ 'L' РАВНО 1: ТАБЛИЦА НЕ СОСТАВЛЕНА ИЗ ВЫБОРОК
   IF (L.EQ.1) KL=0
C  ФОРМИРОВАНИЕ КОДА ОШИБКИ
   IERW(2) = KL
C
C--  ПРОВЕРКА 'T' ( 5-й РАЗРЯД )
C  KT=0 ОЗНАЧАЕТ, ЧТО 'T' - НОРМАЛЬНАЯ ТАБЛИЦА,
C  Т.Е. СУММА ЭЛЕМЕНТОВ В ЛЮБОМ СТРОКЕ НЕ= 0 И
C  СУММА ЭЛЕМЕНТОВ В ЛЮБОМ СТОЛБЦЕ НЕ= 0.
   KT = 11
C  ВОЗМОЖНА ЛИ ПРОВЕРКА 'T' ?
   IF (KN.GT.1 .OR. KM.GT.1) GO TO 70
C  ПРЕДЕЛЫ ДЛЯ ИНДЕКСОВ ПРАВИЛЬНЫЕ .
   KT = 0
   DO 63 I=1,N
     S = ZERO
     DO 61 J=1,M
       E = T(I,J)
       IF (E .EQ. ZERO) GO TO 61
       IF (E .EQ. ONE) GO TO 61
       GO TO 69
     61  S = S + E
       IF (S.EQ.ZERO .OR. S.EQ.FLOAT(M)) KT=1
   63  CONTINUE
     DO 67 J=1,M
       S = ZERO
       DO 65 I=1,N
         65  S = S + T(I,J)
           IF (S.EQ.ZERO .OR. S.EQ.FLOAT(N)) KT=1
       67  CONTINUE
         GO TO 70
     69  KT = 3
C  ФОРМИРОВАНИЕ КОДА ОШИБКИ
   70 IERW(5) = KT
C
C--  ПРОВЕРКА 'Q' ( 6-й РАЗРЯД )
C  НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'Q' : 0.025 <= Q <= 0.2
   KQ = 0
C  ЕСЛИ 'Q' БОЛЬШЕ ЧЕМ 0.2 ИЛИ МЕНЬШЕ ЧЕМ 0.025,
C  ТО СЧИТАЕМ ЧТО ДОПУЩЕНА ЛОГИЧЕСКАЯ ОШИБКА
   IF (Q.LT.0.025) KQ = 1
   IF (Q.GT.0.2) KQ = 1
C  ВВОДИМ ОГРАНИЧЕНИЕ : 0.005 <= Q <= 0.5
   IF (Q.LT.0.005) KQ = 3
   IF (Q.GT.0.5) KQ = 3

```

С ФОРМИРОВАНИЕ КОДА ОШИБКИ

IERW(6) = KQ

IERW(8) = 6

С

С ОПРЕДЕЛЕНИЕ МАКСИМАЛЬНОГО КОДА ОШИБКИ

MWC = MAX0 (KN,KM,KL,KT,KQ)

С

ШКАЛА КОДИРОВКИ ОШИБОЧНЫХ ПАРАМЕТРОВ :

С

СЕ РАЗРЯД КОДА ОШИБКИ I 1 I 2 I 3 I 4 I 5 I 6

С

С ПАРАМЕТР В ОБРАЩЕНИИ I NFWR I L I N I M I T I Q

С

С ВОЗМОЖНЫЕ ЗНАЧЕНИЯ I 0 I 0 I 0 I 0 I 0 I 0 I 0

С

I - I - I - I - I - I - I - I -

С

I - I - I - I - I - I - I - I -

С

I 3 I 3 I 3 I 3 I 3 I 3 I 3

С;

I - I - I - I - I - I - I - I -

С

С

С ПЕЧАТЬ ВХОДНЫХ ДАННЫХ

IF (KF.GT.1) GO TO 999

IF (KEY(2).EQ.0) GO TO 225

WRITE (NFWR,220) NAME, MWC

220 FORMAT(/' ',A8,4X,'КОД МАКСИМАЛЬНОЙ ОШИБКИ = ',I2)

225 IF(MWC.LT.1) GO TO 310

С

J=IERW(8)

WRITE (NFWR,120) (IERW(I),I=1,J)

120 FORMAT(/' ',8X,'ИМЕНА ПАРАМЕТРОВ : '

* ' ',NFWR, L, N, M, T, Q'

* /' ',8X,'К О Д Ы ОШИБОК : ',6(3X,I2)/)

С

310 IF (KEY(3).EQ.0) GOTO 710

WRITE (NFWR,340) Q

340 FORMAT(/' ',8X,'УРОВЕНЬ ЗНАЧИМОСТИ РИСКА = ',F8.3)

С

710 IF((KEY(7).EQ.0).OR.(L.EQ.1)) GOTO 810

WRITE (NFWR,750) ((KSN(I,J),I=1,4),J=1,L)

750 FORMAT(/' ',8X,'ТАБЛИЦА СФОРМИРОВАНА ИЗ ВЫБОРК: '

* /(' ',8X,5X,4A8))

С

810 IF (KEY(8).EQ.0) GOTO 999

WRITE (NFWR,860) N,M

860 FORMAT(/' ',8X,'ЗНАЧЕНИЯ ЭЛЕМЕНТОВ '

* ' ',ТАБЛИЦЫ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ : '

* /' ',8X

* ' ',СТРОКА (СТРОК = ',15,6X,'СТОЛБЦОВ = ',15,' ')/)

С

IF (KN.GT.1.OR.KM.GT.1) GO TO 999

DO 870 I=1,N

WRITE (NFWR,880) I,(T(I,J),J=1,M)

870 CONTINUE

880 FORMAT(' ',8X,I2,1X,6F8.0/(' ',11X,6F8.0))

С

999 RETURN

CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0

END

```

CC; PR370 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP370
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫВОДА В ФАЙЛ 'NFWR'
C РЕЗУЛЬТАТОВ РАБОТЫ П/П SP370
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP370
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ,
C Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE PR370(KEY,NFWR,KSQN,L,HIQ,THI,IND,IER)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CP ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
C KSQN R*8(4,L) 4A8 7 МАССИВ КОДОВ ВЫБОРОК
C L I*4 1 -- - ДЛИНА МАССИВА "KSQN" /L=N ИЛИ L=1/
C HIQ R*4 1 G12.5 3 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ КРИТЕР. КОКРЕНА
C THI R*4 1 G12.5 3 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ РАСПР. ХИ-КВАД.
C IND I*4 1 I1 2 ИНДЕКС ПРОВЕРКИ ПО КРИТЕРИЮ КОКРЕНА
C IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
C IND = 1 - НУЛЕВАЯ ГИПОТЕЗА ОТВЕРГАЕТСЯ.
C IER I*4 1 I3 2 КОД ОШИБКИ, ВЫРАБОТАННЫЙ SP370
C
DIMENSION KEY(8),KSQN(4,L)
REAL*8 TNK(3),TS1(4),TS2(4),TS3(4),NAME,KSQN
C
DATA TNK /' Q - КРИ', 'ТЕРИИ К', 'ОКРЕНА' /
* ,TS1 /'ОЦЕНКА Р', 'АЗЛИЧИЯ', 'М СВЯЗ.', 'ВЫБОРОК' /
* ,TS2 /'М УСЛОВИ', 'Я ОПЫТА', 'НАД Н О', 'Б"ЕКТАМИ' /
* ,TS3 /'ДВУХВАРИ', 'АНТНЫЙ', 'ХАРАКТЕР', 'ОТВЕТОВ' /
C
DATA NAME/' PR370' /
IF (NFWR.LE.0.OR.NFWR.GE.100) GO TO 999
C
KLU=KEY(1)
IF (KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
KLU=KLU/2
13 GOTO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF (KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
C
IF (KEY(1).EQ.0) GO TO 210
C
WRITE (NFWR,120) TS1,TNK,TS2,TS3
120 FORMAT(/ 1H ,26(1H*)
* /2H *,24(1H ),3H* ,4A8
* /2H *,3A8 ,3H* ,4A8
* /2H *,24(1H ),3H* ,4A8
* / 1H ,26(1H*) )

```

```

C
210 IF (KEY(2) .EQ. 0) GO TO 310
    WRITE (NFWR,230) NAME, IND, IER
230 FORMAT(/' ',A8,' - СПРАВЕДЛИВА ГИПОТЕЗА Н',I1
*          ', ' ', ' ;      КОД ОШИБКИ =' ,I3)
C
310 IF (KEY(3) .EQ. 0) GO TO 710
    WRITE (NFWR,340) C, D, Z
340 FORMAT(/' ',13X,'КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ =' ,G20.5
*          /' ',13X,'СТАНДАРТНОЕ ОТКЛОНЕНИЕ =' ,G20.5
*          /' ',13X,'СТАТИСТИКА      =' ,G20.5)
C
710 IF (KEY(7) .EQ. 0) GO TO 999
    WRITE (NFWR,760)
760 FORMAT(' ',20X,'К О Д Ы   В Ы Б О Р О К   : ')
C
    DO 770 J=1,2
770   WRITE (NFWR,775) J, (KSON(I,J), I=1,4)
775   FORMAT(' ',8X,I2,' - ',4A8)
C
999 RETURN
CV: ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
    END

```

§ 6. Оценка корреляционной связи двух выборок

В этой задаче различают два случая. Первый — обе выборки есть результат измерений в номинальных шкалах, второй — одна выборка есть результат измерения в номинальной шкале, а другая измерена в шкале отношений (результат количественных измерений). В первом случае для оценки корреляционной связи двух выборок используют коэффициент ассоциации, тетрафорический коэффициент корреляции, коэффициенты Чупрова и Крамера, во втором — бисериальный и точечный бисериальный коэффициенты корреляции.

6.1. Коэффициент ассоциации Юла K_a . В случае измерений с двумя градациями ($k = 2$) коэффициент ассоциации Юла вычисляется как отношение

$$K_a = \frac{ad - bc}{ad + bc},$$

где a, b, c, d — абсолютные значения частот (долей) в таблице «2 × 2»:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

в которую сводятся результаты измерений.

При достаточно большой суммарной частоте величина K_a имеет нормальное распределение с ошибкой

$$\sigma_{K_a} = \frac{1 + K_a^2}{2} \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}.$$

Доверительный интервал для «истинного» значения связи определяется как $K_a \pm t_{q,v} \sigma_{K_a}$; $v = n - 2$, при $n \leq 20$, или $K_a \pm Z \sigma_{K_a}$ при $n > 20$, где Z — критическое значение нормированного нормального распределения при уровне значимости риска q ; $t_{q,v}$ — критическое значение распределения Стьюдента.

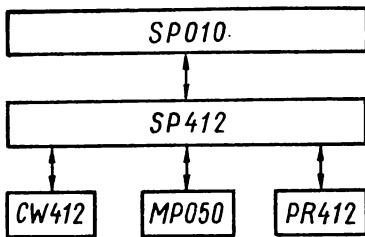


Рис. 58

Если доверительный интервал накрывает значение 0, то нуль-гипотеза $H_0 : K_a = 0$ не отклоняется с доверительной вероятностью $1 - q$.

6.2. Программная реализация. Оценка корреляционной связи двух выборок — результатов измерений в номинальных шкалах с двумя градациями может быть представлена в виде блок-схемы на рис. 67. На рисунках $SP010$ — модуль управления комплек-

са; $SP412$ — подпрограмма вычисления коэффициента ассоциации Юла и проверки нуль-гипотезы; $CW412$ — подпрограмма контроля и печати входных данных для $SP412$; $PR412$ — подпрограмма печати результатов работы $SP412$; $MP050$ — подпрограмма вычисления критического значения нормального распределения.

Тексты подпрограмм $SP412$, $CW412$, $PR412$ приведены после примера.

6.3. Пример. Для подготовки задания в программу «СОКИ-Н» необходимо ввести $IRP = 006$ или $IRP = 016$, если данные — экспериментальные выборки и $IRP = 106$ или 116 , если данные — сформированная исследователем таблица мультиномиальных частот. В нашем случае данные — выборки, поэтому задано $IRP = 006$. Количество градаций обеих выборок равно двум по умолчанию.

Необходимо также задать:

- коды выборок и их длины;
- форматы данных и сами массивы данных;
- массивы значений уровней градаций;
- уровень значимости риска;
- ключи печати.


```

CC; SP412 - КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ (АССОЦИАЦИИ) ЮЛА
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫЧИСЛЕНИЯ СТЕПЕНИ КОРРЕЛЯЦИОННОЙ
C СВЯЗИ МЕЖДУ 2-Я СОВОКУПНОСТЯМИ ПРИЗНАКОВ, ИМЕЮЩИХ
C ТОЛЬКО ПО 2 ГРАДАЦИИ (КОЭФФИЦИЕНТ АССОЦИАЦИИ ЮЛА)
C
CR ЛИТЕРАТУРА ПО АЛГОРИТМУ : М.ДЖ.КЕНДАЛ, А.СТЬЮАРТ
C СТАТИСТИЧЕСКИЕ ВЫВОДЫ И СВЯЗИ, ТОМ 2, М., НАУКА, 1973.
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММО,
C Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE SP412(KEY,NFWR,KSQN,L,T,N,Q,C,D,S,TZ
* ,IND,IER,IERW)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CF ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C
C KEY I*4 * (8) МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 * 1 НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<NFWR<=99/
C KSQN R*8 * (4,L) МАССИВ К О Д О В ВЫБОРОК
C L I*4 * 1 ДЛИНА "KSQN" /L=1 ИЛИ L=2/
C T R*4 * (2,2) ТАБЛИЦА МУЛЬТИНОМИНАЛ. ЧАСТОТ ПРИЗНАКОВ
C N I*4 * 1 СУММАРНЫЙ ОБЪЕМ ВЫБОРОК / N >= 4 /
C Q R*4 * 1 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА /.005<=Q<=.5 /
C C R*4 * 1 ВЫЧИСЛ. КОЭФФИЦИЕНТ КОРРЕЛЯЦ. СВЯЗИ
C D R*4 * 1 ПОЛУШИРИНА ДОВЕРИТЕЛЬНОГО ИНТЕРВАЛА
C S R*4 * 1 СТАНДАРТ. ОШИБКА ВЫБОРОЧН. КОЭФФИЦИЕНТА
C TZ R*4 * 1 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ РАСПРЕДЕЛЕНИЯ
C IND I*4 * 1 ИНДЕКС ОТВЕТА :
C IND = 1, ЕСЛИ "0" НЕ ВКЛЮЧ. В ДОВЕРИТ. ГРАНИЦЫ(ЕСТЬ КОРР.);
C IND = 0, ЕСЛИ "0" ВКЛЮЧАЕТСЯ В ДОВЕРИТ. ГРАНИЦЫ (НЕТ КОРР.)
C IER I*4 * 1 ВЫРАБОТАННЫЙ КОД ОШИБКИ
C IERW I*4 * (8) МАССИВ КОДОВ ОШИБОК, ПРИНЯТЫХ ИЗ SW412
C
C DIMENSION KEY(8),IERW(8),T(2,2),KSQN(4,L)
C REAL*8 NAME,KSQN
C DATA NAME /' SP412 '/, ZERO/0.0/, EPS /0.001/
C * - HALF/0.5/, ONE/1.0/
C
C IND = 11
C IER = 0
C KLP = KEY(2)
C IF (KEY(1).EQ.0) GO TO 120
C CALL SW412(KEY,NFWR,KSQN,L,T,N,Q,IERW,MWC)
CL; SW412 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ SP412;
C
C АНАЛИЗ КОДОВ ОШИБОК : "KF" И "MWC"
C
C KF = IERW(1)
C IF (MWC.LT.3) GO TO 120
C ЕСЛИ "MWC" >= 3, ТО IER = MWC
C IER = MWC
C IF (MWC.GE.3) GO TO 999
C
C 120 C = (T(1,1)*T(2,2)-T(1,2)*T(2,1))
C * / (T(1,1)*T(2,2)+T(1,2)*T(2,1))
C S = (ONE-C*C)*HALF*SQRT(ONE/T(1,1)
C * + ONE/T(1,2)+ONE/T(2,1)+ONE/T(2,2))
C ЕСЛИ L=1, ТО ДЛИНА ВЫБОРОК НЕИЗВЕСТНА -
C - "TZ" И "IND" НЕ ВЫЧИСЛЯЕТСЯ
C IF ( L .EQ. 2 ) GO TO 210

```

```

      TZ = ZERO
      D = ZERO
      GO TO 810
210  IF ( N .GT. 20 ) GOTO 250
C
      CALL MP010(Q,N-2,EPS,TZ,IERT)
CL; MP010 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ Т-РАСПРЕДЕЛЕНИЯ;
      IF ( IERT .EQ. 0 ) GOTO 300
      IER = 4+IERT
      GOTO 810
250  V = ONE-Q*HALF
      CALL MP050(0,V,EPS,U,TZ,IERT)
CL; MP050 - ВЫЧИСЛ. КРИТИЧ. ЗНАЧЕНИЙ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ;
      IF ( IERZ .EQ. 0 ) GOTO 300
      IER = 4+IERZ
      GOTO 810
C  ПРОВЕРКА ЗНАЧИМОСТИ КОЭФФИЦИЕНТА КОРРЕЛЯЦИОННОЙ СВЯЗИ
300  CONTINUE
      D = TZ*S
      IND = 0
      IF ( (ABS(C)-ABS(D)) .LE. ZERO ) IND=1
C
810  IF (KF .NE. 3) GO TO 830
      IER = -IER
      GO TO 999
830  IF (KLP .EQ. 0) GOTO 999
      KEY(1) = KLP
      CALL PR412(KEY,NFWR,KSQN,L,C,D,S,TZ,IND,IER)
CL; PR412 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP412.
999  RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0          02/02/88
      END

```

```

C; CW412 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP412
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ КОНТРОЛЯ ВХОДНЫХ ДАННЫХ ДЛЯ SP412
C И ВЫДАЧИ В ФАЙЛ С НОМЕРОМ "NFWR" РЕЗУЛЬТАТА ПРОВЕРКИ
C; И ВХОДНЫХ ДАННЫХ
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP412
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММОФ
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE CW412(KEY,NFWR,KSQN,L,T,N,Q,IERW,MWC)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
C ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
C KSQN R*8 (4,L) 4A8 7 МАССИВ К О Д О В ВЫБОРОК
C L I*4 1 -- - ДЛИНА "KSQN" /L=1 ИЛИ L=2/
C T R*4 (2,2) G20.6 8 ТАБ.МУЛЬТИНОМИН. ЧАСТОТ ПРИЗНАКОВ
C N I*4 1 I5 7 СУММАРНЫЙ ОБЪЕМ ВЫБОРОК /N>=4/
C Q R*4 1 F8.3 3 УРОВЕНЬ ЗНАЧ. РИСКА /.005<=Q<=.5/
C IERW I*4 (8) S12 [2]ВЫРАБОТАННЫЙ КОД ОШИБКИ : 5 РАЗР.
C; MWC I*4 1 I1 2 КОД МАКСИМАЛЬНОЙ ОШИБКИ
C
DIMENSION KEY(8),IERW(8),T(2,2),KSQN(4,L)
REAL*8 NAME, KSQN
DATA NAME /' CW412 '/, ZERO/0.0/, TMAX/1.0E18/
C
KLU=KEY(1)
IF(KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
13 KLU=KLU/2
GOTO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF(KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
C
C КОНТРОЛЬ ВХОДНЫХ ДАННЫХ
C
C-- ПРОВЕРКА 'NFWR' ( 1-й РАЗРЯД )
C НАЧАЛЬНАЯ УСТАНОВКА : ОШИБОК НЕТ
KF = 0
C ЕСЛИ 'NFWR' БОЛЬШЕ 99 ИЛИ МЕНЬШЕ 1, ЭТО ГРУБАЯ ОШИБКА
IF (NFWR.GT.99) KF = 3
IF (NFWR.LT.1) KF = 3
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
IERW(1) = KF
C
C-- ПРОВЕРКА 'T' ( 2-й РАЗРЯД )
C КТ=0 ОЗНАЧАЕТ, ЧТО 'T' - НОРМАЛЬНАЯ ТАБЛИЦА,
C Т.Е. ЭЛЕМЕНТЫ 'T' БОЛЬШЕ НУЛЯ
KT = 0

```

```

      I=0 23 I=1,2
      DO 23 J=1,2
        IF (T(I,J) .GT. TMAX) GO TO 25
        IF (T(I,J) .LE. ZERO) GO TO 25
      23 CONTINUE
      GO TO 30
      25 KT = 3
      C ФОРМИРОВАНИЕ КОДА ОШИБКИ
      30 IERW(2) = KT
      C
      C-#- ПРОВЕРКА 'L' ( 3-й РАЗРЯД )
      KL = 3
      IF (L.EQ.1 .OR. L.EQ.2) KL = 0
      C ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(3) = KL
      C
      C-#- ПРОВЕРКА 'N' ( 4-й РАЗРЯД )
      KN = 0
      IF (N .LT. 4 .OR. N .GT. 32000) KN = 3
      C ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(4) = KN
      C
      C-#- ПРОВЕРКА 'Q' ( 5-й РАЗРЯД )
      C НОРМАЛЬНОЕ ЗНАЧЕНИЕ "Q" : 0.025 <= Q <= 0.2
      KQ = 0
      C ЕСЛИ "Q" БОЛЬШЕ ЧЕМ 0.2 ИЛИ МЕНЬШЕ ЧЕМ 0.025,
      C ТО СЧИТАЕМ ЧТО ДОПУЩЕНА ЛОГИЧЕСКАЯ ОШИБКА
      IF (Q .LT. 0.025) KQ = 1
      IF (Q .GT. 0.2) KQ = 1
      C ВВОДИМ ОГРАНИЧЕНИЕ : 0.005 <= Q <= 0.5
      IF (Q .LT. 0.005) KQ = 3
      IF (Q .GT. 0.5) KQ = 3
      C ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(5) = KQ
      IERW(8) = 5
      C
      C ОПРЕДЕЛЕНИЕ МАКСИМАЛЬНОГО КОДА ОШИБКИ
      MWC = MAX0( KT, KL, KN, KQ )
      C
      C ШКАЛА КОДИРОВКИ ОШИБОЧНЫХ ПАРАМЕТРОВ
      C
      C SE РАЗРЯД КОДА ОШИБКИ I 1 I 2 I 3 I 4 I 5
      C -----
      C ПАРАМЕТР В ОБРАЩЕНИИ I NFWR I T I L I N I Q
      C -----
      C ВОЗМОЖНЫЕ ЗНАЧЕНИЯ I 0 I 0 I 0 I 0 I 0
      C I - I - I - I - I 1
      C I - I - I - I - I -
      C I 3 I 3 I 3 I 3 I 3
      C
      C ПЕЧАТЬ ВХОДНЫХ ДАННЫХ
      IF (KF .GT. 0) GOTO 999
      IF (KEY(2) .EQ. 0) GOTO 225
      C
      WRITE (NFWR,220) NAME, MWC
      220 FORMAT(' ',A8,4X,'КОД МАКСИМАЛЬНОЙ ОШИБКИ =',I1)
      C
      225 IF (MWC .LT. 1) GOTO 310
      C
      J=IERW(8)
      WRITE (NFWR,227) (IERW(I),I=1,J)
      227 FORMAT(/' ',8X,'ИМЕНА ПАРАМЕТРОВ :',
      * , NFWR, T, L, N, Q'
      * /' ',8X,'К О Д Ы ОШИБОК :',5(3X,I2)/)
      C
      310 IF (KEY(3) .EQ. 0) GOTO 710

```

```

      WRITE (NFWR,330) Q
330  FORMAT(/' ',8X,'УРОВЕНЬ ЗНАЧИМОСТИ РИСКА =' ,F8.3)
*
710  IF (KEY(7).EQ.0) GOTO 999
      IF (L .EQ. 1) GOTO 999
      WRITE (NFWR,770) ((KSGN(I,J),I=1,4),J=1,2)
770  FORMAT(/' ',8X,'ТАБЛИЦА ДАННЫХ СФОРМИРОВАНА ИЗ ВЫБОРОК :')
      *
      *      /' ',8X,'1 - ',4A8
      *      /' ',8X,'2 - ',4A8)
C
      WRITE (NFWR,775) N
775  FORMAT(/' ',13X,'СУММА ДЛИН ВЫБОРОК =' ,I5)
C
810  IF (KEY(8) .EQ. 0) GOTO 999
C
      WRITE(NFWR,880) ((T(I,J),J=1,2),I=1,2)
880  FORMAT(/' ',8X,'ТАБЛИЦА (2*2) ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ :')
      *
      *      /(' ',8X,2G20.6))
999  RETURN
CV1  ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
      END

```

```

CC; PR412 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP412
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫВОДА В ФАЙЛ 'NFWR'
C; РЕЗУЛЬТАТОВ РАБОТЫ П/П SP412
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP412
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE PR412(KEY,NFWR,KSQN,L,C,D,S,TZ,IND,IERY
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C - ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
C Ksqn R*8(4,L) 4A8 7 МАССИВ К О Д О В ВЫБОРОК
C L I*4 1 -- - ДЛИНА "Ksqn" /L=1 ИЛИ L=2/
C C R*4 1 615.5 3 ВЫЧИСЛ. КОЭФФИЦИЕНТ КОРРЕЛЯЦ. СВЯЗИ
C D R*4 1 615.5 [3]ПОЛУШИРИНА ДОВЕРИТЕЛЬНОГО ИНТЕРВАЛА
C S R*4 1 615.5 [3]СТАНДАРТНАЯ ОШИБКА КОЭФФ. КОРР.СВЯЗИ
C TZ R*4 1 615.5 [3]КРИТИЧЕСКОЕ ЗНАЧЕНИЕ РАСПРЕДЕЛЕНИЯ
C IND I*4 1 11 2 ИНДЕКС ОТВЕТА :
C IND = 1 - ЕСЛИ "0" НЕ ВКЛ. В ДОВЕРИТЕЛ. ГРАНИЦЫ(ЕСТЬ КОРР.)
C IND = 0 - ЕСЛИ "0" ВКЛ. В ДОВЕРИТЕЛЬНЫЕ ГРАНИЦЫ(НЕТ КОРР.).
C IERY I*4 1 13 2 КОД ОШИБКИ, ВЫРАБОТАННЫЙ SP412
C
DIMENSION KEY(8),Ksqn(4,L)
REAL*8 NAME, Ksqn
REAL*8 TNK(3),TS1(4),TS2(4),TS3(4)
DATA TNK / ' КОЭФФ. ', 'АССОЦИАЦ', 'ИИ ЮЛА ' /
* ,TS1 / ' ВЫЧИСЛЕ', 'НИЕ КОЭ', 'ФФ. КОРР.', ' СВЯЗИ ' /
* ,TS2 / ' МЕЖДУ ', ' ДВУМЯ ', 'СОВОКУПН', 'ОСТЯМИ ' /
* ,TS3 / ' для 2-Х ', ' ПРИЗНАКО', 'В ПО 2 ', 'ГРАДАЦИИ' /
C
DATA NAME / ' PR412 ' /
C
KLU=KEY(1)
IF(KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
13 KLU=KLU/2
GO TO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF(KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
C
IF (KEY(1).EQ. 0) GO TO 210
IF ( NFWR.LE.0 .OR. NFWR.GE.100 ) GO TO 999
C
WRITE (NFWR,120) TS1,TNK,TS2,TS3
120 FORMAT(/ 1H ,26(1H*))
* /2H *,24(1H ),3H* ,4A8
* /2H *,3A8 ,3H* ,4A8

```

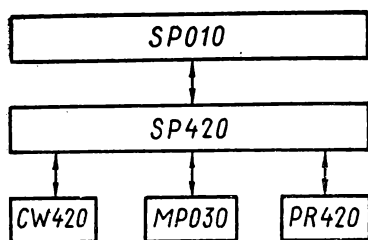



Рис. 59

6.5. Программная реализация.

Программа вычисления коэффициентов Чупрова и Крамера может быть представлена в виде блок-схемы (рис. 59).

Назначение подпрограмм:

SP420 — вычисление коэффициента корреляции Чупрова — Крамера и оценка значимости статистической связи;

CW420 — контроль и печать входных значений для *SP420*;

PR420 — печать результатов работы *SP420*;

MP030 — вычисление критических значений распределения χ^2 -квadrата. Тексты подпрограмм *SP420*, *CW420*, *PR420* приведены ниже, после примера.

6.6. Пример. Проведем вычисление коэффициента Чупрова — Крамера с целью оценки значимости статистической связи между двумя выборками — результатами измерений в номинальных шкалах при числе градаций $K > 2$. Приведем задание «входной поток» для программы «СОКИ-Н» на нахождение корреляционной связи между двумя выборками с числом градаций $K = 4$.

Необходимо задание следующих параметров и массивов:

IRP-006 или *IRP-026* — нахождение корреляционной связи (данные — выборки); $K = 04$ — количество градаций выборки 1; $m = 04$ — количество градаций выборки 2. Остальные параметры и данные задаются, как и в предыдущем примере.

Возможно задание не самих выборок, а уже сформированной исследователем матрицы мультиномиальных частот. В этом случае необходимо задать *IRP-106* или *IRP-126* и матрицу частот вместо элементов выборки и уровней градаций, а также отключить ключи печати для *SP120*. Второй пример иллюстрирует работу программы «СОКИ-Н» в случае задания таблицы мультиномиальных частот « 4×4 ».

**ЗАДАНИЕ НА ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТА КОРРЕЛЯЦИОННОЙ СВЯЗИ
ДВУХ ВЫБОРОЧНЫХ СОВОКУПНОСТЕЙ, ИЗМЕРЕННЫХ В
НОМИНАЛЬНЫХ ШКАЛАХ. ДАННЫЕ - ВЫБОРКИ**

```

=====  ВЫЧИСЛЕНИЕ КОРРЕЛЯЦИОННОЙ СВЯЗИ ДВУХ ВЫБОРОК  120
****  МАКЕТ ПОДГОТОВКИ ДАННЫХ ДЛЯ IR=006:  ****SP420****
IRP -006 (I3)- КОД ФУНКЦИИ
NS -02 /ПО УМОЛЧАНИЮ/ (I2)- КОЛИЧЕСТВО ВЫБОРОК
K -04 (I2)- КОЛИЧЕСТВО ГРАДАЦИЙ ВЫБОРКИ 1
M -04 (I2)- КОЛИЧЕСТВО ГРАДАЦИЙ ВЫБОРКИ 2
(I2)-НОМЕР; (4A8)- К О Д ВЫБОРКИ; (I4)-ДЛИНА
XX XXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXX
KSNQ-01 ТЕСТОВАЯ /4 ГРАДАЦИИ/ ВЫБОРКА 61 0100
KSNQ-02 ТЕСТОВАЯ /4 ГРАДАЦИИ/ ВЫБОРКА 62 0100
(I2, I3, I2)C 4A8 J:НОМЕР ВЫБ.,КОД,КАНАЛ,ИМЯ ФАЙЛА
FILX-01 000 08-INF:DATA61.DAT
C СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) J
FORX-01 (4X,30I2)
1 1 1 1 2 2 2 4 3 2 2 4 2 1 3 4 3 2 3 4 3 1 3 4 3 2 3 4 4
3 1 3 3 2 3 3 4 3 3 3 4 3 1 3 4 3 2 3 4 3 1 3 4 3 2 3 4 4
3 1 3 3 2 3 3 4 3 3 3 4 3 1 3 4 3 2 3 4 3 1 3 4 3 2 3 4 4
3 1 3 3 2 3 3 4 3
(I2, I3, I2)C 4A8 J:НОМЕР ВЫБ.,КОД,КАНАЛ,ИМЯ ФАЙЛА
FILX-02 000 08-INF:DATA5.DAT
C СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) J
FORX-02 (4X,30I2)
1 1 1 1 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
4 4 4 4 4 4 4 4 4 4
(I2, I3, I2)C 4A8 J:НОМЕР ВЫБ.,КОД,КАНАЛ,ИМЯ ФАЙЛА
FILC-01 000 08-INF:DATA64.DAT
C СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) J
FORC-01 (4X,9I2)
1 2 3 4 5 6 7 8 9
(I2, I3, I2)C 4A8 J:НОМЕР ВЫБ.,КОД,КАНАЛ,ИМЯ ФАЙЛА
FILC-02 000 08-INF:DATA65.DAT
C СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) J
FORC-02 (4X,9I2)
1 2 3 4 5 6 7 8 9
Q - 0.05 <- (014.7)- УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
NFWR-02 (I2) - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
(I1, I2)-ТИП КЛЮЧЕЙ ПЕЧАТИ: 0-ПОРАЗРЯДНЫЙ; И ИХ КОЛИЧЕСТВО
KEYC-0 04 1-ИНТЕРВАЛЬНЫЙ.
ИНДЕКСЫ/ИНТЕРВАЛЬНОЕ ЗНАЧЕНИЕ КЛЮЧА ПЕЧАТИ/ПОРАЗРЯДНОЕ/
12345678 - РАЗРЯДЫ КЛЮЧА ПЕЧАТИ :
KEY - 1 2 1/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SP120 ;
KEY - 2 1/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP120 ;
KEY - 1 2 4/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SP420 ;
KEY - 2 2 4/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP420 .
MD -20 (I2)- МАКСИМАЛЬНАЯ ВЫСОТА ДИАГРАММЫ
(I1)-РАСПРЕДЕЛЕНИЕ ПАМЯТИ: =1-ВЫДАЕТСЯ ВСЕГДА;
END -1 КОНЕЦ ДАННЫХ; =0-ВЫДАЕТСЯ ТОЛЬКО АВАРИЙНО.

```

РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ ЗАДАНИЯ (РАСПЕЧАТКА РЕШЕНИЯ)

```

IRP - 6
K - 4
M - 4
KSNQ NN XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX NNNN
KSNQ- 1 ТЕСТОВАЯ /2 ГРАДАЦИЙ/ ВЫБОРКА 61 100
KSNQ- 2 ТЕСТОВАЯ /2 ГРАДАЦИЙ/ ВЫБОРКА 62 100
FILX- 1 0 8-INF:DATA61.DAT
FORX- 1:(4X,30I2)
1 1 1 1 1 2 2 2 4 3 2 2 4 2 1 3 4 3 2 3 4 3 1 3 4 3 2 3 4 4
3 1 3 3 3 2 3 3 4 3 3 3 4 3 1 3 4 3 2 3 4 3 1 3 4 3 2 3 4 4
3 1 3 3 3 2 3 3 4 3 3 3 4 3 1 3 4 3 2 3 4 3 1 3 4 3 2 3 4 4
3 1 3 3 3 2 3 3 4 3
FILX- 2 0 8-INF:DATA5.DAT
FORX- 2:(4X,30I2)
1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
4 4 4 4 4 4 4 4 4 4
FILC- 1 0 8-INF:DATA4.DAT
FORC- 1:(4X,9I2)
1 2 3 4
FILC- 2 0 8-INF:DATA5.DAT
FORC- 2:(4X,9I2)
1 2 3 4
Q 0.5000000E-01
NFWR 2
KEYC 0 4 4
KEY 1 2 1/8/11111111/
KEY 2 2 1/8/11111111/
KEY 1 2 4/8/11111111/
KEY 2 2 4/8/11111111/
ОБЪЕМ РАБОЧЕЙ ПАМЯТИ - 4000 ЕДИНИЦ INTEGER
ТРЕБУЕМЫЙ ОБЪЕМ ПАМЯТИ - 449 ЕДИНИЦ INTEGER

```

ИНДЕКСЫ НАЧАЛА ПЕРЕМЕННЫХ В РАБОЧЕЙ ПАМЯТИ :

| | | | | | | | |
|-----|---|------|-----|------|-----|------|-----|
| 1 - | 1 | 6 - | 6 | 11 - | 107 | 16 - | 175 |
| 2 - | 2 | 7 - | 7 | 12 - | 307 | 17 - | 183 |
| 3 - | 3 | 8 - | 36 | 13 - | 315 | 18 - | 420 |
| 4 - | 4 | 9 - | 19 | 14 - | 158 | 19 - | 423 |
| 5 - | 5 | 10 - | 105 | 15 - | 174 | 20 - | 426 |

CH110 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

```

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА Q = 0.050
КОД ВЫБОРКИ : ОБЪЕМ :
ТЕСТОВАЯ /2 ГРАДАЦИЙ/ ВЫБОРКА 61 100
КОЛИЧЕСТВО ГРАДАЦИЙ - K = 4
ЗНАЧЕНИЯ ЭЛ. ВЫБОРОЧНОЙ СОВОКУПНОСТИ
1 1 1 1 1 2 2 4 3
2 2 4 2 1 3 4 3 2 3
4 3 1 3 4 3 2 3 4 4
3 1 3 3 3 2 3 3 4 3
3 3 4 3 1 3 4 3 2 3
4 3 1 3 4 3 2 3 4 4
3 1 3 3 3 2 3 3 4 3
3 3 4 3 1 3 4 3 2 3
4 3 1 3 4 3 2 3 4 4
3 1 3 3 3 2 3 3 4 3

```

МАССИВ ЗНАЧЕНИЙ УРОВНЕЙ:

1 2 3 4

SW110 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА $\alpha = 0.050$

К О Д В Ы Б О Р К И :

ОБЪЕМ :

ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА 62

100

КОЛИЧЕСТВО ГРАДАЦИЙ - К = 4

ЗНАЧЕНИЯ ЭЛ. ВЫБОРОЧНОЙ СОВОКУПНОСТИ

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 |
| 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 |
| 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 |
| 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |

МАССИВ ЗНАЧЕНИЙ УРОВНЕЙ:

1 2 3 4

* * * * *

ВЫЧИСЛЕНИЕ

* СТАТ.ОБРАБОТКА ВЫБОРОК *

СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК

* * * * *

ДВУХ ВЫБОРОЧНЫХ СОВОКУПНОСТЕЙ

PR120 - ВЫЧИСЛЕННЫЕ ХАРАКТЕРИСТИКИ: ВЫБОРКА 1 ВЫБОРКА 2

СПРАВЕДЛИВА ГИПОТЕЗА

H1

H1

КОД ОШИБКИ

0

0

ПОКАЗАТЕЛЬ ИЗМЕНЧИВОСТИ -

0.218

0.163

КОЭФФИЦИЕНТ ВАРИИРОВАНИЯ -

0.873

0.651

КВАДРАТИЧЕСКАЯ ОШИБКА ДОЛИ -

0.218E-01

0.163E-01

МОДА

0.490

0.600

ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ H12 -

32.2

75.3

КРИТИЧЕСКОЕ ЗНАЧЕНИЕ H12 -

7.81

7.81

К О Д Ы В Ы Б О Р О К

ОБЪЕМЫ ГРАДАЦИЙ

1-ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА 61

N1- 100 K1- 4

2-ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА 62

N2- 100 K2- 4

МАССИВ ЗНАЧЕНИЯ ВЫЧИСЛЕННЫХ ДОЛЕЙ 1-й ВЫБОРКИ :

0.140

0.150

0.490

0.220

МАССИВ ЗНАЧЕНИЯ ВЫЧИСЛЕННЫХ ДОЛЕЙ 2-й ВЫБОРКИ :

0.500E-01

0.900E-01

0.600

0.260

ТАБЛИЦА МУЛЬТИНОМИНАЛЬНЫХ ЧАСТОТ :

| | | | | |
|----|----|----|-----|-----|
| 1: | 5. | 0. | 6. | 3. |
| 2: | 0. | 6. | 6. | 3. |
| 3: | 0. | 1. | 34. | 14. |
| 4: | 0. | 2. | 14. | 6. |

SW420 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА = 0.050

ТАБЛИЦА СФОРМИРОВАНА ИЗ ВЫБОРОК:

ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА 61

ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА 62

СТРОК (ГРАДАЦИЯ КАЧЕСТВЕННОГО ПРИЗНАКА "А") = 4
 СТОЛБЦОВ (ГРАДАЦИИ КАЧЕСТВЕННОГО ПРИЗНАКА "Б") = 4

ТАБЛИЦА МУЛЬТИНОМИНАЛЬНЫХ ЧАСТОТ ПРИЗНАКОВ "А" И "Б":

| | | | | |
|---|----|----|-----|-----|
| 1 | 5. | 0. | 6. | 3. |
| 2 | 0. | 6. | 6. | 3. |
| 3 | 0. | 1. | 34. | 14. |
| 4 | 0. | 2. | 14. | 6. |

 * ВЫЧИСЛЕНИЕ КОЭФФ. КОРР. СВЯЗИ
 * КОЭФФ. ЧУПРОВА/КРАМЕРА * МЕЖДУ ДВУМЯ СОВОКУПНОСТЯМИ
 * В ТАБЛИЦЕ СОПРЯЖЕННОСТИ R*C

PR420 - СПРАВЕДЛИВА ГИПОТЕЗА H1 КОД ОШИБКИ = 0

КОЭФФИЦИЕНТ КОРРЕЛЯЦИОННОЙ СВЯЗИ 126.74
 СТАНДАРТНАЯ ОШИБКА 0.33333E-02
 КРИТ. ЗНАЧЕНИЕ РАСПРЕД. ХИ-КВАД. 16.919

ТАБЛИЦА СФОРМИРОВАНА ИЗ ВЫБОРОК:
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 61
 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 62

ЗАДАНИЕ НА ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТА КОРРЕЛЯЦИОННОЙ СВЯЗИ
 ДВУХ ВЫБОРОЧНЫХ СОВОКУПНОСТЕЙ, ИЗМЕРЕННЫХ
 В НОМИНАЛЬНЫХ ШКАЛАХ. ДАННЫЕ - ТАБЛИЦА
 МУЛЬТИНОМИНАЛЬНЫХ ЧАСТОТ ПРИЗНАКОВ

----- ВЫЧИСЛЕНИЕ КОРРЕЛЯЦИОННОЙ СВЯЗИ ДВУХ ВЫБОРОК
 **** МАКЕТ ПОДГОТОВКИ ДАННЫХ ДЛЯ IR=106: ****SP420****
 IRP -106 Q (I3)- КОД ФУНКЦИИ
 MS -02 /ПО УМОЛЧАНИЮ/ (I2)- КОЛИЧЕСТВО ВЫБОРОК
 K -04 (I2)- КОЛИЧЕСТВО ГРАДАЦИИ ВЫБОРКИ 1
 M -04 (I2)- КОЛИЧЕСТВО ГРАДАЦИИ ВЫБОРКИ 2
 (I2)-НОМЕР; (4A8)- К О Д ВЫБОРКИ; (I4)-ДЛИНА
 XX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXX
 KSNQ-01 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 61 0100
 KSNQ-02 ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 62 0100
 (I2, I3, I2) [4A8] :НОМЕР ВЫБ., КОД, КАНАЛ, ИМЯ ФАЙЛА
 FILT-01 000 08-INF:DATA60.DAT
 [СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4)]
 FORT-01 (4X, 12F4.0)
 5 0 0 0 0 6 1 2 6 6 34 14
 3 3 14 6
 Q - 0.05 <- (G14.7)- УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
 NFWR-02 (I2) - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
 (I1, I2)-ТИП КЛЮЧЕЙ ПЕЧАТИ: 0-ПОРАЗРЯДНЫЕ; И ИХ КОЛИЧЕСТВО
 KEYS-02 02 1-ИНТЕРВАЛЬНЫЕ.
 ИНДЕКСЫ/ИНТЕРВАЛЬНОЕ ЗНАЧЕНИЕ КЛЮЧА ПЕЧАТИ/ПОРАЗРЯДНОЕ/
 12345678 - РАЗРЯДЫ КЛЮЧА ПЕЧАТИ ;
 KEY - 1 2 4/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SP420 ;
 KEY - 2 2 4/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP420 .
 (I1)-РАСПРЕДЕЛЕНИЕ ПАМЯТИ:=1-ВЫДАЕТСЯ ВСЕГДА;
 END -1 КОНЕЦ ДАННЫХ: =0-ВЫДАЕТСЯ ТОЛЬКО АВАРИЙНО.)

РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ ЗАДАНИЯ (РАСПЕЧАТКА РЕШЕНИЯ) :

IRP -106

K - 4

M - 4

KSQN NN XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX NNNN

KSQN- 1 ТЕСТОВАЯ /4 ГРАДАЦИИ/ ВЫБОРКА 61 100

KSQN- 2 ТЕСТОВАЯ /4 ГРАДАЦИИ/ ВЫБОРКА 62 100

FILT- 1 0 8-INF:DATA60.DAT

FORT- 1: (4X,12F4.0)

5. 0. 0. 0. 0. 6. 1. 2. 6. 6. 34. 14.

3. 3. 14. 6.

Q 0.5000000E-01

NFWR 2

KEYC 0 2 2

KEY 1 2 4/8/11111111/

KEY 2 2 4/8/11111111/

ОБЪЕМ РАБОЧЕЙ ПАМЯТИ - 4000 ЕДИНИЦ INTEGER

ТРЕБУЕМЫЙ ОБЪЕМ ПАМЯТИ - 177 ЕДИНИЦ INTEGER

ИНДЕКСЫ НАЧАЛА ПЕРЕМЕННЫХ В РАБОЧЕЙ ПАМЯТИ :

| | | | | | | | |
|-----|---|------|-----|------|-----|------|-----|
| 1 - | 1 | 6 - | 6 | 11 - | 107 | 16 - | 72 |
| 2 - | 2 | 7 - | 7 | 12 - | 108 | 17 - | 73 |
| 3 - | 3 | 8 - | 36 | 13 - | 109 | 18 - | 168 |
| 4 - | 4 | 9 - | 19 | 14 - | 55 | 19 - | 169 |
| 5 - | 5 | 10 - | 105 | 15 - | 71 | 20 - | 170 |

SW420 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА = 0.050

ТАБЛИЦА СФОРМИРОВАНА ИЗ ВЫБОРОК:

ТЕСТОВАЯ /4 ГРАДАЦИИ/ ВЫБОРКА 61

ТЕСТОВАЯ /4 ГРАДАЦИИ/ ВЫБОРКА 62

СТРОК (ГРАДАЦИИ КАЧЕСТВЕННОГО ПРИЗНАКА "А") = 4

СТОЛБЦОВ (ГРАДАЦИИ КАЧЕСТВЕННОГО ПРИЗНАКА "Б") = 4

ТАБЛИЦА МУЛЬТИНОМИНАЛЬНЫХ ЧАСТОТ ПРИЗНАКОВ " А"И "Б":

| | | | | |
|---|----|----|-----|-----|
| 1 | 5. | 0. | 6. | 3. |
| 2 | 0. | 6. | 6. | 3. |
| 3 | 0. | 1. | 34. | 14. |
| 4 | 0. | 2. | 14. | 6. |

* ВЫЧИСЛЕНИЕ КОЭФФ. КОРР. СВЯЗИ

* КОЭФФ. ЧУПРОВА/КРАМЕРА * МЕЖДУ ДВУМЯ СОВОКУПНОСТЯМИ

* В ТАБЛИЦЕ СОПРЯЖЕННОСТИ R*С

PR420 - СПРАВЕДЛИВА ГИПОТЕЗА H1 КОД ОШИБКИ = 0

КОЭФФИЦИЕНТ КОРРЕЛЯЦИОННОЙ СВЯЗИ

126.74

СТАНДАРТНАЯ ОШИБКА

0.333333E-02

КРИТ. ЗНАЧЕНИЕ РАСПРЕД. ХИ-КВАД.

16.919

ТАБЛИЦА СФОРМИРОВАНА ИЗ ВЫБОРОК:

ТЕСТОВАЯ /4 ГРАДАЦИИ/ ВЫБОРКА 61

ТЕСТОВАЯ /4 ГРАДАЦИИ/ ВЫБОРКА 62

```

CS; SP420 - КОЭФФИЦИЕНТЫ КОРРЕЛЯЦИИ ЧУПРОВА И КРАМЕРА
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫЧИСЛЕНИЯ СТЕПЕНИ КОРРЕЛЯЦИОННОЙ
C СВЯЗИ МЕЖДУ 2-Я СОВОКУПНОСТЯМИ ПРИЗНАКОВ, ИМЕЮЩИХ НЕ
C МЕНЕЕ ЧЕМ ПО 3 ГРАДАЦИИ, В ТАБЛИЦЕ СОПРЯЖЕННОСТИ "R*С"
C; (КОЭФФИЦИЕНТЫ ЧУПРОВА И КРАМЕРА)
C
CR ЛИТЕРАТУРА ПО АЛГОРИТМУ : М.ДЖ.КЕНДАЛ, А.СТЮАРТ
C; СТАТИСТИЧЕСКИЕ ВЫВОДЫ И СВЯЗИ, ТОМ 2, М., НАУКА, 1973.)
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ,
C;
C
SUBROUTINE SP420(KEY,NFWR,KSQN,L,K,M,T,Q,X,Y,C,S,THI
* ,IND,IER,IERW)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CP ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C
C KEY I*4 * (8) МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 * 1 НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
C KSQN R*8 * (4,L) МАССИВ К О Д О В ВЫБОРОК
C L I*4 * 1 ДЛИНА "KSQN" /L=1 ИЛИ L=2/
C K I*4 * 1 КОЛИЧЕСТВО ГРАДАЦИЙ ПРИЗНАКА "А" /K>2/
C M I*4 * 1 КОЛИЧЕСТВО ГРАДАЦИЙ ПРИЗНАКА "В" /M>2/
C T R*4 * (K,M) ТАБЛИЦА МУЛЬТИНОМИНАЛ. ЧАСТОТ "А" И "В"
C Q R*4 * 1 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА/0.005<=Q<=0.5/
C X R*4 * (M) РАБОЧИЙ МАССИВ: СУММЫ ПО СТОЛБЦАМ
C Y R*4 * (K) РАБОЧИЙ МАССИВ: СУММЫ ПО СТРОКАМ
C C R*4 * 1 ВЫЧИСЛ. КОЭФФИЦИЕНТ КОРРЕЛЯЦ. СВЯЗИ
C S R*4 * 1 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ СТАНДАРТНОЙ ОШИБКИ
C THI R*4 * 1 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ РАСПРЕД. ХИ-КВАДР.
C IND I*4 * 1 ИНДЕКС ПРОВЕРКИ ПО КРИТЕРИЮ ХИ-КВАДРАТ:
C IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
C IND = 1 - НЕЛЕВАЯ ГИПОТЕЗА ОТКЛОНЯЕТСЯ.
C IER I*4 * 1 ВЫРАБОТАННЫЙ КОД ОШИБКИ
C; IERW I*4 * (8) МАССИВ КОДОВ ОШИБОК ПРИНЯТЫЙ ИЗ CW420
C
DIMENSION KEY(8), IERW(8), T(K,M), KSQN(4,L), X(M), Y(K)
REAL*8 NAME, KSQN
DATA NAME /' SP420 '/, ZERO, HALF, ONE/0.0, 0.5, 1.0/
* EPS/0.001/
C
IND = 11
IER = 0
KLP = KEY(2)
IF (KEY(1).EQ. 0) GO TO 120
CALL CW420(KEY,NFWR,KSQN,L,K,M,T,Q,IERW,MWC)
CL; CW420 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ;
C АНАЛИЗ КОДОВ ОШИБОК "KF" И "MWC"
KF = IERW(1)
IF (MWC.LT.3) GO TO 120
C ЕСЛИ MWC >= 2, TO IER=MWC
IER = MWC
IF (MWC.GE.3) GO TO 999
C ПОДГОТОВКА
120 Z = ZERO
DO 220 I=1,K
H = ZERO
DO 210 J=1,M
210 H = H + T(I,J)
Z = Z + H

```

```

220 Y(I) = H
DO 240 J=1,M
H = ZERO
DO 230 I=1,K
230 H = H + T(I,J)
240 X(J) = H
C ВЫЧИСЛЕНИЕ X2
R=ZERO
DO 310 J=1,M
DO 310 I=1,K
S = T(I,J)
R = R + S/Y(I)*(S/X(J))
310 CONTINUE
R=(R-ONE)*Z
IF (K.EQ.M) GO TO 410
C ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТА КРАМЕРА
S=ONE/(Z*AMIN0(K-1,M-1))
GO TO 420
C ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТА ЧУПРОВА
410 S=ONE/( Z * SQRT(FLOAT(K-1)*FLOAT(M-1)) )
420 C=SQRT(R/S)
C ВЫЧИСЛЕНИЕ THI - КРИТИЧЕСКОГО ЗНАЧЕНИЯ HI2
NU=(K-1)*(M-1)
CALL MP030(Q,NU,EPS,THI,IERTHI)
CL; MP030 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ РАСПРЕД. ХИ-КВАДРАТ
IF (IERTHI .EQ. 0) GOTO 510
IER = 4+IERTHI
GOTO 810
510 IND = 0
IF (C .GE. HALF) IND=1
C
810 IF (KF .NE. 3) GO TO 830
IER = -IER
GO TO 999
830 IF (KLP .EQ. 0) GO TO 999
KEY(1) = KLP
CALL PR420(KEY,NFWR,KSQN,L,C,S,THI,IND,IER)
CL; PR420 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP420.
999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0 02/02/88
END

```

```

СC; SW420 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP420
С
СF П/П ПРЕДНАЗНАЧЕНА ДЛЯ КОНТРОЛЯ ВХОДНЫХ ДАННЫХ ДЛЯ SP420
С И ВЫДАЧИ В ФАЙЛ С НОМЕРОМ "NFWR" РЕЗУЛЬТАТА ПРОВЕРКИ
С; И ВХОДНЫХ ДАННЫХ
С
СR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP420
С
СТ; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
С
СА П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ
С; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
С
SUBROUTINE SW420(KEY,NFWR,KSQN,L,K,M,T,Q,IERW,MWC)
С
С ОПИСАНИЕ ПАРАМЕТРОВ :
С -----
С
СР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
С ПЕЧАТИ
С
С KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
С NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
С KSQN R*8 (4,L) 4A8 7 МАССИВ К О Д О В ВЫБОРОК
С L I*4 1 -- - ДЛИНА "KSQN" /L=1 ИЛИ L=2/
С K I*4 1 I4 8 КОЛ-ВО ГРАДАЦИИ ПРИЗНАКА "А" /K>2/
С M I*4 1 I4 8 КОЛ-ВО ГРАДАЦИИ ПРИЗНАКА "Б" /M>2/
С T R*4 (K,M) 6F8.0 8 ТАБЛИЦА МУЛЬТИНОМИН. ЧАСТОТ "А" И "Б"
С Q R*4 1 F8.3 3 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА/.005<=Q<=.5/
С IERW I*4 (8) 6I2 [2]ВЫРАБОТАННЫЙ КОД ОШИБКИ : 6 РАЗРЯД00
С;MWC I*4 1 I2 2 КОД МАКСИМАЛЬНОЙ ОШИБКИ
С
DIMENSION KEY(8),IERW(8),T(K,M),KSQN(4,L)
С
REAL*8 NAME,KSQN
DATA NAME /' SW420 '/, EPS /0.001/, ZERO/0.0/
С
KLU=KEY(1)
IF(KLU) 11,999,15
С - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
13 KLU=KLU/2
GO TO 19
С - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF(KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
С
С КОНТРОЛЬ ВХОДНЫХ ДАННЫХ
С
С -*- ПРОВЕРКА 'NFWR' ( 1-Я РАЗРЯД )
С НАЧАЛЬНАЯ УСТАНОВКА : ОШИБОК НЕТ
KF = 0
С ЕСЛИ 'NFWR' БОЛЬШЕ 99 ИЛИ МЕНЬШЕ 3, ЭТО ГРУБАЯ ОШИБКА
IF (NFWR.GT.99) KF = 3
IF (NFWR.LT.1) KF = 3
С ФОРМИРОВАНИЕ КОДА ОШИБКИ
IERW(1) = KF
С
С -*- ПРОВЕРКА 'K' ( 3-Я РАЗРЯД )
С НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'K' : 2 <= K <= 99

```



```

      KK = 0
C   ЕСЛИ K > 100, ЭТО УСЛОВНО СЧИТАЕТСЯ ЛОГИЧЕСКОЙ ОШИБКОЙ
      IF (K.GE.100) KK = 1
C   НЕПРАВИЛЬНОЕ ЗНАЧЕНИЕ 'K' : K < 2
      IF (K.LT.2) KK = 3
C   ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(3) = KK
C
C-1-   ПРОВЕРКА 'M' ( 4-й РАЗРЯД )
C   НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'M' : 2 <= M <= 99
      KM = 0
C   ЕСЛИ M >= 100, ЭТО СЧИТАЕТСЯ ЛОГИЧЕСКОЙ ОШИБКОЙ
      IF (M.GE.100) KM = 1
C   НЕПРАВИЛЬНОЕ ЗНАЧЕНИЕ 'M' : M < 2
      IF (M.LT.2) KM = 3
C   ПРОВЕРКА НА СОВМЕСТИМОСТЬ :
C   ОДНОВРЕМЕННО НЕ ДОЛЖНО БЫТЬ : K=2 И M=2 .
      IF (K.EQ.2 .AND. M.EQ.2) KM = 3
C   ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(4) = KM
C
C-1-   ПРОВЕРКА 'L' ( 2-й РАЗРЯД )
C   KL=0 ОЗНАЧАЕТ, ЧТО 'L'- НОРМАЛЬНОЕ,
C   Т.Е. L=1 ИЛИ L=2
      KL = 3
      IF (L.EQ.1) KL = 0
      IF (L.EQ.2) KL = 0
C   ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(2) = KL
C
C-1-   ПРОВЕРКА 'T' ( 5-й РАЗРЯД )
C   KT=0 ОЗНАЧАЕТ, ЧТО 'T'- НОРМАЛЬНАЯ ТАБЛИЦА,
C   Т.Е. СУММА ЭЛЕМЕНТОВ В ЛЮБОЙ СТРОКЕ НЕ= 0 И
C   СУММА ЭЛЕМЕНТОВ В ЛЮБОМ СТОЛБЦЕ НЕ= 0.
      KT=11
C   ПРОВЕРКА ВОЗМОЖНОСТИ РАБОТЫ С МАССИВОМ 'T' :
      IF (KM.GT.1 .OR. KK.GT.1) GO TO 70
      KT=0
C   ПРЕДЕЛЫ НОРМАЛЬНЫЕ .
      DO 63 I=1,K
      S = ZERO
      DO 61 J=1,M
61      S = S+T(I,J)
          IF (S .EQ. ZERO) GO TO 69
63      CONTINUE
      DO 67 J=1,M
      S = ZERO
      DO 65 I=1,K
65      S = S+T(I,J)
          IF (S .EQ. ZERO) GO TO 69
67      CONTINUE
      GO TO 70
69      KT = 3
C   ФОРМИРОВАНИЕ КОДА ОШИБКИ
      70 IERW(5) = KT
C
C-1-   ПРОВЕРКА 'Q' ( 6-й РАЗРЯД )
C   НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'Q' : 0.025 <= Q <= 0.2
      KQ = 0
C   ЕСЛИ 'Q' БОЛЬШЕ ЧЕМ 0.2 ИЛИ МЕНЬШЕ ЧЕМ 0.025,
C   ТО СЧИТАЕМ ЧТО ДОПУЩЕНА ЛОГИЧЕСКАЯ ОШИБКА
      IF (Q.LT.0.025) KQ = 1
      IF (Q.GT.0.2) KQ = 1
C   ВВОДИМ ОГРАНИЧЕНИЕ : 0.005 <= Q <= 0.5
      IF (Q.LT.0.005) KQ = 3
      IF (Q.GT.0.5) KQ = 3

```

C ФОРМИРОВАНИЕ КОДА ОШИБКИ

IERW(6) = KQ

IERW(8) = 6

C ОПРЕДЕЛЕНИЕ МАКСИМАЛЬНОГО КОДА ОШИБКИ

MWC = MAX0 (KL,KK,KM,KT,KQ)

C

C ШКАЛА КОДИРОВКИ ОШИБОЧНЫХ ПАРАМЕТРОВ :

C SE РАЗРЯД КОДА ОШИБКИ I 1 I 2 I 3 I 4 I 5 I 6

C ПАРАМЕТР В ОБРАЩЕНИИ I NFWR I L I K I M I T I Q

C ВОЗМОЖНЫЕ ЗНАЧЕНИЯ I 0 I 0 I 0 I 0 I 0 I 0 I 0

C I - I - I 1 I 1 I - I 1

C I - I - I - I - I - I -

C I 3 I 3 I 3 I 3 I 3 I 3 I 3

C I - I - I - I - I 11 I -

C

C ПЕЧАТЬ ВХОДНЫХ ДАННЫХ

IF (KF.GT.0) GO TO 999

IF (KEY(2).EQ.0) GO TO 225

C

WRITE (NFWR,220) NAME,MWC

220 FORMAT(/' ',A8,4X,'КОД МАКСИМАЛЬНОЙ ОШИБКИ =' ,I2)

C

225 IF (MWC.LT.1) GO TO 310

J=IERW(8)

WRITE (NFWR,227) (IERW(I),I=1,J)

227 FORMAT(/' ',8X,'ИМЕНА ПАРАМЕТРОВ :'

* , ' NFWR, L, K, M, T, Q'

* /' ',8X,'К О Д Ы ОШИБОК :',6(3X,I2)/)

C

310 IF (KEY(3).EQ.0) GOTO 710

WRITE (NFWR,330) Q

330 FORMAT(/' ',8X,'УРОВЕНЬ ЗНАЧИМОСТИ РИСКА =' ,F8.3)

C

710 IF((KEY(7).EQ.0).OR.(L.EQ.1)) GOTO 810

WRITE (NFWR,750) ((KSON(I,J),I=1,4),J=1,L)

750 FORMAT(/' ',8X,'ТАБЛИЦА СФОРМИРОВАНА ИЗ ВЫБОРОК:'

* /(' ',8X,5X,4A8))

C

810 IF (KEY(8).EQ.0) GOTO 999

WRITE (NFWR,860) K,M

860 FORMAT(' ',8X

* , 'СТРОК (ГРАДАЦИЯ КАЧЕСТВЕННОГО ПРИЗНАКА "А") ='

* , I4 /' ',8X

* , 'СТОЛБЦОВ (ГРАДАЦИЯ КАЧЕСТВЕННОГО ПРИЗНАКА "Б") ='

* , I4)

C

WRITE (NFWR,870)

870 FORMAT(/' ',8X

* , 'ТАБЛИЦА МУЛЬТИНОМИНАЛЬНЫХ ЧАСТОТ ПРИЗНАКОВ " А"И "Б":')

C

IF (KK.GT.1.OR.KM.GT.1) GO TO 999

DO 880 I=1,K

WRITE(NFWR,890) I,(T(I,J),J=1,M)

880 CONTINUE

890 FORMAT(' ',8X,I2,1X,6F8.0/(' ',11X,6F8.0))

C

999 RETURN

CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0

02-DEC-87

END

```

CC; PR420 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP420
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫВОДА В ФАЙЛ 'NFWR'
C; РЕЗУЛЬТАТОВ РАБОТЫ П/П SP420
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP420
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE PR420(KEY,NFWR,KSQN,L,C,S,THI,IND,IER)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CP ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
C KSNQ R*8(4,L) 4A8 7 МАССИВ К О Д О В ВЫБОРОК, ОБРАЗ. Т
C L I*4 1 -- - ДЛИНА "KSNQ" /L=1 ИЛИ L=2/
C C R*4 1 G20.5 3 ВЫЧИСЛЕННЫЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ
C S R*4 1 G20.5 3 ВЫЧИСЛ. ЗНАЧЕНИЕ СТАНДАРТНОЙ ОШИБКИ
C THI R*4 1 G20.5 3 КРИТ. ЗНАЧЕНИЕ РАСПРЕДЕЛ. ХИ-КВАДРАТ
C IND I*4 1 I1 2 ИНДЕКС ПРОВЕРКИ ПО КРИТ. ХИ-КВАДРАТ:
C IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
C IND = 1 - НУЛЕВАЯ ГИПОТЕЗА ОТВЕРГАЕТСЯ.
C; IER I*4 1 I3 2 КОД ОШИБКИ, ВЫРАБОТАННЫЙ SP420
C
DIMENSION KEY(8),KSNQ(4,L)
C
REAL*8 TNK(3),TS1(4),TS2(4),TS3(4),NAME,KSNQ
DATA TNK /' КОЭФФ.', 'ЧУПРОВА', 'КРАМЕРА' /
* ,TS1 /' ВЫЧИСЛЕ.', 'НИЕ КОЭ.', 'ФФ. КОРР.', ' СВЯЗИ' /
* ,TS2 /' МЕЖДУ.', ' ДВУМЯ', ' СОВОКУПН.', ' ОСТЯМИ' /
* ,TS3 /' В ТАБЛ.', ' ИЦЕ СОП', ' РЯЖЕННОС', 'ТИ R*С' /
DATA NAME /' PR420' /
C
IF (NFWR.LE.0.OR.NFWR.GE.100) GO TO 999
C
KLU=KEY(1)
IF (KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
13 KLU=KLU/2
GO TO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF (KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
IF (KEY(1).EQ. 0) GO TO 210
C
WRITE (NFWR,120) TS1,TNK,TS2,TS3
120 FORMAT(/ 1H ,26(1H*)
* /2H *,24(1H ),3H* ,4A8
* /2H *,3A8 ,3H* ,4A8
* /2H *,24(1H ),3H* ,4A8
* / 1H ,26(1H*) )

```


6.9. Пример. Оценки статистической связи между двумя выборками, одна из которых дихотомия, а другая измерена в шкале отношений, приведены в [000]. Необходимо проверить предложение о том, что существует связь между полами и ростом мужчин и женщин. Как всегда, нулевая гипотеза формулируется как отсутствие значимой корреляционной связи между двумя этими признаками.

Задание на обработку («входной поток») должно содержать следующие данные: *IRP-027* — код выполняемой функции; коды и длины выборок (длины должны быть равны); местонахождение и формат записи дихотомической выборки; элементы дихотомической выборки в виде чисел 0 и 1 (пол); местонахождение и формат записи выборки в шкале отношений (рост); уровень значимости риска (обычно 5 % или 10 %); ключи управления печатью.

Далее приведен полный текст файла «входного потока» и распечатка результата работы программ.

```

ЗАДАНИЕ НА ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТА КОРРЕЛЯЦИОННОЙ СВЯЗИ
ДВУХ ВЫБОРОЧНЫХ СОВОКУПНОСТЕЙ, ОДНА ИЗ КОТОРЫХ
РЕЗУЛЬТАТ ИЗМЕРЕНИЙ В ШКАЛЕ ОТНОШЕНИЙ,
А ДРУГАЯ — В НОМИНАЛЬНОЙ ШКАЛЕ С ДВУМЯ ГРАДАЦИЯМИ

-----  ВЫЧИСЛЕНИЕ КОРРЕЛЯЦИОННОЙ СВЯЗИ ДВУХ ВЫБОРОК :
****  МАКЕТ ПОДГОТОВКИ ДАННЫХ ДЛЯ IR= 27:      ****SP482****
IRP -027      (I3)- КОД ФУНКЦИИ
NS -02 /ПО УМОЛЧАНИЮ/ (I2)- КОЛИЧЕСТВО ВЫБОРОК
K -01 /ПО УМОЛЧАНИЮ/ (I2)- КОЛИЧЕСТВО ГРАДАЦИЙ ВЫБОРКИ 1
M -01 /ПО УМОЛЧАНИЮ/ (I2)-
(I2)-НОМЕР; (4A8)- К О Д ВЫБОРКИ; (I4)-ДЛИНА
XX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXX
KSGN-01 ТЕСТОВАЯ /2 ГРАДАЦИЙ/ ВЫБОРКА 7A 0015
KSGN-02 ТЕСТОВАЯ /ШК.ОТНОШЕНИЯ/ВЫБОРКА 7B 0015
(I2, I3, I2)[ 4A8 ]:НОМЕР ВЫБ.,КОД,КАНАЛ,ИМЯ ФАЙЛА
FILX-01 000 08-INF:DATA71.DAT
[ СПИСОК СПЕЦИФИКАЦИЙ ОПЕРАТОРА F O R M A T (14A4) ]
FORX-01 (4X,30I2)
1 0 1 1 0 1 0 0 1 1 1 0 1 0 0
FILZ-01 000 08-INF:DATA71.DAT
FORZ-01 (4X,6F10.2)
150.00 170.00 160.00 165.00 140.00 183.00
157.00 152.00 163.00 168.00 160.00 155.00
157.00 160.00 152.00
Q - 0.05 <- (G14.7)- УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
NFWR-02 (I2) - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
(I1, I2)-ТИП КЛЮЧЕЙ ПЕЧАТИ: 0-ПОРАЗРЯДНЫЕ; И ИХ КОЛИЧЕСТВО
KEYC-0 02 1-ИНТЕРВАЛЬНЫЕ.
ИНДЕКСЫ/ИНТЕРВАЛЬНОЕ ЗНАЧЕНИЕ КЛЮЧА ПЕЧАТИ/ПОРАЗРЯДНОЕ/
12345678 - РАЗРЯДЫ КЛЮЧА ПЕЧАТИ :
KEY - 1 8 4/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SP482 ;
KEY - 2 8 4/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SP482 .
(I1)-РАСПРЕДЕЛЕНИЕ ПАМЯТИ:=1-ВЫДАЕТСЯ ВСЕГДА;
END -1 КОНЕЦ ДАННЫХ; =0-ВЫДАЕТСЯ ТОЛЬКО АВАРИЙНО.

```

РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ ЗАДАНИЯ (РАСПЕЧАТКА РЕШЕНИЯ) :

```

IRP - 27
KSNQ NN XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX NNNN
KSNQ- 1 ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА 7А 15
KSNQ- 2 ТЕСТОВАЯ/ШК.ОТНОШЕНИЙ/ВЫБОРКА 7В 15
FILX- 1 0 8-INF:DATA71.DAT
FORX- 1:(4X,30I2)
      1 0 1 1 0 0 1 1 1 0 1 0 0
FILZ- 1 0 8-INF:DATA71.DAT
FORZ- 1:(4X,6F10.2)
      150.00 170.00 160.00 165.00 140.00 183.00
      157.00 152.00 163.00 168.00 160.00 155.00
      157.00 160.00 152.00
Q      0.5000000E-01
NFWR 2
KEYC 0 2 4
KEY 1 8 4/8/11111111/
KEY 2 8 4/8/11111111/
ОБЪЕМ РАБОЧЕЙ ПАМЯТИ - 4000 ЕДИНИЦ INTEGER
ТРЕБУЕМЫЙ ОБЪЕМ ПАМЯТИ - 185 ЕДИНИЦ INTEGER

ИНДЕКСЫ НАЧАЛА ПЕРЕМЕННЫХ В РАБОЧЕЙ ПАМЯТИ :
1 - 1 6 - 6 11 - 107 16 - 78
2 - 2 7 - 7 12 - 122 17 - 79
3 - 3 8 - 36 13 - 123 18 - 166
4 - 4 9 - 19 14 - 62 19 - 168
5 - 5 10 - 105 15 - 63 20 - 170

CW482 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

УРОВЕНЬ ЗНАЧИМОСТИ РИСКА = 0.050
К О Д Ы ВЫБОРОК :

1 - ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 7А
2 - ТЕСТОВАЯ/ШК.ОТНОШЕНИЙ/ВЫБОРКА 7В

ДЛИНА ВЫБОРОК = 15
МАССИВ ЗНАЧЕНИЙ В ШКАЛЕ ОТНОШЕНИЙ :
150.00 170.00 160.00 165.00
140.00 183.00 157.00 152.00
163.00 168.00 160.00 155.00
157.00 160.00 152.00
МАССИВ ЗНАЧЕНИЙ В ШКАЛЕ НАИМЕНОВАНИЙ :
1 0 1 1
0 1 0 0
1 1 1 1
1 0 0 0

*****
* * ВЫЧИСЛ. ТОЧЕЧНОГО БИСЕРИАЛЬНОГО
*Т0Ч. БИСЕР. КОЭФФ. КОРР.* КОЭФФ. КОРРЕЛЯЦИИ 2 ВЫБ. ОДНА -В
* * ШКАЛЕ НАИМЕНОВАНИЙ,2-Я ОТНОШЕНИЙ
*****

PR482 - СПРАВЕДЛИВА ГИПОТЕЗА Н0 ; КОД ОШИБКИ = 0

КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ = 0.43432
КРИТИЧЕСКОЕ ЗНАЧЕНИЕ Т-РАСПРЕДЕЛЕНИЯ = 2.1450
К О Д Ы ВЫБОРОК :
1 - ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 7А
2 - ТЕСТОВАЯ/ШК.ОТНОШЕНИЙ/ВЫБОРКА 7В

```

```

С; SP482 - ТОЧЕЧНЫЙ БИСЕРИАЛЬНЫЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ПИРСОНА
С
СF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫЧИСЛЕНИЯ ТОЧЕЧНОГО БИСЕРИАЛЬНОГО
С КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ ПИРСОНА МЕЖДУ ДВУМЯ ВЫБОРКАМИ,
С ОДНА ИЗ КОТОРЫХ ИЗМЕРЕНА В ШКАЛЕ ОТНОШЕНИЯ, ДРУГАЯ -
С; В ШКАЛЕ НАИМЕНОВАНИЙ В ПРЕДПОЛОЖЕНИИ ДИХОТОМИИ
С
СR ЛИТЕРАТУРА ПО АЛГОРИТМУ : ДЖ. ГЛАСС, ДЖ. СТЭНЛИ,
С СТАТИСТИЧЕСКИЕ ВЫВОДЫ В ПЕДАГОГИКЕ И ПСИХОЛОГИИ, М.,
С; ПРОГРЕСС, 1976.
С
СТ; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
С
СА П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММО,
С; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
С
SUBROUTINE SP482(KEY,NFWR,KSQN,N,X,M,Q,C,IND,IER,IERW)
С
С ОПИСАНИЕ ПАРАМЕТРОВ :
С
С -----
СР ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
С
С KEY I*4 * (8) МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
С NFWR I*4 * 1 НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
С KSQN R*8 * (4,2) МАССИВ К О Д О В ВЫБОРОК
С N I*4 * 1 ДЛИНА ВЫБОРОК "X" И "M"
С X R*4 * (N) МАССИВ ЗНАЧЕНИЙ, ИЗМЕР. В ШК. ОТНОШЕНИЯ
С M I*4 * (N) МАССИВ ЗНАЧЕНИЙ, ИЗМЕР. В ШК. НАИМЕНОВАНИЙ
С Q R*4 * 1 УРОВНЬ ЗНАЧИМОСТИ РИСКА /0.05, 0.1/
С C R*4 * 1 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ КОЭФФ. КОРРЕЛЯЦИИ
С IND I*4 * 1 ИНДЕКС ПРОВЕРКИ ПО ИНТЕРВ. ОЦЕНИВАНИЮ :
С IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
С IND = 1 - НУЛЕВАЯ ГИПОТЕЗА ОТКЛОНЯЕТСЯ.
С IER I*4 * 1 ВЫРАБОТАННЫЙ КОД ОШИБКИ
С; IERW I*4 * (8) МАССИВ КОДОВ ОШИБОК, ПРИНЯТЫЙ ИЗ SW482
С
DIMENSION KEY(8),IERW(8),X(N),M(N),KSQN(4,2)
REAL*8 NAME, KSQN
DATA NAME /' SP482 '/
* ,ZERO/0.0/,CONST/0.05/,ONE/1.0/
IND = 11
IER = 0
KLP = KEY(2)
IF (KEY(1) .EQ. 0) GO TO 120
CALL SW482(KEY,NFWR,KSQN,N,X,M,Q,IERW,MWC)
CL; SW482 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ SP482;
С
С АНАЛИЗ КОДОВ ОШИБОК : "KF" И "MWC"
С
KF = IERW(1)
IF (MWC.LT.3) GO TO 120
С ЕСЛИ MWC >= 3, TO IER=MWC
IER = MWC
IF (MWC.GE.3) GO TO 999
С ВЪЧИСЛЕНИЕ СРЕДНЕГО ЗНАЧЕНИЯ X, СООТВ. '1' ИЛИ '0' ПО Y
120 N0=0
N1=0
X0=ZERO
X1=ZERO
DO 220 I=1,N
IF (M(I).EQ.1) GO TO 210
N0=N0+1
X0=X0+X(I)
GO TO 220
210 CONTINUE

```

```

      N1=N1+1
      X1=X1+X(I)
220  CONTINUE
      C=(X0+X1)/FLOAT(N)
      X0=X0/N0
      X1=X1/N1
C  ВЫЧИСЛЕНИЕ СРЕДНЕКВАДРАТИЧЕСКОГО ОТКЛОНЕНИЯ ДЛЯ X
      S=ZERO
      DO 310 I=1,N
      S=S+(X(I)-C)**2
,310  CONTINUE
      C=FLOAT(N)
      S=SQRT(S/C)
C  ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ
      C=(X1-X0)/S*SQRT((N1/C)*(N0/(C-1)))
      IF (ABS(C).LT.CONST) GO TO 450
      CALL MP010(Q,N-1,0.001,TQV,IER)
CL; MP010 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ T-РАСПРЕДЕЛЕНИЯ
      IF(IERT.EQ. 0) GOTO 430
      IER = 4+IER
      GOTO 810
430  IND = 0
      IF (ABS(C).GT.TQV) IND=1
      GO TO 510
450  CONTINUE
      T=C/SQRT((ONE-C*C)/FLOAT(N-2))
      CALL MP010(Q,N-2,0.001,TQV,IER)
CL; MP010 - ВЫЧИСЛЕНИЕ КРИТИЧЕСКИХ ЗНАЧЕНИЙ T-РАСПРЕДЕЛЕНИЯ
      IF(IERT.EQ. 0) GOTO 480
      IER = 4+IER
      GOTO 810
480  IND = 0
      IF (ABS(T).GT.TQV) IND=1
      510 CONTINUE
C
810  IF (KF.NE. 3) GO TO 830
      IER = -IER
      GO TO 999
830  IF (KLP.EQ. 0) GO TO 999
      KEY(1) = KLP
      CALL PR482(KEY,NFWR,KSQN,C,TQV,IND,IER)
CL; PR482 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP482.
999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0      2-FEB-88
      END

```



```

CC; CW482 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP482
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ КОНТРОЛЯ ВХОДНЫХ ДАННЫХ ДЛЯ SP482
C И ВЫДАЧИ В ФАЙЛ С НОМЕРОМ "NFWR" РЕЗУЛЬТАТА ПРОВЕРКИ
C И ВХОДНЫХ ДАННЫХ
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP482
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE CW482(KEY,NFWR,KSQN,N,X,M,Q,IERW,MWC)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 --- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
C KSQN R*8 (4,2) 4A8 7 МАССИВ КОДОВ ВЫБОРОК
C N I*4 1 I5 8 ДЛИНА ВЫБОРОК "X" И "M"
C X R*4 1 I4 8 МАССИВ ЗНАЧЕНИЙ, ИЗМЕР. В ШК.ОТНОШЕНИИ
C M I*4 1 I4 8 МАССИВ ЗНАЧ., ИЗМЕР. В ШК.НАИМЕНОВАНИИ
C Q R*4 1 F8.3 3 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
C IERW I*4 (8) 5I2 2 ВЫРАБОТАННЫЙ КОД ОШИБКИ : 5 РАЗРЯДОВ
C;MWC I*4 1 I2 2 КОД МАКСИМАЛЬНОЙ ОШИБКИ
C
DIMENSION KEY(8),IERW(8),X(N),M(N),KSQN(4,2)
REAL*8 NAME, KSQN
DATA NAME /' CW482 '/
C
KLU=KEY(1)
IF(KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
13 KLU=KLU/2
GOTO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF(KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
C
C КОНТРОЛЬ ВХОДНЫХ ДАННЫХ
C
C-*-- ПРОВЕРКА 'NFWR' ( 1-Я РАЗРЯД )
C НАЧАЛЬНАЯ УСТАНОВКА : ОШИБОК НЕТ
KF = 0
C ЕСЛИ 'NFWR' БОЛЬШЕ 99 ИЛИ МЕНЬШЕ 1, ЭТО ГРУБАЯ ОШИБКА
IF (NFWR.GT.99) KF = 3
IF (NFWR.LT.1) KF = 3
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
IERW(1) = KF
C
C-*-- ПРОВЕРКА 'N' ( 2-Я РАЗРЯД )
C KN=0 ОЗНАЧАЕТ, ЧТО 'N' - НОРМАЛЬНОЕ, Т.Е. БОЛЬШЕ 9
KN = 0
C ЕСЛИ 'N' МЕНЬШЕ ЧЕМ 10, ТО ЭТО ЛОГИЧЕСКАЯ ОШИБКА

```

```

      IF (N.LT.10) KN = 1
C   ЕСЛИ 'N' МЕНЬШЕ ЧЕМ 2, ЭТО ГРУБАЯ ОШИБКА
      IF (N.LT. 2) KN = 3
C   ПО ВЕРХНЕМУ ПРЕДЕЛУ ДЛЯ 'N' ПРОВЕРКИ НЕТ
C   ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(2) = KN
C-#-      ПРОВЕРКА 'X' ( 3-Я РАЗРЯД )
C   ЕСЛИ 'N' - НЕПРАВИЛЬНОЕ, ТО ПРОВЕРКА НЕВОЗМОЖНА
      KX = 11
      IF (KN.GT.1) GO TO 30
C   ПОЛОЖИМ, ЧТО МАССИВ 'X' = ПРАВИЛЬНЫЙ
      KX = 0
C   ПОИСК 'MIN' И 'MAX' ЭЛЕМЕНТОВ В 'X'
      XMIN = X(1)
      XMAX = X(1)
      DO 25 I=2,N
      R = X(I)
      IF (R.GT.XMAX) XMAX=R
      IF (R.LT.XMIN) XMIN=R
25  CONTINUE
      IF (XMAX.EQ.XMIN) KX=3
C   ФОРМИРОВАНИЕ КОДА ОШИБКИ
      30 IERW(3) = KX
C
C-#-      ПРОВЕРКА 'M' ( 4-Я РАЗРЯД )
C   KM = 0 ОЗНАЧАЕТ ЧТО МАССИВ 'M' ПРАВИЛЬНЫЙ
C   ,Т.Е. 'M' ПРИНИМАЕТ ЗНАЧЕНИЯ ИЗ (0,1)
C   ,В 'M' СУЩЕСТВУЕТ ХОТЯ БЫ ОДИН ЭЛЕМЕНТ = '1'
C   ,И В 'M' СУЩЕСТВУЕТ ХОТЯ БЫ ОДИН ЭЛЕМЕНТ = '0'
C   ЕСЛИ 'N' - НЕПРАВИЛЬНОЕ, ТО ПРОВЕРКА НЕВОЗМОЖНА
      KM = 11
      IF (KN.GT.1) GO TO 40
C   ДОПУСТИМ, ЧТО 'M' - НОРМАЛЬНЫЙ
      KM = 0
      N0 = 0
      N1 = 0
      DO 35 I=1,N
      L = M(I)
      IF (L.EQ.0) N0=N0+1
      IF (L.EQ.1) N1=N1+1
35  CONTINUE
      IF (N0+N1.NE.N.OR.N0.EQ.0.OR.N1.EQ.0) KM = 3
      40 IERW(4) = KM
C
C-#-      ПРОВЕРКА 'Q' ( 5-Я РАЗРЯД )
C   НОРМАЛЬНОЕ ЗНАЧЕНИЕ 'Q' : 0.025 <= Q <= 0.2
      KQ = 0
C   ЕСЛИ 'Q' БОЛЬШЕ ЧЕМ 0.2 ИЛИ МЕНЬШЕ ЧЕМ 0.025,
C   ТО СЧИТАЕМ ЧТО ДОПУЩЕНА ЛОГИЧЕСКАЯ ОШИБКА
      IF (Q.LT.0.025) KQ = 1
      IF (Q.GT.0.2) KQ = 1
C   ВВОДИМ ОГРАНИЧЕНИЕ : 0.005 <= Q <= 0.5
      IF (Q.LT.0.005) KQ = 3
      IF (Q.GT.0.5) KQ = 3
C   ФОРМИРОВАНИЕ КОДА ОШИБКИ
      IERW(5) = KQ
      IERW(8) = 5
C
C   ОПРЕДЕЛЕНИЕ МАКСИМАЛЬНОГО КОДА ОШИБКИ
C
      MNC = MAX0(KQ,KM,KX,KN)
C
C   ШКАЛА КОДИРОВКИ ОШИБОЧНЫХ ПАРАМЕТРОВ :
C
C   РАЗРЯД КОДА ОШИБКИ  I  1  I  2  I  3  I  4  I  5
C   -----

```

```

С ПАРАМЕТР В ОБРАЩЕНИИ I NFWR I N I X I M I Q
С -----
С ВОЗМОЖНЫЕ ЗНАЧЕНИЯ I 0 I 0 I 0 I 0 I 0
С I - I 1 I - I - I 1
С I - I - I - I - I -
С I 3 I 3 I 3 I 3 I 3
С I - I - I 11 I 11 I -
С
С ПЕЧАТЬ ВХОДНЫХ ДАННЫХ
С
      IF (KF .GT. 0) GOTO 999
      IF (KEY(2) .EQ. 0) GOTO 225
С
      WRITE (NFWR,220) NAME, MWC
      220 FORMAT(/' ',A8,4X,'КОД МАКСИМАЛЬНОЙ ОШИБКИ =' ,I2)
С
      225 IF (MWC .LT. 1) GOTO 310
С
      J=IERW(10)
      WRITE (NFWR,227) (IERW(I),I=1,J)
      227 FORMAT(/' ',8X,'ИМЕНА ПАРАМЕТРОВ : '
      *      ' , NFWR, N, X, M, Q'
      *      /' ',8X,'К О Д Ы ОШИБОК : ',5(3X,I2)/)
С
      310 IF (KEY(3) .EQ. 0) GOTO 710
      WRITE (NFWR,330) Q
      330 FORMAT(/' ',8X,'УРОВЕНЬ ЗНАЧИМОСТИ РИСКА =' ,F8.3)
С
      710 IF (KEY(7) .EQ. 0) GO TO 810
      WRITE (NFWR,760)
      760 FORMAT(' ',20X,'К О Д Ы ВЫБОРОК : '/')
С
      DO 770 J=1,2
      770 WRITE (NFWR,775) J,(KSON(I,J),I=1,4)
      775 FORMAT(' ',8X,I2,' - ',4A8)
С
      810 IF (KEY(8) .EQ. 0) GO TO 999
      WRITE (NFWR,880) N
      880 FORMAT(/' ',20X,'ДЛИНА ВЫБОРОК =' ,I5)
С
      IF (KN .EQ. 3) GO TO 999
С
      WRITE (NFWR,885) X
      885 FORMAT(' ',13X
      *      , 'МАССИВ ЗНАЧЕНИЙ В ШКАЛЕ ОТНОШЕНИЯ : ')
      *      /,(' ',8X,4G13.5))
С
      WRITE (NFWR,890) M
      890 FORMAT(' ',13X
      *      , 'МАССИВ ЗНАЧЕНИЙ В ШКАЛЕ НАИМЕНОВАНИЙ : ')
      *      /,(' ',8X,4I13))
С
      999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
      END

```

```

C; PR482 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP482
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫВОДА В ФАЙЛ 'NFWR'
C; РЕЗУЛЬТАТОВ РАБОТЫ П/П SP482
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP482
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММОФ,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE PR482(KEY,NFWR,KSQN,C,TQV,IND,IER)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CP ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- 4 МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- 1 НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
C Ksqn R*8(4,2) 4A8 7 МАССИВ К О Д О В ВЫБОРОК
C C R*4 1 G16.5 3 КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ
C TQV R*4 1 G16.5 3 КРИТИЧЕСКОЕ ЗНАЧЕНИЕ Т-РАСПРЕДЕЛЕНИЯ
C IND I*4 1 I1 2 ИНДЕКС ПРОВЕРКИ:
C IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ (НЕТ КОРР. СВЯЗИ);
C IND = 1 - НУЛЕВАЯ ГИПОТЕЗА ОТВЕРГАЕТСЯ (ЕСТЬ КОРР. СВЯЗЬ);
C IER I*4 1 I3 2 КОД ОШИБКИ, ВЫРАБОТАННЫЙ SP482
C
C DIMENSION KEY(8),Ksqn(4,2)
C REAL*8 TNK(3),TS1(4),TS2(4),TS3(4),NAME,Ksqn
C DATA TNK /'ТОЧ. БИС','ЕР. КОЭФ','Ф. КОРР.'/
C * ,TS1 /'ВЫЧИСЛ. ','ТОЧЕЧНОГ','О БИСЕРИ','АЛЬНОГО '/
C * ,TS2 /'КОЭФФ. К','ОРРЕЛЯЦИ','И 2 ВЫБ.','ОДНА -В'/
C * ,TS3 /'ШКАЛЕ НА','ИМЕНОВАН','ИИ,2-Я О','ТНОШЕНИИ'/
C DATA NAME /' PR482 '/
C
C IF (NFWR.LE.0.OR.NFWR.GE.100) GO TO 999
C
C KLU=KEY(1)
C IF (KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 11 KLU=IABS(KLU)
C DO 13 J=1,8
C I=9-J
C KEY(I)=MOD(KLU,2)
C KLU=KLU/2
C 13 GOTO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 15 DO 17 I=1,8
C KEY(I)=1
C IF (KLU.LT.I) KEY(I)=0
C 17 CONTINUE
C 19 CONTINUE
C
C IF (KEY(1) .EQ. 0) GOTO 210
C
C WRITE (NFWR,120) TS1,TNK,TS2,TS3
C 120 FORMAT(/ 1H ,26(1H*))
C * /2H *,24(1H ),3H* ,4A8
C * /2H *,3A8 ,3H* ,4A8
C * /2H *,24(1H ),3H* ,4A8
C * / 1H ,26(1H*) )
C
C 210 IF (KEY(2) .EQ. 0) GOTO 310

```

**ЗАДАНИЕ - ТЕСТ НА ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТА КОРРЕЛЯЦИОННОЙ
СВЯЗИ ДВУХ ВЫБОРОЧНЫХ СОВОКУПНОСТЕЙ, ОДНА ИЗ
КОТОРЫХ РЕЗУЛЬТАТ ИЗМЕРЕНИЙ В ШКАЛЕ ОТНОШЕНИЙ,
А ДРУГАЯ - В НОМИНАЛЬНОЙ ШКАЛЕ, В ПРЕДПОЛОЖЕНИИ
ДИХОТОМИИ НА ОСНОВЕ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ**

```

----      ВЫЧИСЛЕНИЕ КОРРЕЛЯЦИОННОЙ СВЯЗИ ДВУХ ВЫБОРОК :
****      МАКЕТ ПОДГОТОВКИ ДАННЫХ ДЛЯ IR= 17:      ****SF472****
IRP -017      (I3)- КОД ФУНКЦИИ
NS -02 /ПО УМОЛЧАНИЮ/ (I2)- КОЛИЧЕСТВО ВЫБОРОК
K -01 /ПО УМОЛЧАНИЮ/ (I2)- КОЛИЧЕСТВО ГРАДАЦИЙ ВЫБОРКИ 1
M -01 /ПО УМОЛЧАНИЮ/ (I2)-
      (I2)-НОМЕР; (4A8)- К О Д ВЫБОРКИ;      (I4)-ДЛИНА
      XX      XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX XXXX
KSON-01      ТЕСТОВАЯ /2 ГРАДАЦИИ/ ВЫБОРКА 7A 0018
KSON-02      ТЕСТОВАЯ/ШК.ОТНОШЕНИЙ/ВЫБОРКА 7B 0018
      (I2, I3,I2) [ 4A8 I:НОМЕР ВЫБ.,КОД,КАНАЛ;ИМЯ ФАЙЛА
FILX-01 000 08-INF:DATA7A.DAT
      [ СПИСОК СПЕЦИФИКАЦИИ ОПЕРАТОРА F O R M A T (14A4) ]
FORX-01 (4X,30I2)
      . 1 0 0 1 1 1 0 0 1 0 1 0 1 1 1 1
FILZ-01 000 08-INF:DATA71.DAT
FORZ-01 (4X,6F10.2)
      16.00      12.00      11.00      7.00      15.00      14.00
      10.00      11.00      15.00      9.00      13.00      7.00
      13.00      11.00      10.00      11.00      10.00      11.00
Q - 0.05      <- (G14.7)- УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
NFWR-02      (I2) - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА
      (I1,I2)-ТИП КЛЮЧЕЙ ПЕЧАТИ: 0-ПОРАЗРЯДНЫЙ; И ИХ КОЛИЧЕСТВО
KEYC-0 02      1-ИНТЕРВАЛЬНЫЙ.
      ИНДЕКСЫ/ИНТЕРВАЛЬНОЕ ЗНАЧЕНИЕ КЛЮЧА ПЕЧАТИ/ПОРАЗРЯДНОЕ/
      12345678 - РАЗРЯДЫ КЛЮЧА ПЕЧАТИ :
KEY - 1 7 4/8/11111111/ ВХОДНЫХ ДАННЫХ ДЛЯ SF472 ;
KEY - 2 7 4/8/11111111/ РЕЗУЛЬТАТОВ РАБОТЫ SF472 ;
      (I1)-РАСПРЕДЕЛЕНИЕ ПАМЯТИ:=1-ВЫДАЕТСЯ ВСЕГДА;
END -1      КОНЕЦ ДАННЫХ;      =0-ВЫДАЕТСЯ ТОЛЬКО АВАРИЙНО.

```

РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ ЗАДАНИЯ (РАСПЕЧАТКА РЕШЕНИЯ) :

```

IRP - 17
KSNQ NN      XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX      NNNN
KSNQ- 1      ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА 7С      18
KSNQ- 2      ТЕСТОВАЯ/ШК.ОТНОШЕНИЯ/ВЫБОРКА 7D      18
FILX- 1      0      8-INF:DATA7A.DAT
FORX- 1:(4X,30I2)
      1 0 0 1 1 1 0 0 1 0 1 0 1 1 0 1 1 1
FILZ- 1      0      8-INF:DATA7I.DAT
FORZ- 1:(4X,6F10.2)
      16.00      12.00      11.00      7.00      15.00      14.00
      10.00      11.00      15.00      9.00      13.00      7.00
      13.00      11.00      10.00      11.00      10.00      11.00
Q          0.5000000E-01
NFWR 2
KEYC 0      2      4
KEY 1 7 4/8/11111111/
KEY 2 7 4/8/11111111/
ОБЪЕМ РАБОЧЕЙ ПАМЯТИ - 4000 ЕДИНИЦ INTEGER
ТРЕБУЕМЫЙ ОБЪЕМ ПАМЯТИ - 195 ЕДИНИЦ INTEGER

```

ИНДЕКСЫ НАЧАЛА ПЕРЕМЕННЫХ В РАБОЧЕЙ ПАМЯТИ :

| | | | | | | | |
|-----|---|------|-----|------|-----|------|-----|
| 1 - | 1 | 6 - | 6 | 11 - | 107 | 16 - | 83 |
| 2 - | 2 | 7 - | 7 | 12 - | 125 | 17 - | 84 |
| 3 - | 3 | 8 - | 36 | 13 - | 126 | 18 - | 176 |
| 4 - | 4 | 9 - | 19 | 14 - | 64 | 19 - | 178 |
| 5 - | 5 | 10 - | 105 | 15 - | 65 | 20 - | 180 |

СW472 КОД МАКСИМАЛЬНОЙ ОШИБКИ = 0

К О Д Ы ВЫБОРОК :

1 - ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА 7С
 2 - ТЕСТОВАЯ/ШК.ОТНОШЕНИЯ/ВЫБОРКА 7D
 ДЛИНА ВЫБОРОК = 18

МАССИВ ЗНАЧЕНИЙ В ШКАЛЕ ОТНОШЕНИЯ :

| | | | |
|--------|--------|--------|--------|
| 15.000 | 12.000 | 11.000 | 7.0000 |
| 15.000 | 14.000 | 10.000 | 11.000 |
| 15.000 | 9.0000 | 13.000 | 7.0000 |
| 13.000 | 11.000 | 10.000 | 11.000 |
| 10.000 | 11.000 | | |

МАССИВ ЗНАЧЕНИЙ В ШКАЛЕ НАИМЕНОВАНИЙ :

| | | | |
|---|---|---|---|
| 1 | 0 | 0 | 1 |
| 7 | 1 | 0 | 0 |
| 1 | 0 | 1 | 0 |
| 1 | 1 | 0 | 1 |
| 1 | 1 | | |

```

*****
*                               *
* БИСЕРИАЛЬНЫЙ КОЭФФ. КОРР*   * ВЫЧИСЛЕНИЕ БИСЕРИАЛЬНОГО КОЭФФ.
*                               *   * КОРРЕЛЯЦИИ ДВУХ ВЫБОРОК, ОДНА -В
*                               *   * ШКАЛЕ НАИМЕНОВАНИЙ, 2-Я ОТНОШЕНИЙ
*****

```

PR472 - СПРАВЕДЛИВА ГИПОТЕЗА Н0 ; КОД ОШИБКИ = 0

КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ = 0.60860
 СТАНДАРТНОЕ ОТКЛОНЕНИЕ = 0.29969
 СТАТИСТИКА = 2.0308

1 - ТЕСТОВАЯ /2 ГРАДАЦИЯ/ ВЫБОРКА 7С
 2 - ТЕСТОВАЯ/ШК.ОТНОШЕНИЯ/ВЫБОРКА 7D

```

C; SP472 - БИСЕРИАЛЬНЫЙ КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ
C
C; П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫЧИСЛЕНИЯ БИСЕРИАЛЬНОГО КОЭФФИЦИЕНТА
C КОРРЕЛЯЦИИ МЕЖДУ ДВУМЯ ВЫБОРКАМИ, ОДНА ИЗ КОТОРЫХ
C ИЗМЕРЕНА В ШКАЛЕ ОТНОШЕНИЙ, ДРУГАЯ - В ШКАЛЕ
C НАИМЕНОВАНИЙ В ПРЕДПОЛОЖЕНИИ ДИХОТОМИИ НА ОСНОВЕ
C НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ
C
C; ЛИТЕРАТУРА ПО АЛГОРИТМУ : ДЖ. ГЛАСС, ДЖ. СТЭНЛИ,
C СТАТИСТИЧЕСКИЕ ВЫВОДЫ В ПЕДАГОГИКЕ И ПСИХОЛОГИИ, М.,
C ПРОГРЕСС, 1976.
C
C; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
C; П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММО,
C Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
C SUBROUTINE SP472(KEY,NFWR,KSQN,N,X,M,C,D,Z,IND,IER,IERW)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
C; ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C
C KEY I*4 * (8) МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 * 1 НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<NFWR<=99/
C KSQN R*8 * (4,2) МАССИВ К О Д О В ВЫБОРОК
C N I*4 * 1 ДЛИНА ВЫБОРОК "X" И "M" /N>2/
C X R*4 * (N) МАССИВ ЗНАЧ., ИЗМЕР. В ШКАЛЕ ОТНОШЕНИЙ
C M I*4 * (N) МАССИВ ЗНАЧ., ИЗМЕР. В ШК. НАИМЕНОВАНИЙ
C C R*4 * 1 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ КОЭФФ. КОРРЕЛЯЦИИ
C D R*4 * 1 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ СТАНД. ОТКЛОНЕНИЯ
C Z R*4 * 1 НОРМАЛЬНО РАСПРЕД. СТАТИСТИКА: M=0, S=1
C IND I*4 * 1 ИНДЕКС ПРОВЕРКИ ЗНАЧИМОСТИ :
C IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ (НЕТ КОРР. СВЯЗИ);
C IND = 1 - НЕЛЕВАЯ ГИПОТЕЗА ОТКЛОНЯЕТСЯ (ЕСТЬ КОРР. СВЯЗЬ).
C IER I*4 * 1 ВЫРАБОТАННЫЙ КОД ОШИБКИ
C; IERW I*4 * (8) МАССИВ КОДОВ ОШИБОК ПРИНЯТЫЙ ИЗ SW472.
C
C DIMENSION KEY(8),IERW(8),X(N),M(N),KSQN(4,2)
C REAL*8 NAME, KSQN
C DATA NAME /' SP472 '/,ZERO/0.0/
C IND = 11
C IER = 0
C KLP = KEY(2)
C IF (KEY(1).EQ.0) GO TO 120
C CALL SW472(KEY,NFWR,KSQN,N,X,M,IERW,MWC)
C; SW472 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP472;
C
C АНАЛИЗ КОДОВ ОШИБОК : "KF" И "MWC"
C
C KF = IERW(1)
C IF (MWC.LT.3) GO TO 120
C ЕСЛИ MWC >= 3, ТО IER=MWC
C IER = MWC
C IF (MWC.GE.3) GO TO 999
C ВЫЧИСЛЕНИЕ СРЕДНЕГО ЗНАЧЕНИЯ X, СООТВ. '1' ИЛИ '0' ПО 'Y'
120 N0=0
N1=0
X0=ZERO
X1=ZERO
DO 220 I=1,N
IF (M(I).EQ.1) GO TO 211
N0=N0+1
X0=X0+X(I)
GO TO 220

```

6.10. Бисериальный коэффициент корреляции. Бисериальный коэффициент r_{bis} — мера корреляционной связи между двумя выборками X и Y — применяется в том случае, когда одна из выборок (X) — дихотомия на основе нормального распределения, а вторая (Y) — результат количественных измерений по шкале отношений. Бисериальный коэффициент корреляции []

$$r_{\text{bis}} = \frac{\bar{X}_{\bullet 1} - \bar{X}_{\bullet 0}}{S_x} \frac{n_1 n_0}{un_{\bullet\bullet} \sqrt{n_{\bullet\bullet}^2 - n_{\bullet\bullet}}},$$

где $n_1, n_0, n_{\bullet\bullet}, \bar{X}_{\bullet 1}, \bar{X}_{\bullet 0}, S_x$ — определены в п. 6.7, u — ордината нормированного нормального распределения в точке, выше которой находится доля $n_1/n_{\bullet\bullet}$ площади под кривой.

Оценка значимости бисериального коэффициента корреляции осуществляется путем проверки нулевой гипотезы об отсутствии связи между выборками. Для больших выборок ($n_{\bullet\bullet} > 20$), когда выполняется нуль-гипотеза $H_0: r_{\text{bis}} = 0$, выборочное значение r_{bis} имеет распределение, близкое к нормальному, с нулевым средним и стандартным отклонением

$$\sigma_{r_{\text{bis}}} = \frac{\sqrt{n_1 n_0}}{un_{\bullet\bullet} \sqrt{n_{\bullet\bullet}}}.$$

Если нуль-гипотеза выполняется, то $Z = r_{\text{bis}}/\sigma_{\text{bis}}$ будет иметь нормальное распределение с нулевым средним и единичным нормальным отклонением по всем выборкам объема $n_{\bullet\bullet}$.

При отклонении H_0 связь считается существенной при $|Z| \geq Z_q$, где q — квантиль нормированного нормального распределения.

6.11. Программная реализация. Вычисления бисериального коэффициента корреляции в комплексе «СОКИ-Н» может быть представлено в виде блок-схемы на рис. 61,

где *SP472* — подпрограмма вычисления r_{bis} и проверки нулевой гипотезы; *CW472* — подпрограмма контроля и печати входных данных для *SP472*; *PR472* — подпрограмма печати результатов работы *SP472*; *MP050* — подпрограмма вычисления квантилей нормального распределения q .

Тексты этих подпрограмм приведены ниже, после примера.

6.12. Пример. Вычисление коэффициента корреляционной связи двух выборок, одна из которых — результат измерений по шкале отношений, другая — в номинальной шкале, в предложении дихотомии на основе нормального распределения возьмем из [32]. При подготовке файла «входного потока» следует задать:

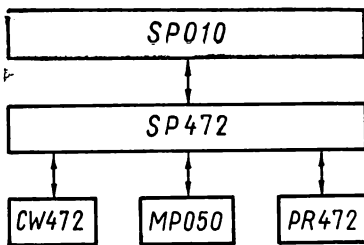


Рис. 61

IRP-017 — код выполняемой функции; коды и длины выборок (длины должны быть равны); указания на местонахождения элементов выборок; форматы записи выборок; элементы выборок (выборка в шкале отношений — числа в форме с плавающей точкой); уровень зна-

чимости риска; ключи печати исходных данных и результатов счета. Ниже приведены текст файла «входного потока» и распечатка выполнения задания на ЭВМ.

```

210  CONTINUE
      N1=N1+1
      X1=X1+X(I)
220  CONTINUE
      FN = FLOAT(N)
      C=(X0+X1)/FN
      X0=X0/N0
      X1=X1/N1
C  ВЬЧИСЛЕНИЕ СРЕДНЕКВАДРАТИЧЕСКОГО ОТКЛОНЕНИЯ ДЛЯ X
      S=ZERO
      DO 310 I=1,N
      S=S+(X(I)-C)**2
310  CONTINUE
      C=FN
      S=SQRT(S/C)
C  ВЬЧИСЛЕНИЕ ОРДИНАТЫ НОРМИРОВАННОГО НОРМАЛЬНОГО РАСПРЕД.
      V=FLOAT(N1)/FN
      CALL MP050(0.0,V,0.001,U,T,IERU)
CL; MP050 - ВЬЧ. ОРДИНАТЫ НОРМИРОВАННОГО НОРМАЛЬНОГО РАСПРЕД.
      IF (IERU.EQ. 0) GO TO 400
      IER = 4+IERU
      GO TO 810
C  ВЬЧИСЛЕНИЕ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ
400  C=(X1-X0)/S*(N1*N0/(U*FN*SQRT(FN*FN-FN)))
C  ВЬЧИСЛЕНИЕ СТАНДАРТНОГО ОТКЛОНЕНИЯ
      D=SQRT(FLOAT(N1*N0))/(U*FN*SQRT(FN))
C  ВЬЧИСЛЕНИЕ Z
      Z=C/D
      IND = 0
      IF (Z.GT.3.0) IND=1
C
810  IF (KF.NE. 3) GO TO 830
      IER = -IER
      GO TO 999
830  IF (KLP.EQ. 0) GO TO 999
      KEY(1) =KLP
      CALL PR472(KEY,NFWR,KSQN,C,D,Z,IND,IER)
CL; PR472 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP472.
C
999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0          2-ФЕВ-88
      END

```

```

C; SW472 - КОНТРОЛЬ И ПЕЧАТЬ ВХОДНЫХ ДАННЫХ ДЛЯ SP472
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ КОНТРОЛЯ ВХОДНЫХ ДАННЫХ ДЛЯ SP472
C И ВЫДАЧИ В ФАЙЛ С НОМЕРОМ "NFWR" РЕЗУЛЬТАТА ПРОВЕРКИ
C И ВХОДНЫХ ДАННЫХ
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP472
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE SW472(KEY,NFWR,KSQN,N,X,M,IERW,MWC)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CР ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
C K SQN R*8 (4,2) 4A8 7 МАССИВ К О Д О В ВЫБОРОК
C N I*4 1 8 ДЛИНА ВЫБОРОК "X" И "M" /N>=2/
C X R*4 (N) -- 8 МАССИВ ЗНАЧЕНИЙ,ИЗМЕР. В ШК.ОТНОШЕНИИ
C M I*4 (N) -- 8 МАССИВ ЗНАЧ.,ИЗМЕР. В ШК.НАИМЕНОВАНИЯ
C IERW I*4 (8) 4I2 [2]ВЫРАБОТАННЫЙ КОД ОШИБКИ : 4 РАЗРЯДОВ
C;MWC I*4 1 I2 2 КОД МАКСИМАЛЬНОЙ ОШИБКИ
C
DIMENSION KEY(8),IERW(8),X(N),M(N),KSQN(4,2)
REAL*8 NAME, KSQN
DATA NAME /' SW472 '/
C
KLU=KEY(1)
IF(KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
11 KLU=IABS(KLU)
DO 13 J=1,8
I=9-J
KEY(I)=MOD(KLU,2)
13 KLU=KLU/2
GOTO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
15 DO 17 I=1,8
KEY(I)=1
IF(KLU.LT.I) KEY(I)=0
17 CONTINUE
19 CONTINUE
C
C КОНТРОЛЬ ВХОДНЫХ ДАННЫХ
C
C-*-- ПРОВЕРКА 'NFWR' ( 1-й РАЗРЯД )
C НАЧАЛЬНАЯ УСТАНОВКА : ОШИБОК НЕТ
KF = 0
C ЕСЛИ 'NFWR' БОЛЬШЕ 99 ИЛИ МЕНЬШЕ 1, ЭТО ГРУБАЯ ОШИБКА
IF (NFWR.GT.99) KF = 3
IF (NFWR.LT.1) KF = 3
C ФОРМИРОВАНИЕ КОДА ОШИБКИ
IERW(1) = KF
C
C-*-- ПРОВЕРКА 'N' ( 2-й РАЗРЯД )
C KN=0 ОЗНАЧАЕТ, ЧТО 'N'- НОРМАЛЬНОЕ, Т.Е. БОЛЬШЕ 9
KN = 0
C ЕСЛИ 'N' МЕНЬШЕ ЧЕМ 10, ТО ЭТО ЛОГИЧЕСКАЯ ОШИБКА
IF (N.LT.10) KN = 1

```

347

```

220 FORMAT(/' ',A8,4X,'КОД МАКСИМАЛЬНОЙ ОШИБКИ =' ,I2)
C
225 IF (MWC .LT. 1) GOTO 710
C
      J=IERW(8)
      WRITE (NFWR,227) (IERW(I),I=1,J)
227  FORMAT(/' ',8X,'ИМЕНА ПАРАМЕТРОВ : '
      *           , ' NFWR,  N,    X,    M'
      *           /' ',8X,'К О Д Ы  ОШИБОК : ',4(3X,I2)/)
C
710 IF (KEY(7) .EQ. 0) GO TO 810
      WRITE (NFWR,760)
760  FORMAT(' ',20X,'К О Д Ы  ВЫБОРОК : '/')
C
      DO 770 J=1,2
770  WRITE (NFWR,775) J,(KSGN(I,J),I=1,4)
775  FORMAT(' ',8X,I2,' - ',4A8)
C
810 IF (KEY(8) .EQ. 0) GO TO 999
      WRITE (NFWR,880) N
880  FORMAT(' ',20X,'ДЛИНА ВЫБОРОК =' ,I5)
C
      IF (KN .EQ. 3) GO TO 999
C
      WRITE (NFWR,885) X
885  FORMAT(/' ',13X
      *           , 'МАССИВ ЗНАЧЕНИЙ В ШКАЛЕ ОТНОШЕНИЙ      *
      *           /,( ' ',8X,4G13.5))
C
      WRITE (NFWR,890) M
890  FORMAT(' ',13X
      *           , 'МАССИВ ЗНАЧЕНИЙ В ШКАЛЕ НАИМЕНОВАНИЙ *
      *           /,( ' ',8X,4I13))
C
999 RETURN
CV1 ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
      END

```

```

CC; PR472 - ПЕЧАТЬ РЕЗУЛЬТАТОВ РАБОТЫ SP472
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫВОДА В ФАЙЛ 'NFWR'
C; РЕЗУЛЬТАТОВ РАБОТЫ П/П SP472
C
CR; ЛИТЕРАТУРА ПО АЛГОРИТМУ : СМ. ОПИСАНИЕ SP472
C
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММОФ,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE PR472(KEY,NFWR,KSQN,C,D,Z,IND,IER)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CP ИМЯ ТИП РАЗМ.ФОРМАТ КЛЮЧ ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C ПЕЧАТИ
C
C KEY I*4 (8) -- - МАССИВ КЛЮЧЕЙ УПРАВЛЕНИЯ
C NFWR I*4 1 -- - НОМЕР ФАЙЛА ДЛЯ ВЫВОДА /1<=NFWR<=99/
C Ksqn R*8(4,2) 4A8 7 МАССИВ К О Д О В ВЫБОРОК
C C R*4 1 G20.5 3 ВЫЧИСЛЕН. ЗНАЧЕНИЕ КОЭФФ. КОРРЕЛЯЦИИ
C D R*4 1 G20.5 3 ВЫЧИСЛЕН. ЗНАЧЕНИЕ СТАНД. ОТКЛОНЕНИЯ
C Z R*4 1 G20.5 3 НОРМАЛЬНО РАСПРЕД.СТАТИСТИКА:M=0,S=1
C IND I*4 1 I1 2 ИНДЕКС ПРОВЕРКИ ЗНАЧИМОСТИ:
C IND = 0 - НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ;
C IND = 1 - НУЛЕВАЯ ГИПОТЕЗА ОТВЕРГАЕТСЯ.
C; IER I*4 1 I3 2 КОД ОШИБКИ, ВЫРАБОТАННЫЙ SP472
C
C
C DIMENSION KEY(8),Ksqn(4,2)
C REAL*8 TNK(3),TS1(4),TS2(4),TS3(4),NAME,Ksqn
C DATA TNK /'БИСЕРИАЛ','ЫННЙ КОЭ','ФФ. КОРР'/
C * ,TS1 /'ВЫЧИСЛЕН','ИЕ БИСЕР','ИАЛЬНОГО',' КОЭФФ. '/
C * ,TS2 /'КОРРЕЛЯЦ','ИИ ДВУХ ','ВЫБОРОК',' ОДНА -В'/
C * ,TS3 /'ШКАЛЕ НА','ИМЕНОВАН','ИИ,2-Я О','ТНОШЕНИИ'/
C DATA NAME /' PR472 '/
C
C IF (NFWR .LE. 0 .OR. NFWR .GE. 100) GO TO 999
C
C KLU=KEY(1)
C IF (KLU) 11,999,15
C - ПОРАЗРЯДНОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 11 KLU=IABS(KLU)
C DO 13 J=1,8
C I=9-J
C KEY(I)=MOD(KLU,2)
C 13 KLU=KLU/2
C GO TO 19
C - ПОЗИЦИОННОЕ КОДИРОВАНИЕ КЛЮЧЕЙ УПРАВЛЕНИЯ
C 15 DO 17 I=1,8
C KEY(I)=1
C IF (KLU.LT.I) KEY(I)=0
C 17 CONTINUE
C 19 CONTINUE
C
C IF (KEY(1) .EQ. 0) GO TO 210
C
C WRITE (NFWR,120) TS1,TNK,TS2,TS3
C 120 FORMAT(/ 1H ,26(1H*))
C * /2H *,24(1H ),3H* ,4A8
C * /2H *,3A8 ,3H* ,4A8
C * /2H *,24(1H ),3H* ,4A8
C * / 1H ,26(1H*) )

```

```

C
210 WRITE (NFWR,230) NAME, IND, IER
230 FORMAT(/' ',A8,' - СПРАВЕДЛИВА ГИПОТЕЗА H',I1
*      ', ' ', '      КОД ОШИБКИ =' ,I3)

C
310 IF (KEY(3).EQ.0) GO TO 710
      WRITE (NFWR,340) HIQ,THI
340 FORMAT(
*      ', ',8X,'ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ КРИТЕРИЯ КОКРЕНА      =' ,B12.5
*/' ',8X,'КРИТИЧЕСКОЕ ЗНАЧЕНИЕ РАСПРЕД. ХИ-КВАДРАТ      =' ,B12.5
*      )

C
710 IF ((KEY(7).EQ.0).OR.(L.EQ.1)) GO TO 999
      WRITE (NFWR,750) ((KSN(I,J),I=1,4),J=1,L)
750 FORMAT(/' ',8X,'ТАБЛИЦА СФОРМИРОВАНА ИЗ ВЫБОРОК:
*      /(' ',8X,5X,4A8))

C
999 RETURN
CV: ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
      END

```

```

С; МР010 - ПРОЦЕНТНЫЕ ТОЧКИ Т-РАСПРЕДЕЛЕНИЯ
С
С; П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫЧИСЛЕНИЯ КРИТИЧЕСКИХ
С; ЗНАЧЕНИЙ РАСПРЕДЕЛЕНИЯ СТЬЮДЕНТА
С;
С; ЛИТЕРАТУРА ПО АЛГОРИТМУ :
С;
С; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
С
С; П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ,
С; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
С
С; SUBROUTINE МР010(Q,NU,EPS,TQV,IER)
С
С; ОПИСАНИЕ ПАРАМЕТРОВ :
С; -----
С
С; ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
С
С; Q R*4 * 1 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА /Q=.1,Q=.05/
С; NU I*4 * 1 СТЕПЕНЬ СВОБОДЫ / N>=1 /
С; EPS R*4 * 1 ТОЧНОСТЬ ВЫЧИСЛЕНИЯ (НЕ ИСПОЛЬЗУЕТСЯ).
С; TQV R*4 * 1 ВЫЧИСЛЕННОЕ КРИТИЧЕСКОЕ ЗНАЧЕНИЕ
С; IER I*4 * 1 К О Д ОШИБКИ:
С
С; IER = 0 - ОШИБОК НЕТ;
С; IER = 1 - ТРЕБУЕМАЯ ТОЧНОСТЬ НЕ ДОСТИГНУТА;
С; IER = 2 - ОШИБКА ПРИ ВЫЧИСЛЕНИИ РЕЗУЛЬТАТА;
С; IER = 4 - ОШИБКА В ЗАДАНИИ ПАРАМЕТРА "Q";
С; IER = 5 - ОШИБКА В ЗАДАНИИ ПАРАМЕТРА "NU".
С
С; REAL*4 TQN(40,2)
С; DATA TQN
1/6.314,2.920,2.353,2.132,2.015,1.943,1.895,1.860
2,1.833,1.812,1.796,1.782,1.771,1.761,1.753,1.746
3,1.740,1.734,1.729,1.725,1.721,1.717,1.714,1.711
4,1.708,1.706,1.703,1.701,1.699,1.697,1.690,1.684
5,1.680,1.676,1.671,1.667,1.665,1.662,1.661,1.645
6,12.706,4.303,3.182,2.776,2.571,2.447,2.365,2.306
7,2.262,2.228,2.201,2.179,2.160,2.145,2.131,2.120
8,2.110,2.101,2.093,2.086,2.080,2.074,2.069,2.064
9,2.060,2.056,2.052,2.048,2.045,2.042,2.030,2.021
A,2.014,2.008,2.000,1.994,1.989,1.986,1.982,1.960/
С; DATA FIVE/5.0/,TEN/10.0/,TS/1000.0/
С; IF (Q.EQ.0.05.OR.Q.EQ.0.1) GO TO 10
С; IER = 4
С; GO TO 999
10 IF(NU .GE. 1) GOTO 20
С; IER = 5
С; GOTO 999
20 N = NU
С; IER = 0
С; K=1
С; IF (Q.LT.0.1) K=2
С; IF (N.GT.30) GO TO 100
С; TQV=TQN(N,K)
С; GO TO 999
100 IF (N.GT.50) GO TO 200
С; I=(N-30)/5
С; J=11+I
С; L=J-1
С; P=FLOAT(N-30-5*I)
С; TQV=(TQN(J,K)-TQN(L,K))*(P/FIVE)+TQN(L,K)
С; GO TO 999
200 IF (N.GT.100) GO TO 300
С; I=(N-50)/10
С; J=15+I

```

```

L=J-1
P=FLOAT(N-50-10*I)
TQV=(TQN(J,K)-TQN(L,K))* (P/TEN)+TQN(L,K)
GO TO 999
300 IF (N.GT.1100) N=1100
P=FLOAT(N-100)
TQV=(TQN(40,K)-TQN(39,K))* (P/TS)+TQN(39,K)
999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
END
CC; МР020 - ВЫЧИСЛЕНИЕ Q-КВАНТИЛЕЙ БИНОМИАЛЬНОГО РАСПРЕДЕЛЕНИЯ
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫЧИСЛЕНИЯ КРИТИЧЕСКИХ ЗНАЧЕНИЙ
C; БИНОМИАЛЬНОГО РАСПРЕДЕЛЕНИЯ
C
CR ЛИТЕРАТУРА ПО АЛГОРИТМУ :
C;
CT; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММОФ,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE МР020(Q,NU,EPS,KBR,IER)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C
C -----
C
CR ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C
C Q R*4 * 1 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА / Q=0.05, 0.01 /
C NU I*4 * 1 ЧИСЛО СТЕПЕНЕЙ СВОБОДЫ /6<=NU<=50/
C EPS R*4 * 1 ТОЧНОСТЬ ВЫЧИСЛЕНИЯ /НЕ ИСПОЛЗУЕТСЯ/
C KBR R*4 * 1 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ КВАНТИЛЯ
C IER I*4 * 1 ВЫРАБОТАННЫЙ КОД ОШИБКИ:
C IER = 0 - НЕТ ОШИБКИ;
C IER = 3 - НЕТ РЕШЕНИЯ ПРИ ЗАДАННЫХ "Q" И "NU";
C IER = 4 - ОШИБКА В ЗАДАНИИ ПАРАМЕТРА "Q";
C; IER = 5 - ОШИБКА В ЗАДАНИИ ПАРАМЕТРА "NU",
C
INTEGER K81 (45), K85 (45)
DATA K85 /
1 6, 7, 2*8, 9, 2*10, 11, 2*12, 2*13, 14
2 , 2*15, 16, 2*17, 2*18, 19, 2*20, 2*21, 22, 2*23
3 , 2*24, 2*25, 26, 2*27, 2*28, 2*29, 30, 2*31, 2*32
4 , 33
5 ,
6 K81 / ,
7 3*999, 9, 10, 2*11, 12, 2*13, 14, 2*15, 16, 2*17, 18, 2*19,
8 2*20, 21, 2*22, 23, 2*24, 2*25, 26, 2*27, 2*28, 29, 2*30,
9 2*31, 32, 2*33, 34, 2*35 /
C
IER=0
C ПРОВЕРКА ЗНАЧЕНИЯ "Q"
IF (Q.EQ.0.05 .OR. Q.EQ.0.01) GO TO 20
IER = 4
GO TO 999
C ПРОВЕРКА ЗНАЧЕНИЯ "NU"
20 IF (NU.GE.6 .OR. NU.LE.50) GO TO 100
IER = 5
GO TO 999
C ВЫЧИСЛЕНИЕ "K85":
100 IF (Q.EQ.0.05) GO TO 520
C Q=0.01; 9<=NU<=50
IF (NU.GE.9) GO TO 120
IER = 3
GO TO 999
120 KBR = K81 (NU-5)
GO TO 999
C Q=0.05; 6<=NU<=50
520 KBR = K85 (NU-5)
C
999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
END

```



```

C; MP030 - ПРОЦЕНТНЫЕ ТОЧКИ РАСПРЕДЕЛЕНИЯ ХИ-КВАДРАТ
C
C; П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫЧИСЛЕНИЯ КРИТИЧЕСКИХ
C; ЗНАЧЕНИЙ КРИТЕРИЯ ХИ-КВАДРАТ
C
C; ЛИТЕРАТУРА ПО АЛГОРИТМУ : ALGORITHM AS-91,
C; APPLIED STATISTICS, 1975, VOL.24, NO.3, PP.385-388
C
C; ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : ALGORITHM AS-91,
C; /BAMAIN - ALGORITHM AS 32, 1970, VOL.19, NO.3, PP.285-287/
C; /BAUINV - ALGORITHM AS 70, 1974, VOL.23, NO.1, PP.96-97/
C
C; П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ,
C; Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
C      SUBROUTINE MP030(Q,NU,EPS,THI,IER)
C
C      ОПИСАНИЕ ПАРАМЕТРОВ :
C      -----
C
C; ИМЯ      ТИП ВХ.ВЫХ РАЗМ.      ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C
C; Q        R*4 *      1      УРОВЕНЬ ЗНАЧИМОСТИ РИСКА
C; NU       I*4 *      1      ЧИСЛО СТЕПЕНЕЙ СВОБОДЫ
C; EPS      R*4 *      1      ТОЧНОСТЬ
C; THI      R*4 *      1      ВЫЧИСЛЕННОЕ КРИТИЧЕСКОЕ ЗНАЧЕНИЕ
C; IER      I*4 *      1      КОД ОШИБКИ :
C
C          IER = 0 - НЕТ ОШИБКИ ;
C          IER = 1 - ЕСЛИ  $0.999998 < (1-Q) < 0.000002$  ;
C          IER = 2 - ЕСЛИ  $NU \leq 0$  ;
C          IER = 3 - ЕСЛИ КОД ОШИБКИ П/П BAMAIN  $> 0$  .
C
C      EQUIVALENCE (IER,IFault),(THI,PPCHI2)
C      REAL*8 NAME
C      DATA NAME/' MP030 '/
C
C      P=1.0-Q
C      V=FLOAT(NU)
C      G=ALGAMA (V/2.0)
C *** FUNCTION PPCHI2(P,V,G,IFault) ***
C
C      ALGORITHM AS 91 APPL. STATIST. (1975) VOL.24, NO.3
C
C      TO EVALUATE THE PERCENTAGE POINTS OF THE CHI-SQUARED
C      PROBABILITY DISTRIBUTION FUNCTION.
C      P MUST LIE IN THE RANGE 0.000002 TO 0.999998,
C      V MUST BE POSITIVE,
C      G MUST BE SUPPLIED AND SHOULD BE EQUAL TO LN(GAMMA(V/2.0))
C
C      DATA E,AA/0.5E-4,0.6931471805/
C
C      AFTER DEFINING ACCURACY AND LN( 2), TEST ARGUMENTS AND
C      INITIALIZE
C
C      THI = -1.0
C      IER = 1
C      IF (P.LT. 0.000002 .OR. P .GT. 0.999998) RETURN
C      IER = 2
C      IF (V .LE. 0.0) RETURN
C      IER = 0
C      XX=0.5*V
C      C=XX-1.0
C
C      STARTING APPROXIMATION FOR SMALL CHI-SQUARED
C
C      IF (V.GE. -1.24*ALOG(P)) GO TO 1

```

```

CH= ( P * XX * EXP ( B + XX * AA )) ** (1.0/XX)
IF(CH-E) 6,4,4

C
C
C
STARTING APPROXIMATION FOR V LESS THEN OR EQUAL TO 0.32

1 IF (V.GT.0.32) GO TO 3
CH=.4
A=ALOG(1.0-P)
2 D=CH
P1=1.0+CH*(4.67+CH)
P2=CH*(6.73+CH*(6.66+CH))
T=-0.5+(4.67+2.0*CH)/P1-(6.73+CH*(13.32+3.0*CH))/P2
CH=CH-(1.0-EXP(A+B+0.5*CH+C*AA)*P2/P1)/T
IF (ABS(D/CH-1.0)-0.01) 4,4,2

C
C
C
CALL TO ALGORITHM AS 70 - NONE THAT P HAS BEEN TESTED ABOVE

3 X = GAUINV (P,IF1)

C
C
C
STARTING APPROXIMATION USING WILSON AND HILFERTY ESTIMATE

P1 = 0.222222/V
CH = V*(X*SQRT(P1)+1.0-P1)**3

C
C
C
STARTING APPROXIMATION FOR P TENING TO 1

IF (CH .GT. 2.2*V+6.0)
& CH =-2.0*(ALOG(1.0-P)-C*ALOG(0.5*CH)+B)
CALL TO ALGORITHM AS 32 AND CALCULATION OF SEVEN TERM
TAYLOR SERIES

C
C
C
4 D = CH
P1= 0.5*CH
P2= P-GAMAIN(P1,XX,B,IF1)
IF (IF1.EQ.0) GOTO 5
IER = 3
RETURN
5 T = P2*EXP(XX*AA+B+P1-C*ALOG(CH))
B = T/CH
A = 0.5*T-B*C
S1= (210.0+A*(140.0+A*(105.0+A*(84.0+A*(70.0+60.0*A))))/420.0
S2= (420.0+A*(735.0+A*(966.0+A*(1141.0+1278.0*A)))/2520.0
S3= (210.0+A*(462.0+A*(707.0+932.0*A)))/2520.0
S4=(252.0+A*(672.0+1182.0*A)+C*(294.0+A*(889.0+1740.0*A)))/5040.0
S5= (84.0+264.0*A+C*(175.0+606.0*A))/2520.0
S6= (120.0+C*(346.0+127.0*C))/5040.0
CH= CH+T*(1.0+0.5*T*S1-B*C*(S1-B*(S2-B*(S3-B*(S4-B*(S5-B*S6))))))
IF (ABS(D/CH-1.0).GT.E) GOTO 4

6
6 THI = CH
RETURN
CV1 ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
END

```

```

CC: MP040 - КРИТИЧЕСКИЕ ЗНАЧЕНИЯ F-РАСПРЕДЕЛЕНИЯ
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫЧИСЛЕНИЯ КРИТИЧЕСКИХ ЗНАЧЕНИЙ
C F-РАСПРЕДЕЛЕНИЯ
C
CR ЛИТЕРАТУРА ПО АЛГОРИТМУ :
C
C;
CT: ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : Н Е Т
C
CA П/П ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММОФ,
C Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
SUBROUTINE MP040(Q,M,S,EPS,TF,IER)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CR ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C
C Q R*4 * 1 УРОВЕНЬ ЗНАЧИМОСТИ РИСКА /Q=0.05/
C M I*4 * 1 СТЕПЕНЬ СВОБОДЫ /3<=M<=12/
C S R*4 * 1 СТЕПЕНЬ СВОБОДЫ /100.0<=S<=1.E37/
C EPS R*4 * 1 ТОЧНОСТЬ ВЫЧИСЛЕНИЯ (НЕ ИСПОЛЬЗУЕТСЯ)
C TF R*4 * 1 ВЫЧИСЛЕННОЕ ЗНАЧЕНИЕ РАСПРЕДЕЛЕНИЯ
C IER I*4 * 1 ВЫРАБОТАННЫЙ КОД ОШИБКИ:
C IER = 0 - НЕТ ОШИБКИ;
C IER = 1 - ОШИБКА В ЗАДАНИИ ПАРАМЕТРА Q;
C IER = 2 - ОШИБКА В ЗАДАНИИ ПАРАМЕТРА M;
C IER = 3 - ОШИБКА В ЗАДАНИИ ПАРАМЕТРА S.
C IER = 4 - ОШИБКА ПРИ ВЫЧИСЛЕНИИ РЕЗУЛЬТАТА "TF";
C IER = 5 - ТРЕБУЕМАЯ ТОЧНОСТЬ НЕ ДОСТИГНУТА.
C
DIMENSION T(9,10),D(9)
DATA D
*/100.0,125.0,150.0,200.0,300.0,400.0,500.0,1000.0,1.E+37/
DATA T
1/ 2.70, 2.68, 2.67, 2.65, 2.64, 2.63, 2.62, 2.61, 2.60
2, 2.46, 2.44, 2.43, 2.42, 2.41, 2.40, 2.39, 2.38, 2.37
3, 2.30, 2.29, 2.27, 2.26, 2.25, 2.24, 2.23, 2.22, 2.21.
4, 2.19, 2.17, 2.16, 2.14, 2.13, 2.12, 2.11, 2.10, 2.09
5, 2.10, 2.08, 2.07, 2.05, 2.04, 2.03, 2.03, 2.02, 2.01
6, 2.03, 2.01, 2.00, 1.98, 1.97, 1.96, 1.96, 1.95, 1.94
7, 1.97, 1.95, 1.94, 1.92, 1.91, 1.90, 1.90, 1.89, 1.88
8, 1.92, 1.90, 1.89, 1.87, 1.86, 1.85, 1.85, 1.84, 1.83
9, 1.88, 1.86, 1.85, 1.83, 1.82, 1.81, 1.81, 1.80, 1.79
A, 1.85, 1.83, 1.82, 1.80, 1.79, 1.78, 1.77, 1.76, 1.75
*/
C
IF (M .LT. 03) GO TO 20
IF (M .GT. 12) GO TO 20
IF (S .LT. 100.0 ) GO TO 30
IF (S .GT. 1.E+37) GO TO 30
C *** IF (EPS .LT. 5.E-4 ) GO TO 10
IF (Q .NE. 0.05 ) GO TO 10
IER = 0
GO TO 100
C
10 IER = 1
GO TO 999
C
20 IER=2
GOTO 999
C
30 IER=3
GOTO 999
C

```

```

      100 N=M-2
      DO 200 I=1,9
        K=I
        Y=D(I)
        IF (S - Y) 300,200,200
C
      200 CONTINUE
        TF=T(K,N)
        GO TO 999
      300 CONTINUE
        X=D(I-1)
        A=T(I-1,N)
        B=T(I,N)
        TF=A+(B-A)*((S-X)/(Y-X))
C
      999 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
      END

```

```

CC; MP050 - ВЫЧИСЛЕНИЕ Q-КВАНТИЛЕЙ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ
C
CF П/П ПРЕДНАЗНАЧЕНА ДЛЯ ВЫЧИСЛЕНИЯ ОРДИНАТЫ НОРМИРОВАННОГО
C НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ С ПАРАМЕТРАМИ (0,1)
C
CR ЛИТЕРАТУРА ПО АЛГОРИТМУ :
C
CT ИСТОЧНИК ТЕКСТА ПРОГРАММЫ : П/П NDTRI
C ИЗ ПАКЕТА НАУЧНЫХ ПРОГРАММ НА ФОРТРАНЕ
C
CA П/П. ПРОВЕРЕНА : НЕЧАЕВ А.Н., КАФЕДРА ЧММФ,
C Ф-Т КИБЕРНЕТИКИ КГУ /КИЕВ/
C
      SUBROUTINE MP050(Q,V,EPS,U,X,IER)
C
C ОПИСАНИЕ ПАРАМЕТРОВ :
C -----
C
CР ИМЯ ТИП ВХ.ВЫХ РАЗМ. ОПИСАНИЕ / ОГРАНИЧЕНИЯ /
C
C Q R*4 * 1 НЕ ИСПОЛЬЗУЕТСЯ В ДАННОМ СЛУЧАЕ
C V R*4 * 1 ВЕРОЯТНОСТЬ
C EPS R*4 * 1 НЕ ИСПОЛЬЗУЕТСЯ В ДАННОМ СЛУЧАЕ
C U R*4 * 1 ПЛОТНОСТЬ ВЕРОЯТНОСТИ
C X R*4 * 1 ОРДИНАТА
C IER I*4 * 1 ВЫРАБОТАННЫЙ К О Д ОШИБКИ:
C IER = 0 - ОШИБОК НЕТ;
C IER = 1 - ОШИБКА В ЗАДАНИИ ПАРАМЕТРА "P" (P<0).
C
      IER=0
      IF(V) 1,4,2
1 IER = 1
      GOTO 12
2 IF(V-1.0) 7,6,1
4 X=-0.999999E37
5 U = 0.0
      GOTO 12
6 X= 0.999999E37
      GOTO 5
7 U = V
      IF(U-0.5) 9,9,8
8 U=1.0-U
9 T2=ALOG(1.0/(U*U))
      T=SQRT(T2)
      X=T-(2.515517+0.802853*T+0.010328*T2)/
      & (1.0+1.432788*T+0.189269*T2+0.001308*T*T2)
      IF (V-0.5) 10,10,11
10 X=-X
11 U=0.3989423*EXP(-X*X*0.5)
12 RETURN
CV; ВЕРСИЯ / МОДИФИКАЦИЯ : 2.0
      END

```

Сделаем лишь ряд замечаний относительно комплекса программ «СОКИ-П» статистической обработки результатов качественных измерений в шкале порядка (рангов), поскольку вследствие ограниченного объема монографии авторы не смогли поместить этот материал полностью.

Измерение в шкале порядка (рангов) применяется, когда исследователь может установить не только отличие одного состояния объекта измерения от другого, но и отношение следования между различными состояниями, т. е. обнаружить различие в степени качественного признака. При построении шкалы порядка используется свойство упорядоченности чисел. Состояниям объекта измерения приписываются числа (ранги) в порядке следования (возрастания или убывания) степени качественного признака. Структура порядковой шкалы сохраняется при любой взаимно однозначной подстановке, которая сохраняет порядок. Арифметические операции для шкалы порядка (рангов) не имеют смысла. Подходящими мерами положения центра являются **медиана** и **мода**. В шкалу порядка всегда можно преобразовать измерения, произведенные по количественным (метрическим) шкалам.

Комплекс программ «СОКИ-П» статистической обработки результатов качественных измерений в шкале порядка (рангов) применяется для решения следующих типовых статистических задач:

- статистическая обработка одной выборки;
- оценка различия двух независимых выборок;
- оценка различия «М» независимых выборок;
- оценка различия двух связанных выборок (опыт «до — после»);
- оценка корреляционной связи двух выборок;
- обработка коллективной экспертизы по оценке относительной важности объекта;

- обработка коллективной экспертизы по оценке времени совершения определения события.

Для ввода в ЭВМ результаты измерений представляются в виде чисел в форме с плавающей точкой (REAL*4). В случае оценки корреляционной связи между двумя выборками, одна из которых — результат измерения в шкале рангов, а другая — в номинальной шкале, элементы последней вводятся в виде целых чисел типа INTEGER. Кроме результатов измерений, для работы комплекса «СОКИ-П» необходимо задавать код типовой задачи, количество выборок и их длины, уровень значимости риска и другие параметры. Дополнительно задается служебная информация: идентификационные коды выборок, коды управления печатью, номер логического канала ввода. Структура комплекса программ «СОКИ-П» представлена на рис. 62.

Главная программа комплекса «СОКИ-П» резервирует пространство, запрашивает имя входного файла с информацией, вызывает подпрограмму ДАТАР для ввода данных и, затем, подпрограмму SPO20 управления комплексом для обработки данных. После окончания обработки главная программа запрашивает условие прекращения работы и завершает выполнение либо запрашивает снова имя входного файла и имитирует новый цикл обработки данных и т. д.

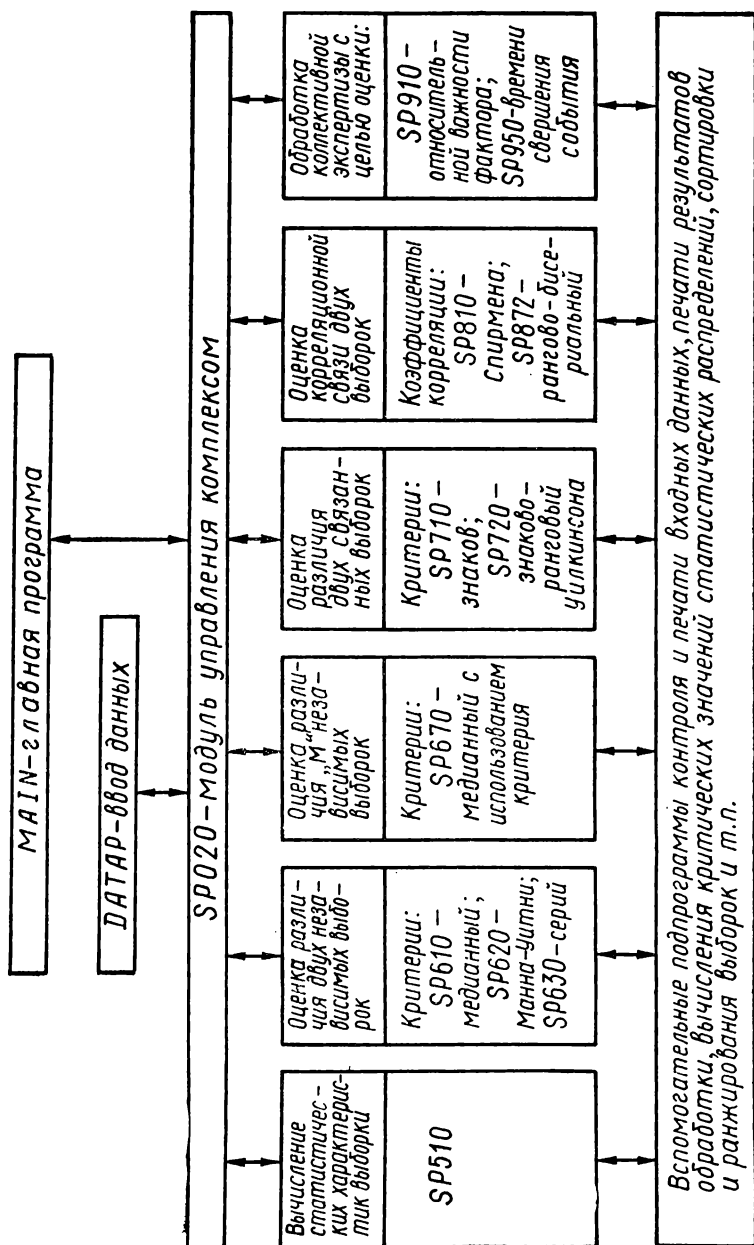


Рис. 62

Подпрограмма ввода данных осуществляет ввод информации из «потока» — файла, имя которого запрашивала главная программа, реализует динамическое распределение рабочей памяти для параметров, данных и рабочих массивов. Информация из «потока» выбирается по ключевым словам, анализируется. В случае обнаружения ошибок выдаются диагностические сообщения.

Информация в «потоке», не запрашиваемая программой ввода, рассматривается как комментарий. В «потоке» возможно использовать ссылки на имена дополнительных файлов, содержащих выборки или таблицы данных.

Подпрограмма ДАТАР осуществляет контрольную печать введенной информации и таблицу распределения динамической памяти.

Модуль управления комплексом (*SP020*) организует вызов функциональных модулей в последовательности, обеспечивающей решение указанной типовой задачи: функциональные модули — подпрограммы, осуществляющие вычисление статистических характеристик, проверки гипотез, нахождение коэффициентов корреляционных связей и оценку их значимости, обработку коллективной экспертизы. Каждой функциональной подпрограмме с префиксом *SP* (кроме *SP020*) соответствует подпрограмма с префиксом *CW* — контроль и печать входных данных и подпрограмма с префиксом *PR* — печать результатов работы. Например, подпрограмме *SP610* соответствует *CW610* и *PR610*.

Подпрограммы вычисления практических значений распределений, ранжирования, сортировки имеют префикс *MP*.

Структура комплекса «СОКИ-П» позволяет выделить три варианта использования программ. Первый — использование отдельных функциональных модулей комплекса, так как каждый такой модуль является самостоятельной единицей, выполняет одну из перечисленных статистических задач. Второй вариант — использование всей группы функциональных модулей в целом под управлением *SP020*, через который и происходит вызов комплекса с указанием кода функции обработки — кода статистической задачи (подзадачи). Оба эти варианта требуют участия программиста в написании главной программы для резервирования пространства, ввода данных, организации вызова комплекса «СОКИ-П» через модуль управления *SP020*. При этом возможно построение весьма сложных систем обработки данных. Третий вариант — использование «автономного варианта» комплекса — загружаемой рабочей программы «СОКИ-П-А», включающей в себя дополнительно универсальную подпрограмму ввода данных ДАТАР и главную программу. Этот вариант не требует знания программирования (не требуется составлять программу). Необходимым условием является только обучение навыкам подготовки информации для решения типовых статистических задач и процедурам запуска программы на выполнение.

В программе «СОКИ-П-А» установлен твердый порядок ввода исходных данных; кроме того, введено единообразие представления информации для большинства решаемых задач. Это позволяет использовать стандартный макет подготовки данных. Такая форма (формуляр, шаблон, макет) обладает определенной избыточностью структуры.

Те элементы макета, которые не требуются данной задаче, рассматриваются как комментарии и пропускаются. Идентификация информации в «потоке» производится по ключевым словам. Данные — элементы выборок, таблицы могут находиться в одном или нескольких файлах. В дополнительных файлах с данными ключевые слова не используются.

В комплексе «СОКИ-П» не предусмотрена программная обработка аварийных ситуаций, поэтому исходные данные необходимо проверять перед обработкой с целью повышения «живучести» комплекса. Эти проверки осуществляют подпрограммы типа *СW*. Проверяются значения параметров и числовые данные на допустимые диапазоны значений и взаимосогласованность. Результаты проверки — числовые коды, в результате анализа которых принимается решение о продолжении работы или аварийном завершении с выдачей на печать кодов проверки данных.

Вывод результатов проверки исходных данных, самих исходных данных и результатов обработки управляется значениями элементов массива ключей печати. Массив ключей печати построен таким образом, что каждой функциональной подпрограмме типа *SP* принадлежат два ключа для подпрограмм *СW* и *PR* соответственно, кроме того, каждый ключ состоит из восьми разрядов, отвечающих разным группам параметров. Это позволяет добиться необходимой детальности вывода.

Комплекс программ «СОКИ-П» написан на языке ФОРТРАН-4. Программа «СОКИ-П-А» выполняется под управлением операционной системы РАФОС (*RT-11*). Для работы «СОКИ-П-А» требуется 20 Кслов оперативной памяти. Объем обрабатываемых выборок — порядка 1000 элементов. Исполняемая программа занимает на диске около 210 блоков (по 512 байт). Программа «СОКИ-П-А» может выполняться на микро-ЭВМ типа ДВК-2 («Электроника — 60 М»), имеющей в своем составе оперативную память 28 Кслов, накопитель на гибких магнитных дисках «Электроника 6022» («Электроника ГМД — 7012»), дисплей. Печатающее устройство не обязательно; длина строки вывода не превышает 64 позиции.

СПИСОК ЛИТЕРАТУРЫ

1. *Автоматизация* сбора, обработки, передачи и отображения радиолокационной информации.— М.: Сов. радио, 1975.— 142 с.
2. *Адлер Ю. П.* Введение в планирование эксперимента.— М.: Металлургия, 1969.— 157 с.
3. *Аркадьев А. Г., Браверман Э. М.* Обучение машины и классификация объектов.— М.: Наука, 1971.— 192 с.
4. *Артемьева Е. П., Мартынов Е. М.* Вероятностные методы психологии.— М.: Изд-во Моск. ун-та, 1975.— 206 с.
5. *Беки Дж. А.* Дискретная модель человека — оператора в системах управления // Тр. II Междунар. конгресса ИФАК, Базель, 27 авг.— 4 сент. 1963 г.— М.: Наука, 1963.— 123 с.
6. *Белов Ю. А., Козлов Н. Н., Ляшко И. И. и др.* Математическое обеспечение сложного эксперимента: В 5-ти т. Т. 2. Математические модели при измерениях.— Киев: Наук. думка, 1982.— 264 с.
7. *Белов Ю. А.* Математическое моделирование и вычислительный эксперимент при исследовании сложных систем.— Киев: Общество «Знание» УССР, 1984.— 16 с.
8. *Бонгард М. М.* Проблема узнавания.— М.: Наука, 1967.— 320 с.
9. *Булгаков А. Я.* Эффективно-вычислительный параметр качества устойчивости системы линейных уравнений с постоянными коэффициентами // Сиб. мат. журн. 1980.— 21, № 3.— 23—30 с.
10. *Бусленко М. П., Калашников В. В., Коваленко И. Н.* Лекции по теории сложных систем.— М.: Сов. радио, 1973.— 432 с.
11. *Валикова Л. Н., Вигдорчик Г. В., Воробьев Ю. А., Лукин А. А.* Операционная система СМ ЭВМ РАФОС: Справочник.— М.: Финансы и статистика, 1984.— 207 с.
12. *Вентцель Е. С.* Теория вероятностей.— М.: Наука, 1969.— 576 с.
13. *Верхаиен К., Дейн Р., Грун и др.* Распознавание образов, состояния и перспективы.— М.: Радио и связь, 1985.— 103 с.
14. *Воскресенский А. Д., Вентцель Н. Д.* Статистический анализ сердечного ритма и показателей гемодинамики в физиологических исследованиях.— М.: Наука, 1974.— 221 с.
15. *Галактионов А. И.* Представление информации оператору.— М.: Энергия, 1969.— 136 с.
16. *Гантмахер Ф. Р.* Теория матриц.— М.: Наука, 1967.— 575 с.
17. *Герасимов Б. М.* Определение количественных характеристик работы оператора методами теории графов // Пробл. инж. психологии.— Ярославль, 1972.— Вып. 4.— 25 с.
18. *Гласс Дж., Стенли Дж.* Статистические методы в педагогике и в психологии.— М.: Прогресс, 1976.— 495 с.
19. *Гнеденко Б. В., Коваленко И. Н.* Введение в теорию массового обслуживания.— М.: Наука, 1966.— 431 с.
20. *Демидович Б. П., Марон И. А.* Основы вычислительной математики.— М.: Наука, 1970.— 664 с.
21. *Джеймс Х., Никольс Н., Филипс Р.* Теория следящих систем.— М.: Изд-во иностр. лит., 1953.— 218 с.
22. *Джури Э.* Импульсные системы автоматического регулирования / Пер. с англ. М. А. Бертмана, Ж. Л. Грина; под ред. Я. З. Цыпкина.— М.: Физматгиз, 1963.— 455 с.

23. *Диамтидес Н. Д.* Оператор как звено в системе управления // *Вопр. ракетной техники.*— 1962.— С. 221—118.
24. *Дэниель К.* Применение статистики в промышленном эксперименте.— М. : Мир, 1979.— 299 с.
25. *Енцков И. С.* Методы, алгоритмы, программы многомерного статистического анализа : Пакет ППСА.— М. : Финансы и статистика, 1986.— 180 с.
26. *Зараковский Г. М.* Психофизиологический анализ трудовой деятельности // *Логико-вероятностный подход при изучении труда управляющего типа.*— М. : Наука, 1966.— 114 с.
27. *Зинченко В. П., Попов Д. Ю.* Узловые проблемы инженерной психологии // *Вопр. психологии.*— 1962.— № 5.— 28 с.
28. *Инженерная психология* / Под ред. Р. К. Середы.— Киев : Вища шк., 1976.— 307 с.
29. *Инженерно-психологическое проектирование АСУ* / Под ред. А. И. Прохорова.— Киев : Будівельник, 1973.— 168 с.
30. *Инженерная психология* // *Теория, методология, практическое применение.*— М. : Наука, 1977.— 304 с.
31. *Картан А.* Дифференциальное исчисление.— М. : Мир, 1971.— 392 с.
32. *Кендалл М., Стюарт А.* Многомерный статистический анализ и временные ряды / Пер. с англ. В. И. Ротаря и Э. Л. Пресмана; под ред. А. Н. Колмогорова, Ю. В. Прохорова.— М. : Наука, 1976.— 736 с.
33. *Кендалл Дж., Стюарт А.* Статистические выводы и связи.— М. : Наука, 1973.— 234 с.
34. *Кокс Д., Смит У.* Теория очередей / Пер. с англ. В. В. Рыкова, Ю. Б. Рождественского; под ред. А. Д. Соловьева.— М. : Мир, 1966.— 218 с.
35. *Конторов Д. С., Голубев-Новожилов Ю. С.* Введение в радиолокационную системотехнику.— М. : Сов. радио, 1971.— 367 с.
36. *Краснощевков П. С., Петров А. А.* Принципы построения модели.— М. : Изд-во Моск. ун-та, 1983.— 264 с.
37. *Ксенз С. П.* Поиск неисправностей в радиоэлектронных системах методом функциональных проб / Под ред. И. Н. Гельфера.— М. : Сов. радио, 1965.— 135 с.
38. *Курдюмов С. П., Малинецкий Г. Г.* Синергетика — теория самоорганизации : Идеи, методы, перспективы.— М. : Знание, 1983.— 64 с.
39. *Курицкий Б. Я.* Математические методы физиологии.— Л. : Наука, 1969.— 292 с.
40. *Линцер Р. Ш., Ширяев А. Н.* Статистика случайных процессов : Нелинейная фильтрация и смежные вопросы.— М. : Наука, 1974.— 696 с.
41. *Ломов Б. Ф., Ширяев А. Н.* Человек и техника : Очерки инженерной психологии.— Л. : Изд-во Ленингр. ун-та, 1963.— 266 с.
42. *Ляпунов А. А., Шестопал Г. А.* Об алгоритмическом описании процессов управления // *Мат. просвещение.*— 1957.— Вып. 2.— 182 с.
43. *Математическое описание характеристик человека — оператора как звена системы управления* // *Вопр. ракетной техники.*— 1965.— № 12.— С. 45—49.
44. *Матюхин Н. Я.* Применение вычислительных машин для проектирования числовых устройств.— М. : Сов. радио, 1968.— 256 с.
45. *Методы анализа данных : Подход, основанный на методике динамических сгущений.*— М. : Финансы и статистика, 1985.— 357 с.
46. *Методические рекомендации по оценке надежности и эффективности систем : (человек — техника)* / Под ред. А. И. Губинского.— Л. : Изд-во ЛГУ, 1972.— 243 с.
47. *Миркин Б. Г.* Анализ качественных признаков и структур.— М. : Статистика, 1980.— 320 с.
48. *Миркин Б. Г.* Группирование социально-экономических исследований.— М. : Финансы и статистика, 1985.— 224 с.
49. *Николаев В. И.* Информационная теория контроля и управления.— Л. : Судостроение, 1973.— 288 с.
50. *Новиков О. А., Петухов С. И.* Прикладные вопросы теории массового обслуживания / Под ред. Б. В. Гнеденко.— М. : Сов. радио, 1969.— 399 с.
51. *Основы автоматического регулирования* / Под ред. В. В. Солодовникова.— М. : Машгиз, 1954.— 1117 с.
52. *Основы инженерной психологии* / Под ред. Б. Ф. Ломова.— М. : Вища шк., 1977.— 335 с.

53. *Перов А. И.* Квазиоптимальная непрерывно-дискретная фильтрация непрерывных процессов // Радиоэлектроника.— 1986.— № 4.— 27—32 с.
54. *Попов Г. П.* Инженерная психология в радиолокации : Система индикатор — оператор.— М. : Сов. радио, 1971.— 143 с.
55. *Портер У.* Современные обоснования общей теории систем / Пер. с англ. Э. Л. Наппельбаума; под ред. С. В. Емельянова.— М. : Наука, 1971.— 555 с.
56. *Помаз Д. Дж.* Справочник по вычислительным методам статистики.— М. : Финансы и статистика, 1982.— 344 с.
57. *Программное обеспечение СМ ЭВМ РАФОС. ФОРТРАН* // Описание языка.— М. : Финансы и статистика.— 1980.— 135 с.
58. *Программное обеспечение СМ ЭВМ РАФОС. ФОРТРАН* // Руководство программиста.— М. : Финансы и статистика.— 1980.— 178 с.
59. *Психологические основы профессионального отбора* // Тез. докл. всесоюз. конф. по инж. психологии.— Киев : Наук. думка, 1973.— 174 с.
60. *Пушкин В. Н.* Психология, кибернетика.— М. : Педагогика, 1971.— 231 с.
61. *Рао С. Р.* Линейные статистические методы и их применение.— М. : Наука, 1968.— 246 с.
62. *Ронжин О. В.* Информационные методы исследования эргатических систем.— М. : Энергия, 1976.— 208 с.
63. *Рунион Р.* Справочник по непараметрической статистике. Современный подход / Пер. с англ. Е. З. Демиденко.— М. : Финансы и статистика, 1982.— 198 с.
64. *Самарский А. А.* Математическое моделирование и вычислительный эксперимент // Вестн. АН СССР.— 1979.— № 5.— 38—49 с.
65. *Самарский А. А.* Введение в численные методы.— М. : Наука, 1987.— 286 с.
66. *Самарский А. А.* Вычислительный эксперимент.— М. : Знание, 1983.— 64 с.
67. *Самарский А. А.* Современные проблемы прикладной математики и математической физики.— М. : Наука, 1988.— 222 с.
68. *Селье Г.* Стресс без дистресса / Пер. с англ. А. Н. Лука, И. С. Хоролла.— М. : Прогресс, 1979.— 124 с.
69. *Сейдж Э., Мелс Дж.* Теория оценивания и ее применение в связи с управлением / Пер. с англ. Б. Р. Левина.— М. : Наука, 1969.— 207 с.
70. *Смирнов Б. А., Самошкина М. Н.* Человек — труд — техника.— Харьков : Пропор, 1975.— 206 с.
71. *Современная теория систем управления* / Пер. с англ. А. Я. Когана.— М. : Наука, 1970.— 511 с.
72. *Солодовников В. В.* Статистическая динамика линейных систем автоматического управления.— М. : Физматгиз, 1960.— 655 с.
73. *Сотсков А. И., Шербаков В. Ю.* Теория и техника калмановской фильтрации при наличии мешающих параметров // Зарубежная радиоэлектроника.— 1985.— № 2.— С. 35—42.
74. *Статистические методы для ЭВМ.*— М. : Наука, 1986.— 459 с.
75. *Суходольский Г. В.* Структурно-алгоритмический анализ и синтез деятельности.— Л. : Изд-во Ленингр. ун-та, 1976.— 120 с.
76. *Фано Р. М.* Передача информации : (Статистическая теория связи).— М. : Мир, 1965.— 438 с.
77. *Физика за рубежом* // Сер. А. Исследования : Сб. ст.— М. : Мир, 1982.— 217 с.
78. *Фокин Ю. Г.* Надежность при эксплуатации технических средств.— М. : Воениздат, 1970.— 224 с.
79. *Френкель А. А.* Математический анализ производительности труда.— М. : Экономика, 1988.— 168 с.
80. *Хон Ф., Сеиу С., Ауфенкампер Д.* Теория сетей // Математика.— 1959.— Вып. 3.— 150 с.
81. *Цибулевский И. Е.* Человек как звено следящей системы.— М. : Наука, 1981.— 288 с.
82. *Цыпкин Я. З.* Теория линейных импульсных систем автоматического регулирования.— М. : Физматгиз, 1963.— 324 с.
83. *Шварц Г.* Выборный метод.— М. : Статистика. 1978.— 213 с.
84. *Шеннон К.* Работы по теории информации и кибернетики.— М. : Изд-во иностр. лит., 1963.— 829 с.

85. *Шеффе Х.* Дисперсионный анализ / Пер. с англ. В. А. Севастьянова, В. П. Чистякова.— М. : Наука, 1980.— 512 с.
86. *Ширяев А. Н.* Вероятность. Статистика. Случайные процессы.— М. : Наука, 1974.— 224 с.
87. *Эргономика.* Принципы и рекомендации.— М. : ВНИИТЭ.— 1971.— Вып. 2.— 210 с.
88. А. с. 1314906 СССР. Адаптивный фильтр / Н. Н. Козлов и др.— 1987.— 234 с.
89. А. с. 1056432 СССР. Регуляризованный фильтр Калмана / Н. Н. Козлов и др.— 1983, Бюл. № 43,— 12 с.

ОГЛАВЛЕНИЕ

| | |
|--|-----|
| Предисловие | 5 |
| ГЛАВА I | 7 |
| <hr/> | |
| МЕТОДЫ И АЛГОРИТМЫ ОБРАБОТКИ ИНФОРМАЦИИ В ИЗМЕРИТЕЛЬНО-ВЫЧИСЛИТЕЛЬНОМ КОМПЛЕКСЕ | |
| § 1. Специальное математическое обеспечение измерительно- вычислительного комплекса | 7 |
| § 2. Адаптивный метод определения параметров движения ЛА по результатам траекторных измерений | 18 |
| § 3. Алгоритм адаптивной дискретно-непрерывной фильтрации траекторных измерений | 37 |
| § 4. Некоторые вопросы коррекции систематических погреш- ностей | 60 |
| ГЛАВА II | 87 |
| <hr/> | |
| МЕТОДЫ АНАЛИЗА ДАННЫХ ПРИ ИССЛЕДОВАНИИ РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА. | |
| § 1. Вводные замечания | 87 |
| § 2. Этапы проведения анализа данных | 88 |
| § 3. Обработка данных, анализ результатов и принятие решений | 94 |
| ГЛАВА III | 118 |
| <hr/> | |
| ЭРГОНОМИЧЕСКИЙ АНАЛИЗ ДЕЯТЕЛЬНОСТИ ОПЕРАТОРА АВТОМАТИЗИРОВАННЫХ СИСТЕМ | |
| § 1. Задачи, решаемые автоматизированной системой сбора и обработки радиолокационной информации. Роль и место че- ловека-оператора в системе | 118 |
| § 2. Основные задачи и принципы эргономического анализа | 123 |
| § 3. Аналитические методы исследования деятельности оператора | 134 |
| § 4. Методы моделирования. Экспериментальные методы исследо- вания деятельности оператора | 173 |
| § 5. Эргономический анализ деятельности оператора АССОРЛИ | 189 |
| ГЛАВА IV | 219 |
| <hr/> | |
| КОМПЛЕКС ПРОГРАММ «СОКИ-Н» СТАТИСТИЧЕСКОЙ ОБРАБОТКИ РЕЗУЛЬТАТОВ КАЧЕСТВЕННЫХ ИЗМЕРЕНИЙ В НОМИНАЛЬНЫХ ШКАЛАХ | |
| § 1. Статистическая обработка выборки | 219 |
| § 2. Оценка различия двух независимых выборок | 243 |
| § 3. Оценка различия «т» независимых выборок | 268 |
| § 4. Оценка различия двух связанных выборок | 292 |
| § 5. Оценка различия «т» связанных выборок | 301 |
| § 6. Оценка корреляционной связи двух выборок | 311 |
| Список литературы | 361 |

Научное издание

*Белов Юрий Анатольевич
Егоров Борис Михайлович
Козлов Николай Николаевич
Ляшко Иван Иванович
Макаров Владимир Леонидович*

МАТЕМАТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ СЛОЖНОГО ЭКСПЕРИМЕНТА

ТОМ 5

**ПРОБЛЕМЫ ПОСТРОЕНИЯ
МАТЕМАТИЧЕСКОГО
И ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ
ИЗМЕРИТЕЛЬНО-ВЫЧИСЛИТЕЛЬНЫХ
КОМПЛЕКСОВ**

Художественный редактор И. П. АНТОНЮК
Технический редактор В. А. КРАСНОВА
Корректоры Л. Н. ЛЕМБАК, И. В. ТОЧАНЕНКО

ИБ № 8845

Сдано в набор 03.05.90. Подп. в печ. 06.09.90. Формат 60×90/16. Бум. офс.
№ 1. Лит. гарн. Выс. печ. Усл. печ. л. 23,0. Усл. кр.-отт. 23,56. Уч.-изд. л.
22,30. Тираж 2170 экз. Зак. № 0—1598. Цена 4 р. 80 к.

Издательство «Наукова думка». 252601 Киев, ул. Репина, 3.

Отпечатано с матриц Головного предприятия республиканского производ-
ственного объединения «Полиграфкнига». 252057 Киев, ул. Довженко, 3 в
Киевской книжной типографии научной книги, 252004 Киев 4, ул. Репи-
на, 4. Зак. 0-684.

