

Федеральное агентство по образованию
Государственное образовательное учреждение
высшего профессионального образования
Владимирский государственный университет

Е. Н. ТАЛИЦКИЙ

**МАТЕМАТИЧЕСКИЕ ОСНОВЫ
ПРОЕКТИРОВАНИЯ
ЭЛЕКТРОННЫХ СРЕДСТВ**

Курс лекций

Владимир 2007

УДК 621.396
ББК 32.844 – 021.1
Т16

Рецензенты:

Доктор технических наук, профессор
зав. кафедрой конструирования и производства
радиоэлектронной аппаратуры
Пензенского государственного университета

Н.К. Юрков

Доктор технических наук,
профессор Московского авиационного института
(государственного технического университета)

А.В. Назаров

Печатается по решению редакционного совета
Владимирского государственного университета

Талицкий, Е. Н.

Т16 Математические основы проектирования электронных средств : курс лекций / Е. Н. Талицкий ; Владим. гос. ун-т. – Владимир : Изд-во Владим. гос. ун-та, 2007. – 160 с. – ISBN 5-89368-732-9.

Рассматриваются вопросы системного анализа, расчета точности и надежности электронных средств, вопросы оптимизации, теории эксперимента и теории массового обслуживания.

Предназначен для студентов 2 – 3-го курсов специальностей 210201 – проектирование и технология радиоэлектронных средств, 210202 – проектирование и технология электронно-вычислительных средств очной и заочной форм обучения.

Табл. 23. Ил. 65. Библиогр.: 14 назв.

УДК 621.396
ББК 32.844 – 021.1

ISBN 5-89368-732-9

© Владимирский государственный
университет, 2007

Оглавление

Предисловие.....	5
Лекция 1. Основные понятия и определения	7
СИСТЕМНЫЕ МЕТОДЫ В ПРОЕКТИРОВАНИИ ЭЛЕКТРОННЫХ СРЕДСТВ.....	13
Лекция 2. Понятие о системе и системном анализе.....	13
Лекция 3. Оценка вариантов систем.....	20
Лекция 4. Моделирование систем.....	24
ОСНОВЫ ТЕОРИИ ТОЧНОСТИ ЭЛЕКТРОННЫХ СРЕДСТВ.....	35
Лекция 5. Погрешности и их математическое описание.....	35
Лекция 6. Уравнения погрешностей. Коэффициенты влияния.....	45
Лекция 7. Методы расчета точности.....	48
Лекция 8. Расчет размерных цепей	52
Лекция 9. Методы матричных испытаний и статистического моделирования	56
ОСНОВЫ ТЕОРИИ НАДЕЖНОСТИ ЭЛЕКТРОННЫХ СРЕДСТВ.....	66
Лекция 10. Введение	66
Лекция 11. Показатели надежности	69
Лекция 12. Показатели надежности ремонтируемых изделий	73
Лекция 13. Законы распределения времени безотказной работы ЭС	76
Лекция 14. Методы расчета надежности	80
Лекция 15. Методы расчета надежности при параллельном и смешанном соединениях элементов	83
МЕТОДЫ ОПТИМИЗАЦИИ ПРИ ПРОЕКТИРОВАНИИ ЭЛЕКТРОННОЙ АППАРАТУРЫ.....	87
Лекция 16. Классические методы оптимизации	87
Лекция 17. Методы регулярного поиска экстремума	93
Лекция 18. Оптимизация методом линейного программирования.....	107

МАТЕМАТИЧЕСКАЯ ТЕОРИЯ ЭКСПЕРИМЕНТА	112
Лекция 19. Введение	112
Лекция 20. Доверительный интервал для дисперсии нормально распределенной случайной величины.....	117
Лекция 21. Обработка результатов пассивного эксперимента	125
Лекция 22. Теория активного многофакторного эксперимента	130
Лекция 23. Проведение и обработка результатов эксперимента	138
Лекция 24. Крутое восхождение по поверхности отклика	143
Лекция 25. Планы второго порядка	147
Лекция 26. Введение в теорию массового обслуживания	151
Библиографический список.....	160

Предисловие [1]

Цель курса – изложение математических методов, используемых в исследованиях конструкций и технологии радиоэлектронных средств (РЭС) и электронно-вычислительных средств (ЭВС), проводимых, прежде всего, на этапах научно-исследовательских и опытно-конструкторских работ. Данный курс призван углубить и расширить фундаментальную подготовку инженеров-конструкторов-технологов, которые обязаны овладеть передовой технологией надежностного проектирования РЭС (ЭВС) с применением ЭВМ. Кроме того, изучение курса должно способствовать выполнению студентами и молодыми специалистами научных работ, а в будущем – и диссертаций в области конструирования, технологии и надежности РЭС (ЭВС).

Курс основывается прежде всего на знаниях по высшей и вычислительной математике, а также физике в объемах, даваемых в технических вузах. Полученные при изучении курса знания предназначены для того, чтобы в последующих конструкторских и технологических курсах использовать современные математические методы системного подхода, вероятностно-статистического анализа, математического моделирования и вычислительного эксперимента, теории надежности и оптимизации. Все это в сочетании с физическими представлениями и эвристическими (интуитивными) приемами, излагаемыми в конструкторских и технологических курсах, позволит студентам овладеть современными диалоговыми системами автоматизированного конструирования и технологической подготовки производства РЭС (ЭВС). Именно на такой основе возможно проведение всесторонних проектных исследований конструкций и технологии с целью получения РЭС (ЭВС) высокой надежности и конкурентоспособности.

Хотя курс носит теоретико-математический характер, являясь прикладной математикой для инженеров-конструкторов-технологов, его следует читать не на математических, а на конструкторско-технологических кафедрах вузов, что позволит глубже увязать его, особенно на практических занятиях, в лабораторных работах и при

курсовом проектировании, с задачами конструкторско-технологической деятельности и иллюстрировать примерами из этой области. Более того, наличие этого курса на конструкторско-технологических кафедрах позволит инициировать теоретико-математическую и функциональную переподготовку и совершенствование преподавательских кадров этих кафедр.

В результате изучения курса *студент должен*:

- познакомиться с сущностью, содержанием и применением системного подхода и анализа в проектировании и исследованиях РЭС (ЭВС);

- освоить основные идеи математического моделирования, а также методы синтеза и анализа конструкций и технологий РЭС (ЭВС) с использованием математических моделей и теории параметрической чувствительности;

- углубить знания в области теории вероятностей и математической статистики, включая инженерные методы решения задач точности, стабильности, серийнопригодности и надежности РЭС (ЭВС);

- овладеть теорией и методами активного эксперимента и обработки его результатов;

- изучить математические методы оптимизации технических решений;

- получить представление о перспективных теоретико-математических проблемах конструирования и технологии РЭС (ЭВС).

При подготовке курса лекций использованы некоторые материалы, написанные и опубликованные другими авторами, о чем есть указания по тексту лекций.

Автор выражает благодарность рецензентам: заведующему кафедрой КиПРа Пензенского государственного университета, доктору технических наук, профессору Н.К. Юркову и профессору Московского авиационного института (государственного технического университета), доктору технических наук А.В. Назарову за ценные замечания, способствующие улучшению настоящего издания.

Автор выражает признательность редактору Т.В. Климовой и аспиранту П.В. Кутровскому за большую работу по редактированию и верстке рукописи.

Лекция 1

ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ [1]

Развитие радиоэлектроники оказывает существенное влияние на научно-технический прогресс и успехи в социально-экономической области. Машиностроение, которое играет ключевую роль в ускорении научно-технического прогресса, нуждается в надежно работающих сложных изделиях радиоэлектроники, органически входящих в качестве комплектующих элементов в станки с числовым программным управлением, роботы и другие устройства гибких производственных систем. Транспорт, энергетика, космические системы и другие направления нашей экономики, где велики последствия одного единственного отказа, также нуждаются в высоконадежных изделиях радиоэлектроники. Надежность этих изделий закладывается при их проектировании, обеспечивается при производстве и поддерживается при эксплуатации в составе РЭС. Изделия радиоэлектроники должны рассматриваться как результат разработки схемы, конструкции и технологии.

Дадим основные определения рассматриваемых понятий и связанных с ними терминов. Исходными являются понятия радиоэлектроники и радиотехники.

Радиоэлектроника – это область науки и техники, охватывающая широкий круг вопросов использования электромагнитной энергии для передачи, приема и преобразования сигналов и содержащейся в них информации.

Радиотехника – область науки и техники, связанная с генерированием, преобразованием, излучением и приемом электромагнитных колебаний и волн радиочастотного диапазона. В соответствии с международным регламентом границы радиочастот лежат в пределах от 3 кГц до 3 ГГц, которым соответствуют длины волн от 100 км до 0,1 мм. Таким образом, радиотехника является частью радиоэлектроники. Кроме того, к радиоэлектронике относят технику инфранизких частот, инфракрасную технику и оптоэлектронику. С технической стороны радиоэлектроника также охватывает автоматику, техническую кибернетику, вычислительную технику и электронику.

Как отрасль промышленности радиотехника связана с разработкой, производством и применением радиоаппаратуры для радиосвязи,

радиовещания, телевидения, радиолокации, радионавигации и других направлений техники передачи информации. В более широком смысле вместо радиоаппаратуры применяется термин *радиоэлектронная аппаратура* (РЭА), под которым подразумеваются приборы и технические устройства, предназначенные для преобразования электрических и электромагнитных сигналов с определенными целями.

Наиболее общим названием изделий радиоэлектроники, рекомендованным в стандартах, является термин *радиоэлектронное средство*, понимаемый как техническое изделие определенной сложности или его составная часть, в основу действия которого положены принципы радиотехники и электроники. По функциональной сложности РЭС разделяют на уровни: система, комплекс, устройство, узел. Собирательное понятие РЭА объединяет последние два уровня РЭС.

Радиоэлектронная система (РЭ система) – это РЭС в виде совокупности функционально взаимодействующих автономных радиоэлектронных комплексов и устройств, которые образуют целостное единство, обладающее свойством перестроения структуры в целях рационального выбора и использования входящих средств нижних уровней при решении технических задач. Примером РЭ системы может служить аэродромная автоматизированная система управления воздушным движением, которая производит измерения траекторных параметров самолетов и управление ими при полетах по кругу в районе аэродрома. В ее состав входят бортовой радиоэлектронный комплекс и наземные радиолокационный и вычислительный комплексы, которые, в свою очередь, состоят из радиоэлектронных устройств (РЭУ), т. е. РЭС более низкого уровня. Перестроение структуры данной РЭ системы вызывается условиями воздушной обстановки в районе аэродрома, в зависимости от интенсивности воздушного движения назначается оптимальная (с точки зрения безопасности) совокупность работающих РЭС.

В зависимости от сложности разрешаемых технических задач РЭ система может быть частью другой РЭ системы более высокого уровня.

Разновидностью РЭ системы является *радиотехническая система*, под которой понимают совокупность РЭС для передачи сообщений и команд по радиоканалам, состоящим из радиопередатчиков, линий радиосвязи и радиоприемников. Другой важной разновидностью РЭ системы служит *вычислительная система*, определяемая как взаимосвязанная совокупность вычислительных комплексов и уст-

ройств, согласованных по быстродействию, пропускной способности и другим параметрам. Отличительный признак как радиотехнической, так и вычислительной системы – возможность перестроения ее структуры с целью наиболее эффективного использования входящих в систему технических средств.

В составе любой РЭ системы могут присутствовать механические, электромеханические и другие средства, без которых невозможна ее эксплуатация. Эти средства могут входить в РЭ систему как отдельно, так и находиться внутри радиоэлектронных комплексов или устройств.

Радиоэлектронный комплекс (РЭК) – это РЭС в виде совокупности функционально связанных РЭУ, обладающих свойством изменения структуры в целях сохранения работоспособности при выполнении технических задач в составе РЭ системы или в режиме самостоятельного применения. Примером РЭК может служить радиолокационный комплекс, состоящий из дальномера, высотомера и устройства опознавания и предназначенный для обнаружения объекта, определения его координат и установления принадлежности. При выходе из строя одного из устройств, входящих в комплекс, изменяются связи между ними, тем самым обеспечивается решение задачи при пониженной точности определения параметров объекта.

В качестве второго примера РЭК рассмотрим вычислительный комплекс, состоящий из двух или большего количества ЭВМ с собственными устройствами ввода-вывода, коммутаторами и дополнительными устройствами для информационного обмена между ЭВМ. Техническая задача такого комплекса – обработка информации. При выходе из строя одной ЭВМ ее функцию берет на себя другая ЭВМ. Работоспособность комплекса при этом сохраняется, но снижается его пропускная способность либо ограничивается круг решаемых задач.

Радиоэлектронное устройство (РЭУ) – это РЭС в виде функционально законченной сборочной единицы, которая выполнена на несущей конструкции и реализует функции приема, преобразования и передачи информации или решает техническую задачу на их основе. В зависимости от сложности решаемых задач РЭУ может быть частью другого РЭУ. Примером РЭУ является метеонавигационный бортовой радиолокатор, который может быть выполнен на несущей конструкции в виде корпуса шкафа или блока. Данный радиолокатор предназначен для выполнения технической задачи по обнаружению с борта летательного аппарата опасных для полета гидрометеорологи-

ческих образований и определению их углового положения и удаленности. Он состоит из совокупности функционально законченных сборочных единиц: антенного передающего и приемного устройства, а также устройств преобразования и отображения информации, каждое из которых также может быть отдельным примером РЭУ.

Еще одной иллюстрацией РЭУ может служить курсовой радиомаяк, предназначенный для изучения радиосигналов, содержащих информацию для управления самолетом по азимуту при заходе на посадку и во время посадки. Функционально законченными сборочными единицами, входящими в радиомаяк, являются антенное и передающее устройства. Обычно радиомаяк конструируется в виде шкафа или блока.

Каждый из рассмотренных РЭУ может эксплуатироваться самостоятельно или в составе РЭК (РЭ системы). Любое РЭУ структурно разделяется на радиоэлектронные функциональные узлы, которые являются РЭС первого уровня, состоящими, в свою очередь, из элементов нулевого уровня функциональной иерархии: радиоэлементов (РЭ) и электротехнических элементов (ЭЭ). Некоторые РЭ и ЭЭ могут входить непосредственно в РЭУ.

Радиоэлектронный функциональный узел (РЭФУ) – это РЭС в виде законченной сборочной единицы, которая выполнена на несущей конструкции, реализует функции преобразования сигнала и не имеет самостоятельного эксплуатационного значения.

РЭФУ, например, являются усилитель, модулятор, источник вторичного электропитания и др. Первый узел выполняет функции преобразования тока или напряжения по амплитуде без изменения их формы, второй – преобразования амплитуды либо частоты сигналов в соответствии с изменениями модулирующего напряжения, а третий – преобразования переменных токов и напряжения в постоянные или постоянных токов и напряжения одного значения в постоянные другого значения. Самостоятельного применения эти РЭФУ не имеют, а могут эксплуатироваться в составе РЭУ. В зависимости от технических характеристик указанные РЭФУ могут иметь конструктивное исполнение в виде радиоэлектронного шкафа, блока или ячейки.

По конструктивной сложности РЭС имеют четыре уровня иерархического деления. Нулевой уровень составляют изделия электронной техники (ИЭТ) и электротехнические изделия (ЭТИ). Они входят в радиоэлектронные ячейки или кассеты, выполненные на основе несущей конструкции первого уровня (НК 1). На несущих кон-

струкциях второго уровня (НК 2) строятся радиоэлектронные блоки или рамы, которые могут включать в себя как ячейки (кассеты), так и ИЭТ с ЭТИ. Наконец, на несущих конструкциях третьего уровня (НК 3) разрабатываются радиоэлектронные шкафы, пультаы или стойки, в которые входят конструкции как второго, так и первого уровня.

Наличие двух видов классификации уровней разукрупнения свидетельствует о том, что в любом РЭС как готовом изделии различают две непосредственно взаимосвязанные части: электрическую схему и конструкцию.

Электрическая схема РЭС – это совокупность РЭ, соединенных между собой в соответствии с заданным принципом действия РЭ системы или в соответствии с функциями, выполняемыми данным РЭС внутри другого РЭС более высокого уровня. Иными словами, электрическая схема отображает те преобразования полезных сигналов и помех, которые происходят с ними при прохождении через РЭС.

Схемотехника – прикладная научная дисциплина, занимающаяся анализом и синтезом электрических схем РЭС. Теоретическим фундаментом схемотехники является радиотехника.

Конструкция РЭС – это совокупность ИЭТ, ЭТИ и конструктивных деталей, находящихся в определенной пространственно-механической, информационной и энергетической взаимосвязи, которая обеспечивает выполнение данным РЭС необходимых функций с высокой надежностью.

Надежность – это свойство РЭС сохранять во времени в установленных пределах значения всех требуемых характеристик, параметров и показателей при заданных режимах и условиях применения, технического обслуживания, ремонтов, хранения и транспортировки.

Различают два понимания термина «конструирование». *Конструирование РЭС как вид инженерной деятельности* есть совокупность работ, выполняемых при создании проекта конструкции аппаратуры. *Конструирование РЭС как прикладная научная дисциплина* – это обобщение методов анализа и синтеза конструкций. Теоретический фундамент конструирования РЭС составляют: системный анализ, математическое моделирование, исследование операций, планирование эксперимента, теория вероятностей и математическая статистика, теория надежности, теория оптимизации и другие научные направления. В настоящее время на основе этих фундаментальных направлений формируется самостоятельная *теория конструирования РЭС* как система научных принципов и общих закономерностей раз-

работки конструкций в радиоэлектронике. Применение ЭВМ как главной технической базы разработки конструкций, усложнение практических задач создания конструкций и усиление взаимосвязи конструирования с системотехникой, схемотехникой, технологией и эксплуатацией РЭС – всё это требует ускоренного развития теории конструирования РЭС.

Технология РЭС как область техники – совокупность способов обработки, изменения состояния, свойств и формы материалов, изготовления полуфабрикатов (деталей, узлов, устройств) и их сборки с определенной точностью и производительностью в процессе производства РЭС. *Технология РЭС как прикладная научная дисциплина* – это обобщенное представление физических, химических, механических и других закономерностей и методов воздействия на материалы соответствующими орудиями производства.

Эксплуатация РЭС как область техники – совокупность способов транспортировки, хранения, установки и технического обслуживания РЭС при их функциональном использовании по назначению. В то же время *эксплуатация РЭС как прикладная научная дисциплина* – это системное представление условий наилучшего использования аппаратуры для заданных целей с высокой надежностью и эффективностью.

СИСТЕМНЫЕ МЕТОДЫ В ПРОЕКТИРОВАНИИ ЭЛЕКТРОННЫХ СРЕДСТВ

Лекция 2

ПОНЯТИЕ О СИСТЕМЕ И СИСТЕМНОМ АНАЛИЗЕ

Конструирование и технология производства электронных средств как система

Совершенствование конструкций электронных средств (ЭС) представляет собой одно из направлений их развития. Правильное понимание задач и перспектив конструирования ЭС возможно только при комплексном рассмотрении вопросов, то есть при учёте взаимосвязи между конструкцией, системой, схемой, технологией, производством, эксплуатацией. Поэтому конструкцию ЭС (а следовательно, и технологию его изготовления) нельзя рассматривать изолированно от ЭС, главное назначение которых состоит в основном в выдаче, обработке и запоминании полезной информации.

Увеличивающаяся сложность ЭС, повышение точности, быстродействия технологического оборудования, а также расширяющиеся специализация и кооперирование предприятий ставят перед конструкторами новые научно-технические проблемы. Среди них существенное место занимают вопросы, связанные с более детальным и глубоким изучением свойств и специфики функционирования отдельных видов ЭС и оборудования, особенно если речь идёт о средствах, работающих на новых малоисследованных принципах.

Однако при разработке крупных комплексов возникают проблемы, относящиеся не только к свойствам отдельных средств, но также и к закономерностям функционирования соответствующей системы в целом.

Более того, появился широкий круг проблем, мало связанных с рассмотрением свойств отдельных элементов: определение общей структуры системы, организация взаимодействия между её элементами, учёт влияния внешней среды, определение оптимальных режимов функционирования и т. д. По мере увеличения сложности систем комплексным, «общесистемным», вопросам отводится всё более значительное место.

Они составляют содержание нового направления инженерной мысли, получившего наименование *системотехники*.

Научной, главным образом математической, базой системотехники является *теория сложных (больших) систем*.

В настоящее время нет возможности дать общее определение сложной системы, обладающее достаточной чёткостью с точки зрения построения строгой теории, поскольку отнесение той или другой реальной системы к разряду «сложных» или «простых» является в значительной мере условным и связано в основном с тем, насколько существенную роль играют при изучении системы комплексные, «общесистемные», вопросы.

Рассмотрение изучаемого объекта материального мира как системы, состоящей из взаимодействующих элементов, построение математической модели для него и исследование его свойств методом моделирования составляет сущность *системного подхода*, а совокупность методов и приёмов исследования входит в арсенал самостоятельного научного направления, получившего название *системный анализ*. Системный подход играет решающую роль в исследовании операций, системотехнике, а также в сфере проектирования и построения РЭС, АСУ и др.

Основные понятия и методология системного анализа

Системный анализ – это методология решения крупных задач и проблем, основанная на концепции систем.

Методология – это логически и процедурно организованная последовательность операций.

Система – это целостное множество взаимосвязанных элементов.

Системы подразделяются на простые и сложные.

Простые обычно состоят из одного элемента (рис. 2.1).

Сложные системы (рис. 2.2) должны удовлетворять ряду требований.

1. Система может быть разделена на подсистемы, подсистема – на более мелкие подсистемы (элементы).

2. Элементы функционируют не изолированно друг от друга, а во взаимодействии.

3. Свойства системы зависят не только от свойств элементов, но и от характера их взаимодействия.

Системы подразделяют также на детерминированные, вероятностные, игровые. *Детерминированной* называется система, состояние которой в будущем однозначно определяется ее состоянием в настоящий момент времени и законами, описывающими переходы элементов и системы из одних состояний в другие. *Вероятностные*, или *стохастические*, системы – это системы, поведение которых определяется законами теории вероятностей. *Игровой* называется система, осуществляющая разумный выбор своего поведения в будущем. Ситуация сдачи продукции изготовителем заказчику может рассматриваться как игровая система.

Системы подразделяются на открытые и закрытые. *Открытые* – имеющие связь с внешней средой, *закрытые* – не имеющие такой связи.

Системы могут быть хорошо организованными, плохо организованными и самоорганизующимися.

Хорошо организованная система – это система, у которой определены все элементы, их взаимосвязь, правила объединения в более крупные компоненты, связи между всеми компонентами и целями системы, ради которых она создается.

Очевидно, что *плохо организованная* система не обладает хотя бы частью этих свойств.

Самоорганизующаяся система обладает свойством адаптации к изменению условий внешней среды.

Системы могут быть линейными и нелинейными, стационарными и нестационарными, инерционными и безынерционными и т.д.

Как установить, входит элемент в систему или нет? Это зависит от того, участвует элемент в процессе, приводящем к появлению выходного результата, или нет.

Каждая система состоит из системных объектов, свойств и связей.

Системные объекты: вход, процесс, выход, обратная связь, ограничения.

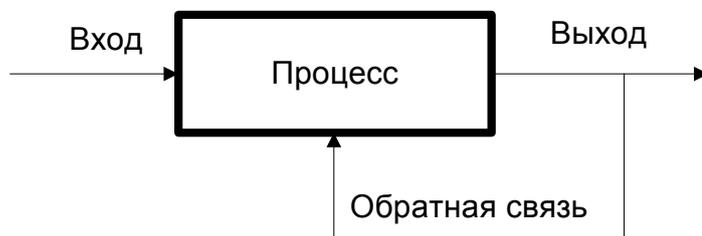


Рис. 2.1

Входом процесса называется то, что изменяется при протекании этого процесса.

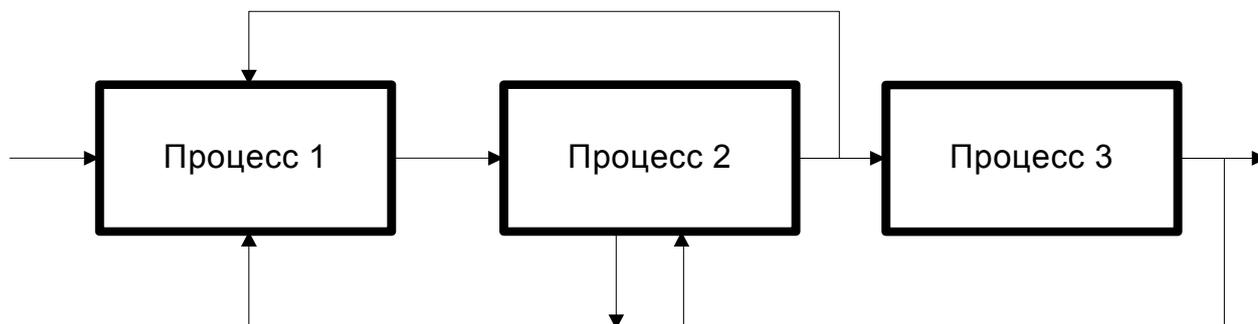


Рис. 2.2

Выходом называется результат или конечное состояние процесса.

Процесс – это то, что переводит вход в выход.

Обратная связь служит для сравнения реального выхода с моделью выхода, выделяет различие и оценивает его, вырабатывает решение, связанное с различием, формирует процесс ввода решения и воздействует на процесс с целью сближения выхода и модели выхода.

Ограничения обеспечивают соответствие выхода системы целям потребителя.

Свойством процесса называется способность системы переводить данный вход в заданный выход.

Связь – это то, что определяет следование процесса, то есть вход некоторого процесса является выходом другого процесса, выход другого – входом следующего и т. д.

Выход является необходимым, если его отсутствие создает угрозу существованию или развитию системы.

Проблемой называется ситуация, характеризующаяся различием между необходимым и существующим выходами.

Решение проблемы – это система, которая устраняет промежуток между существующей и желаемой системами. Система, заполняющая промежуток, является объектом конструирования.

Системный анализ указывает необходимую последовательность операций при решении проблем, которая заключается в следующем:

- 1) обнаружение проблемы;
- 2) оценка ее актуальности;
- 3) определение цели и критериев;
- 4) «вскрытие» структуры существующей системы;
- 5) определение «дефектных» элементов существующей системы, ограничивающих получение заданного выхода;

- 6) оценка «веса» дефектных элементов;
- 7) построение набора альтернатив (вариантов) и их оценка по выбранным критериям;
- 8) выбор варианта для реализации, определение процесса реализации и реализация выбранного решения;
- 9) оценка результатов реализованного решения.

Системный анализ претендует на роль некоторого «каркаса», объединяющего все необходимые методы, знания и действия для решения проблемы.

При системном анализе используются следующие математические методы:

- методы математической статистики;
- методы теории эксперимента;
- методы исследования операции;
- методы теории надежности и теории точности;
- методы распознавания образов;
- методы теории игр;
- методы инженерной психологии, а также другие методы.

Проектирование ЭС как сложная система

Структурная схема проектирования новых изделий (рис. 2.3) показывает, что процесс проектирования ЭВС может рассматриваться как сложная система.

Техническое предложение – совокупность конструкторских документов, которые должны содержать технические и технико-экономические обоснования целесообразности разработки документации изделия на основании анализа технического задания заказчика и различных вариантов возможных решений изделий, сравнительной оценки решений с учетом конструктивных и эксплуатационных особенностей разрабатываемого и существующих изделий и патентных исследований.

Техническое предложение после согласования и утверждения в установленном порядке является основанием для разработки эскизного (технического) проекта.

Эскизный проект – совокупность конструкторских документов, которые должны содержать принципиальные конструктивные решения, дающие общее представление об устройстве и принципе работы изделия, а также данные, определяющие назначение, основные параметры и габаритные размеры разрабатываемого изделия.

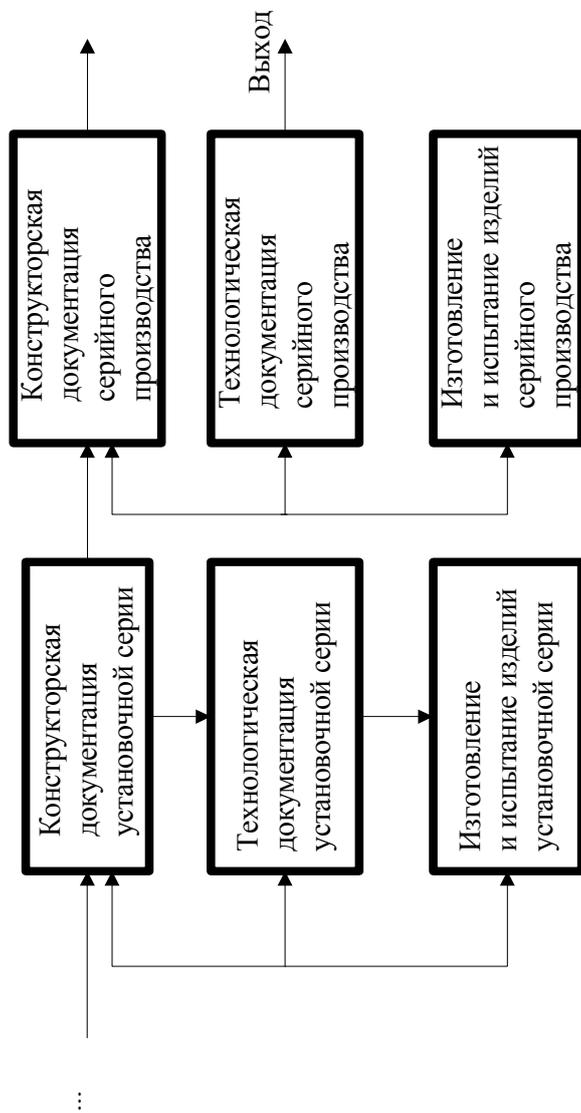
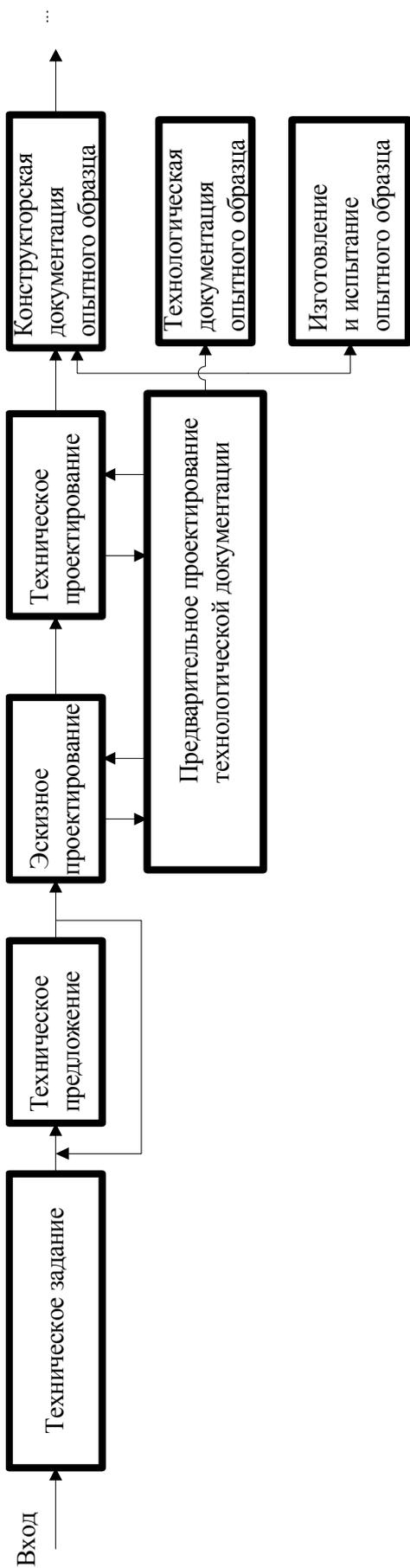


Рис. 2.3

Эскизный проект после согласования и утверждения в установленном порядке служит основанием для разработки технического проекта или рабочей конструкторской документации.

Технический проект – совокупность конструкторских документов, которые должны содержать окончательные технические решения, дающие полное представление об устройстве разрабатываемого изделия, и исходные данные для разработки рабочей документации.

Технический проект после согласования и утверждения в установленном порядке служит основанием для разработки рабочей конструкторской документации.

Другим примером сложной системы может быть структурная схема ЭВМ, показанная на рис. 2.4, где обозначено: $У_{ВВ}$ – устройство ввода информации; ВЗУ – внешнее запоминающее устройство; АЛУ – арифметико-логическое устройство; ОЗУ – оперативное запоминающее устройство; УУ – устройство управления; ПУ – пульт управления; $У_{ВЫВ}$ – устройство вывода информации.

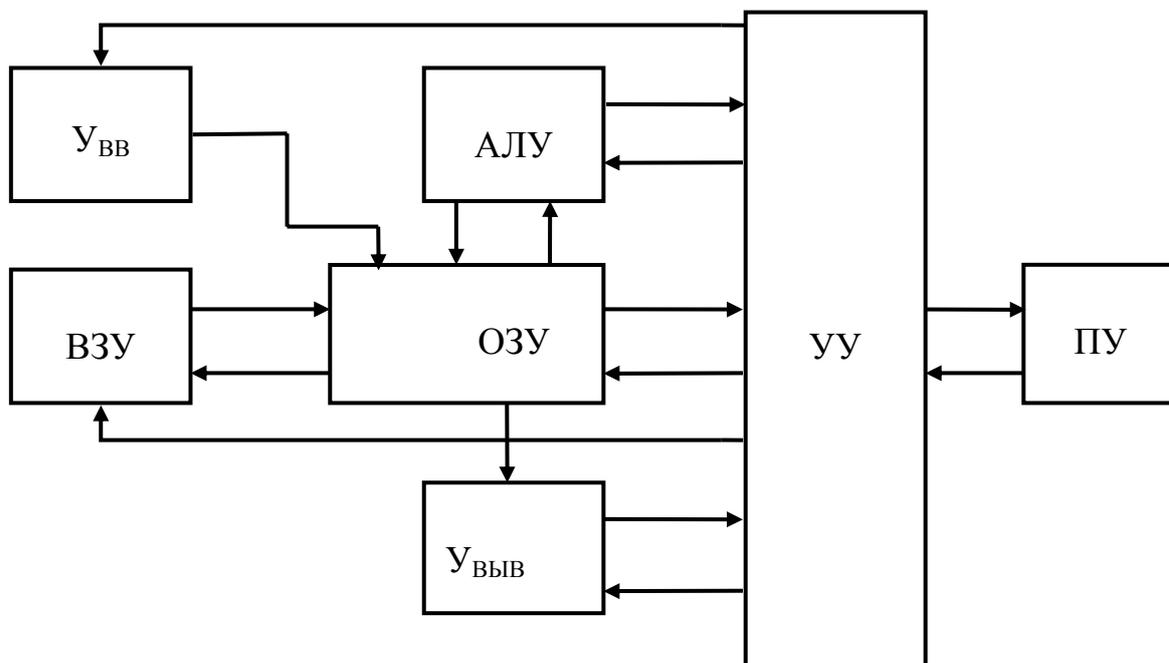


Рис. 2.4

Лекция 3

ОЦЕНКА ВАРИАНТОВ СИСТЕМ

Одна из задач при системном анализе – выбор альтернатив (вариантов) систем по определенным критериям. Из всех вариантов оптимальным является тот, который обладает наилучшим значением критерия, называемым критерием оптимальности. Следовательно, для выбора оптимальной системы необходимо выбрать этот критерий, который, как правило, характеризует качество системы.

Понятие о качестве системы. Показатели качества

В соответствии с ГОСТ Р 50779.11 – 2000 «Статистическое управление качеством. Термины и определения» *качество продукции* – это совокупность свойств и признаков продукции или услуги, которые влияют на их способность удовлетворять установленные или предполагаемые потребности.

Для количественной характеристики качества введены показатели качества. Государственный стандарт предусматривает десять групп показателей качества.

1. *Показатели назначения*, например: мощность, чувствительность, быстродействие, КПД и т.д.

2. *Показатели надежности*. Делятся на четыре группы: показатели безотказности, долговечности, ремонтпригодности и сохраняемости, например: вероятность безотказной работы, интенсивность отказов, средняя наработка на отказ, среднее время восстановления и т. д.

3. *Показатели технологичности*. Делятся на основные, технико-экономические и технические. К основным показателям относятся трудоемкость изготовления и себестоимость изделия, к технико-экономическим – относительная трудоемкость процесса изготовления изделия и другие, к техническим – коэффициент шероховатости поверхности, коэффициент использования материала, коэффициент точности обработки и др.

4. *Эргономические показатели*. Эта обширная группа делится на гигиенические, антропологические, психологические, физиологи-

ческие и психофизические показатели. *Гигиенические* показатели оценивают соответствие изделия гигиеническим условиям жизнедеятельности и работоспособности человека при взаимодействии его с изделием, например: уровень освещенности, температура, влажность, шум и т.д. *Антропологические* показатели оценивают соответствие изделия размерам и форме человеческого тела. *Физиологические* и *психофизические* показатели оценивают соответствие изделия физиологическим свойствам человека и особенностям функционирования его органов чувств, то есть оценивают скоростные, зрительные, слуховые и другие возможности. *Психологические* показатели оценивают соответствие изделия психологическим особенностям человека, то есть оценивают возможности восприятия и переработки информации человеком. В отличие от трех предыдущих групп эргономические показатели не могут быть найдены расчетным путем и определяются обычно экспертами по балльной системе.

5. *Эстетические показатели.* Оценивают цвет, форму изделия, например: показатели цветового колорита, показатель соответствия моде, показатель стилевого соответствия и др.

6. *Патентно-правовые показатели.* Характеризуют количество и весомость изобретений и патентов, используемых в изделии, например: показатель патентной чистоты, показатель патентной защиты и др.

7. *Показатели стандартизации и унификации.* Характеризуют насыщенность продукции стандартными, унифицированными и оригинальными элементами, например: коэффициент стандартизации и унификации, коэффициент повторяемости и др.

8. *Показатели транспортабельности.* Характеризуют приспособленность продукции к перемещению в пространстве, не сопровождающемуся ее использованием или потреблением, например: средняя трудоемкость подготовки продукции к транспортировке, коэффициент использования объема средства транспортировки и др.

9. *Экологические показатели.* Характеризуют уровень вредных воздействий на окружающую среду, возникающих при эксплуатации или потреблении продукции, например: содержание вредных примесей, выбрасываемых в окружающую среду, и др.

10. *Показатели безопасности,* например: вероятность безотказной работы, электрическая прочность изоляции и др.

Оценка уровня качества систем

Применяется несколько методов оценки уровня качества системы, но наибольшее распространение получили дифференциальный и комплексный методы.

Дифференциальный метод является наиболее простым. Для этого определяются отдельные (единичные) показатели качества и сравниваются с базовыми показателями, регламентированными стандартами или техническими условиями. Эти показатели обычно записываются в абсолютной или относительной форме:

$$K_{yi} = \frac{K_i}{K_{\text{б}i}},$$

где K_{yi} – коэффициент уровня по i -му показателю качества; K_i – i -й показатель качества; $K_{\text{б}i}$ – базовое значение i -го показателя качества.

Если показатель характеризует положительное свойство изделия, то желательно чтобы $K_y > 1$, если показатель характеризует отрицательное свойство изделия, то, наоборот, $K_y < 1$. Например, необходимо стремиться, чтобы значение показателя стоимости изделия было меньше базового показателя. Недостаток этого метода в том, что одни показатели могут быть лучше базового значения или таких же показателей других изделий, а другие – хуже, что часто затрудняет выбор наилучшего варианта. Для устранения этого противоречия из всей номенклатуры показателей выделяется один (или несколько), основной, а другие показатели переводятся в разряд ограничений.

Комплексный метод характеризует качество системы одним показателем, учитывающим значения других (единичных) показателей качества, то есть проводится «свертка» единичных показателей в один комплексный показатель качества. Наибольшее применение нашел показатель в виде суммы

$$K = g_1 K_{y1} + g_2 K_{y2} + \dots + g_n K_{yn},$$

где K_{y1} , K_{y2} – единичные показатели качества; g_1 , g_2 – коэффициенты «веса» показателей, определяемые экспертами. Сумма всех коэффициентов «веса» должна равняться единице.

Пример оценки уровня качества. Пусть известны параметры шести судовых радиолокационных станций (РЛС), которые приведены в таблице.

Показатель \ Номер станции	1	2	3	4	5	6
Дальность действия (D), км	45	40	45	50	35	60
Стоимость (C), млн руб.	80	90	120	90	100	200
Масса, $m \cdot 10^3$ кг	2,5	4	2,5	3	2	4

Необходимо выбрать наилучший вариант РЛС при следующих ограничениях: $D_{\min} = 40$ км, $C_{\max} = 100$ млн руб., $m_{\max} = 3000$ кг.

Видно, что этим ограничениям удовлетворяют только первая и четвертая станции. Сравнение их дифференциальным методом по трем указанным параметрам невозможно. Поэтому применим интегральный метод, принимая коэффициенты «веса» равными $g_D = 0,5$, $g_C = 0,25$, $g_m = 0,25$. В качестве «базового» значения выберем среднее по двум сравниваемым вариантам: $D_{\text{ср}} = 47,5$ км, $C_{\text{ср}} = 85$ млн руб., $m_{\text{ср}} = 2750$ кг.

$$\begin{aligned}
 K_{\Sigma 1} &= K_{D1}g_D + K_{C1}g_C + K_{m1}g_m = \\
 &= \frac{45}{47,5} \cdot 0,5 + \frac{85}{80} \cdot 0,25 + \frac{2,75}{2,5} \cdot 0,25 = 1,015.
 \end{aligned}$$

$$K_{\Sigma 4} = \frac{50}{47,5} \cdot 0,5 + \frac{85}{90} \cdot 0,25 + \frac{2,75}{3} \cdot 0,25 = 0,99.$$

Видно, что предпочтение следует отдать первой станции.

Лекция 4

МОДЕЛИРОВАНИЕ СИСТЕМ [8]

Общие сведения о моделировании

Моделирование представляет собой метод научного познания, при котором исследуемый объект замещается другим, более простым, называемым моделью, изучение которого даёт возможность получить новую и ранее неизвестную информацию об исходном объекте. В зависимости от способа воплощения изучаемого объекта в модели различают физическое и математическое моделирование. Оба эти вида используют при проектировании ЭС.

Процесс создания ЭС всегда включает в себя прогнозирование его основных характеристик. Наиболее естественным и, как правило, достоверным методом такого предсказания является макетирование, которое предполагает создание устройства и последующие его испытания. В процессе проведения последних выявляются ошибки и слабые места конструкции, которые далее устраняются. Такой способ отработки изделий называют *физическим моделированием*, поскольку он осуществляется на вещественной (материальной) модели, которая, отображая изучаемый объект, сохраняет его физическую природу. Для удешевления экспериментов физические модели (макеты) изготавливают зачастую в упрощённом виде (например меньшей мощности, производительности и пр.). Естественно, что при этом точность предсказания снижается.

Часто натурный эксперимент практически невыполним, настолько он сложен, дорог и рискован. В ряде случаев создание макета устройства экономически нецелесообразно, особенно когда после изготовления в него практически невозможно внести какие-либо усовершенствования (изменения). Так, например, обстоит дело при создании новых типов больших интегральных схем (БИС). Поэтому естественно стремление разработчиков создавать сразу конструкции ЭС с характеристиками, близкими к оптимальным.

Решение перечисленных задач требует применения методов, с помощью которых можно было бы провести оценку (расчёт) наиболее важных процессов, имеющих место в проектируемом изделии. Это достигается *математическим моделированием*, суть которого в том, что проектируемому изделию ставят в соответствие его математиче-

скую модель (ММ). Под последней обычно понимают упрощённое отображение наиболее существенных свойств реального объекта, выраженное в математической форме.

Математическая модель может быть в общем случае задана системой уравнений, неравенств, логической последовательностью операций и пр. Независимо от способа задания ММ она всегда дополняется ограничениями, связанными, как правило, с техническими возможностями устройства, условиями его эксплуатации и т. п.

Наиболее распространённый способ представления ММ – система (каких-либо) уравнений с необходимыми для её решения данными: начальными и граничными условиями, численными значениями коэффициентов в уравнениях и пр. Варьируя ими, можно провести детальное изучение физических процессов в изделии, отображённых выбранной ММ, выявить основные их закономерности, оценить влияние на них различных факторов, т. е. получить информацию, аналогичную получаемой в ходе физического эксперимента. Однако в данном случае вместо экспериментальной установки мы используем ЭВМ, а вместо изучаемого физического объекта – его математическую модель. Подобные исследования называют *вычислительным экспериментом*.

Вычислительный эксперимент обладает рядом преимуществ по сравнению с натурным. Он, как правило, намного дешевле, легче и быстрее реализуется, допускается (с корректировкой) на любой стадии исследования и позволяет моделировать условия эксперимента, которые зачастую вообще невозможно воспроизвести при натуральных испытаниях.

Математическое моделирование включает в себя математические модели объектов проектирования, методы и алгоритмы выполнения проектных процедур.

Требования к математическим моделям

К математическим моделям предъявляются требования универсальности, адекватности, точности и экономичности.

Степень универсальности ММ характеризует полноту отображения в модели свойств реального объекта. Математическая модель отображает лишь некоторые свойства объекта. Так, большинство ММ, используемых при функциональном проектировании, предназначено для отображения протекающих в объекте физических или

информационных процессов, при этом не требуется, чтобы ММ описывала такие свойства объекта, как геометрическая форма составляющих его элементов. Например, ММ резистора в виде уравнения закона Ома характеризует его свойство пропускать электрический ток, но не отображает габариты резистора как детали, его цвет, механическую прочность, стоимость и т. п.

Точность ММ оценивается степенью совпадения значений параметров реального объекта и значений тех же параметров, рассчитанных с помощью оцениваемой ММ. Пусть отражаемые в ММ свойства оцениваются вектором выходных параметров $y = (y_1, y_2, \dots, y_m)$. Тогда, обозначив истинное и рассчитанное с помощью ММ значения j -го выходного параметра через $y_{j\text{ист}}$ и $y_{j\text{м}}$ соответственно, определим относительную погрешность Δ_j расчёта параметра y_j как

$$\Delta_j = (y_{j\text{м}} - y_{j\text{ист}}) / y_{j\text{ист}}.$$

Адекватность ММ – способность отображать заданные свойства объекта с погрешностью не выше заданной. При этом, как правило, адекватность модели имеет место лишь в ограниченной области изменения внешних переменных – области адекватности (ОА) математической модели.

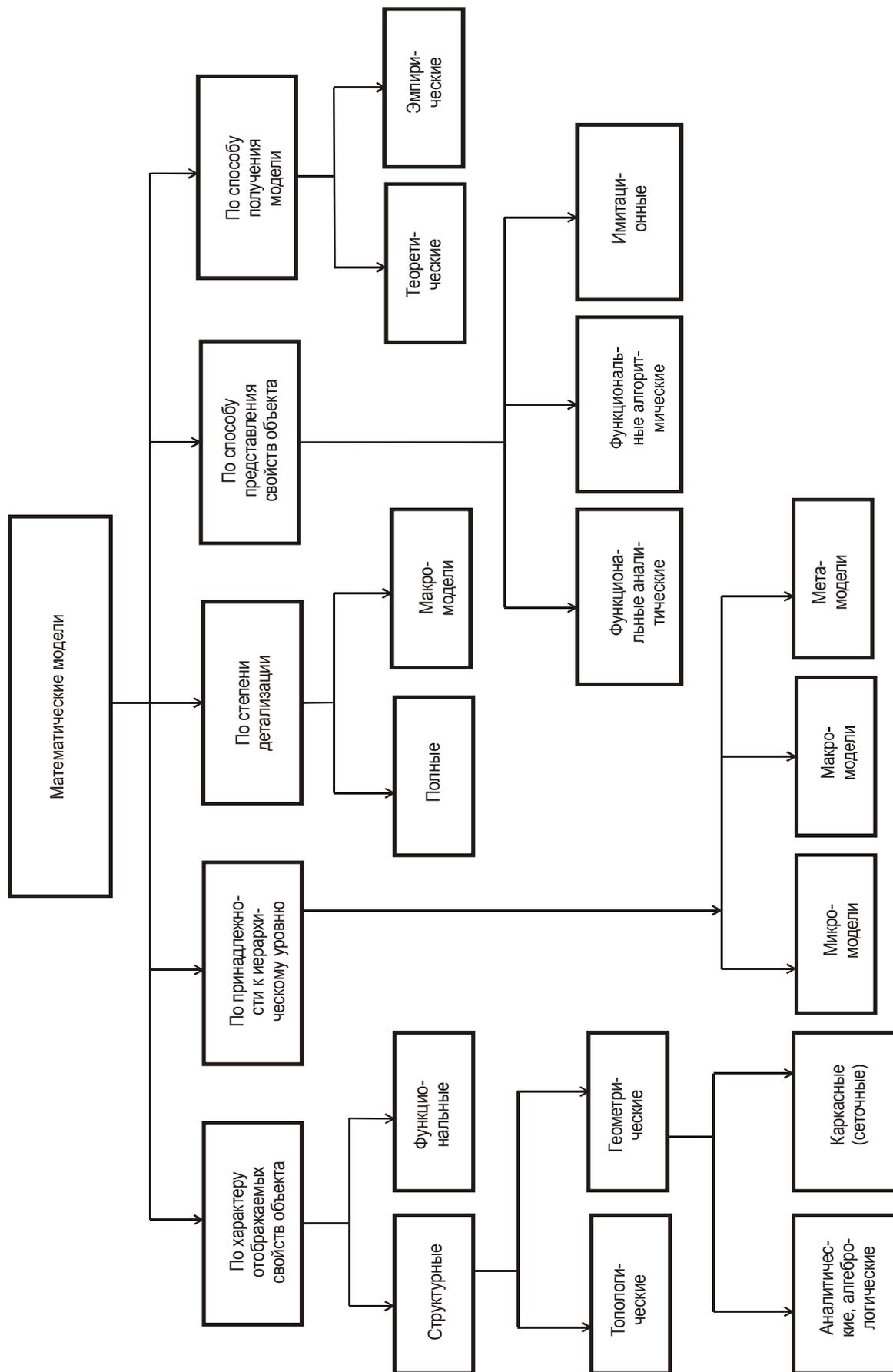
Экономичность ММ характеризуется затратами вычислительных ресурсов (машинного времени $T_{\text{м}}$ и памяти $\Pi_{\text{м}}$) на её реализацию. Чем меньше $T_{\text{м}}$ и $\Pi_{\text{м}}$, тем модель экономичнее. Вместо значений $T_{\text{м}}$ и $\Pi_{\text{м}}$, зависящих не только от свойств модели, но и от особенностей применяемой ЭВМ, часто используют другие величины, например: среднее количество операций, выполняемых при одном обращении к модели, размерность системы уравнений, количество используемых в модели внутренних параметров и т. п.

Необходимо помнить, что требования высокой точности и степени универсальности, широкой области адекватности, с одной стороны, и высокой экономичности, с другой стороны, противоречивы.

Классификация математических моделей

Основные признаки классификации и типы ММ приведены на рисунке.

По характеру отображаемых свойств объекта ММ делятся на структурные и функциональные.



Структурные ММ предназначены для отображения структурных свойств объекта. Различают структурные ММ топологические и геометрические.

В *топологических* ММ отображаются состав и взаимосвязи элементов объекта. Их чаще всего применяют для описания объектов, состоящих из большого числа элементов, при решении задач привязки конструктивных элементов к определённым пространственным позициям (например, задачи компоновки оборудования, размещения деталей, трассировки соединений) или к относительным моментам времени (например, при разработке расписаний, технологических процессов). Топологические модели могут иметь форму графов, таблиц (матриц), списков и т. п.

В *геометрических* ММ отображаются геометрические свойства объектов, в них дополнительно к сведениям о взаимном расположении элементов содержатся сведения о форме деталей. Геометрические ММ могут выражаться совокупностью уравнений линий и поверхностей, алгебрологических соотношений, описывающих области, составляющие тело объекта; графами и списками, отображающими конструкции из типовых конструктивных элементов, и т. п. Геометрические ММ применяют при решении задач конструирования в приборостроении, радиоэлектронике, для оформления конструкторской документации, при задании исходных данных на разработку технологических процессов изготовления деталей. Используют несколько типов геометрических ММ.

Для отображения геометрических свойств деталей со сравнительно несложными поверхностями применяют ММ, представляемые в аналитической или алгебрологической форме (аналитические, алгебрологические). *Аналитические* ММ – уравнения поверхностей и линий. Например, уравнение плоскости имеет вид $ax + by + cz + d = 0$, где x, y, z – пространственные координаты; a, b, c, d – коэффициенты уравнений. В *алгебрологических* ММ тела описываются системами логических выражений, отражающих условия принадлежности точек внутренним областям тел.

Для сложных поверхностей аналитические и алгебрологические модели оказываются слишком громоздкими, их трудно получать и неудобно использовать. Область их применения обычно ограничивается поверхностями плоскими и второго порядка.

В машиностроении для отображения геометрических свойств деталей со сложными поверхностями применяют ММ каркасные и кинематические.

Каркасные (сеточные) ММ представляют собой конечные множества элементов, например точек или кривых, принадлежащих моделируемой поверхности. В частности, выбор каркаса в виде линий, образующих сетку на описываемой поверхности, приводит к разбиению поверхности на отдельные участки. Кусочно-линейная аппроксимация на этой сетке устраняет главный недостаток аналитических моделей, так как в пределах каждого из участков, имеющих малые размеры, возможна удовлетворительная по точности аппроксимация поверхностями с простыми уравнениями. Коэффициенты этих уравнений рассчитываются исходя из условий плавности сопряжений участков.

Функциональные ММ предназначены для отображения физических или информационных процессов, протекающих в объекте при его функционировании или изготовлении. Обычно функциональные ММ представляют собой системы уравнений, связывающих фазовые переменные, внутренние, внешние и выходные параметры.

Деление описаний объектов на аспекты и иерархические уровни непосредственно касается математических моделей. Выделение аспектов описания приводит к выделению моделей электрических, механических, гидравлических, оптических, химических и т. п., причём модели процессов функционирования изделий и модели процессов их изготовления различные, например модели полупроводниковых элементов интегральных схем, описывающих процессы диффузии и дрейфа подвижных носителей заряда в полупроводниковых областях при функционировании прибора и процессы диффузии примесей в полупроводнике при изготовлении прибора.

В зависимости от места в иерархии описаний математические модели делятся на модели, относящиеся к микро-, макро- и метауровням.

Использование принципов блочно-иерархического подхода к проектированию приводит к появлению иерархии математических моделей проектируемых объектов. Количество иерархических уровней при моделировании определяется сложностью проектируемых

объектов и возможностью средств проектирования. Однако для большинства предметных областей имеющиеся иерархические уровни можно отнести к одному из трёх обобщённых уровней, названных выше.

Особенность ММ на *микроуровне* – отражение физических процессов, протекающих в непрерывном пространстве и времени. Типичные ММ на микроуровне – дифференциальные уравнения в частных производных (ДУЧП). В них независимыми переменными являются пространственные координаты и время. С помощью этих уравнений рассчитываются поля механических напряжений и деформаций, электрических потенциалов, давлений, температур и т. п. Возможности применения ММ в виде ДУЧП ограничены отдельными деталями, попытки анализировать с их помощью процессы в многокомпонентных средах, сборочных единицах, электронных схемах не могут быть успешными из-за чрезмерного роста затрат машинного времени и памяти.

На *макроуровне* используют укрупнённую дискретизацию пространства по функциональному признаку, что приводит к представлению ММ на этом уровне в виде систем обыкновенных дифференциальных уравнений (ОДУ). В этих уравнениях независимая переменная – время t , а вектор зависимых переменных V составляют фазовые переменные, характеризующие состояние укрупнённых элементов дискретизированного пространства. Такими переменными являются силы и скорости механических систем, давления и расходы гидравлических и пневматических систем и т. п. Системы ОДУ – универсальные модели на макроуровне, пригодные для анализа как динамических, так и установившихся состояний объектов. Модели для установившихся режимов также можно представить в виде систем алгебраических уравнений. Порядок системы уравнений зависит от числа выделенных элементов объекта. Если порядок системы приближается к 10^3 , то оперирование моделью становится затруднительным и поэтому необходимо переходить к представлениям на *метауровне*.

На *метауровне* в качестве элементов принимают достаточно сложные совокупности деталей. Метауровень характеризуется большим разнообразием типов используемых ММ. Для многих объектов ММ на метауровне по-прежнему представляются системами ОДУ. Однако так как в моделях не описываются внутренние для элементов

фазовые переменные, а фигурируют только фазовые переменные, относящиеся к взаимным связям элементов, то укрупнение элементов на метауровне означает получение ММ приемлемой размерности для существенно более сложных объектов, чем на макроуровне.

В ряде предметных областей удаётся использовать специфические особенности функционирования объектов для упрощения ММ, например для электронных устройств цифровой автоматики возможно применять дискретное представление таких фазовых переменных, как напряжения и токи. В результате ММ становится системой логических уравнений, описывающих процессы преобразования сигналов. Такие логические модели существенно более экономичны, чем модели электрические, описывающие изменения напряжений и сил токов непрерывными функциями времени. Важный класс ММ на метауровне составляют модели массового обслуживания, применяемые для описания процессов функционирования информационных и вычислительных систем, производственных участков, линий и цехов.

Структурные модели также делятся на модели различных иерархических уровней. При этом на низших иерархических уровнях преобладает использование геометрических моделей, на высших иерархических уровнях – топологических моделей.

По степени детализации описания в пределах каждого иерархического уровня выделяют полные ММ и макромодел.

Полная ММ – модель, в которой фигурируют фазовые переменные, характеризующие состояния всех имеющихся межэлементных связей (т. е. состояния всех элементов проектируемого объекта).

Макромодель – ММ, в которой отображаются состояния значительно меньшего числа межэлементных связей, что соответствует описанию объекта при укрупнённом выделении элементов. (Понятия «полная ММ» и «макромодель» относительны и обычно используются для различения двух моделей, отображающих различную степень детальности описания свойств объекта.)

По способу представления свойств объекта функциональные ММ делятся на аналитические, алгоритмические, имитационные.

Аналитические ММ представляют собой явные выражения выходных параметров как функций входных и внутренних параметров. Такие ММ характеризуются высокой экономичностью, однако получение аналитической ММ удаётся лишь в отдельных частных случаях, как правило, при принятии существенных допущений и ограничений, снижающих точность и сужающих область адекватности модели.

Алгоритмические ММ выражают связи выходных параметров с параметрами внутренними и внешними в форме алгоритма. Типичная алгоритмическая ММ – система уравнений, дополненная алгоритмом решения выбранного численного метода и алгоритмом вычисления вектора выходных параметров как функционалов решения системы уравнений $V(z)$.

Имитационная ММ – алгоритмическая модель, отражающая поведение исследуемого объекта во времени при задании внешних воздействий на него. Примерами имитационных ММ могут служить модели динамических объектов в виде систем ОДУ и модели систем массового обслуживания, заданные в алгоритмической форме.

Для получения ММ используют методы неформальные и формальные.

По способу получения модели делятся на теоретические и эмпирические. *Теоретические* ММ создаются в результате исследования процессов и их закономерностей, присущих рассматриваемому классу объектов и явлений; *эмпирические* ММ – в процессе изучения внешних проявлений свойств объекта с помощью измерений фазовых переменных на внешних входах и выходах, обработки результатов измерений.

Понятие об оптимизации систем

В последние годы становится всё более актуальной задача создания не только удовлетворяющих исходным данным, но и оптимальных (наилучших) систем. Это объясняется тем, что с каждым годом возрастают требования, предъявляемые к ЭС, время и средства, затрачиваемые на их изготовление. Поэтому весьма существенно не просто удовлетворить исходным требованиям к показателям качества системы, но и перевыполнить эти требования (уменьшить по сравнению с запланированным стоимость системы и время её разработки, повысить надёжность и т. д.).

Оптимизация ЭС включает в себя оптимизацию как собственно разрабатываемой системы, так и процесса её разработки. Обе эти стороны оптимизации взаимно связаны. Показатели качества разработанной системы существенно зависят от степени оптимальности процесса разработки и от отпущенных на неё времени и средств.

Мы будем рассматривать в основном оптимизацию двух подсистем: конструкции и технологии ЭС, поскольку от степени оптимальности этих подсистем зависят показатели качества ЭС. Очевидно, что в задачу оптимизации конструкции и технологии входит выбор наилучшего варианта конструкции, технологического процесса и наилучших режимов его проведения.

Оптимизация системы осуществляется обычно как на этапах технического предложения и эскизного проекта, так и на всех последующих этапах. Однако важно осуществить оптимизацию в возможно большей степени на самых ранних этапах разработки, так как при этом она может быть наиболее радикальной и требует меньших экономических затрат. В дальнейшем отыскание оптимальной системы (конструкции) называется *синтезом системы* (конструкции).

Синтез конструкции обычно включает решение следующих задач.

1. Синтез оптимальной структуры конструкции, т. е. отыскание оптимальных принципов построения конструкции (синтез структуры).

2. Выбор оптимальных значений параметров конструкции (оптимизация параметров).

Оптимизация технологического процесса включает решение следующих задач.

1. Выбор оптимального варианта технологического процесса (дискретный выбор техпроцесса).

2. Выбор оптимальных режимов проведения техпроцесса (оптимизация параметров техпроцесса).

Оптимизация конструкции и технологического процесса производится обычно коллективами, возглавляемыми главным конструктором и главным технологом, путём сочетания различных методов. Различают следующие методы оптимизации: математический и эвристический.

Математический синтез заключается в математической формулировке (записи) совокупности исходных данных и критерия предпочтения и отыскания чисто математическим путём такого решения, которое обеспечивает наилучшее значение варианта конструкции или технологии.

Под *эвристическим синтезом* понимают сложный и многообразный творческий процесс, заключающийся в отыскании инженерным коллективом приемлемых решений на основе использования накопленных данных, собственного инженерного опыта, приближённых расчётов, инженерной интуиции и творческих способностей членов коллектива.

Сопоставление и взаимная коррекция данных, полученных математическими и эвристическими методами, и позволяют получить оптимальные результаты. Такой комбинированный процесс сочетания математических и эвристических методов будем называть *инженерным синтезом*.

Инженерный синтез связан с информационно-поисковыми работами. Прежде чем приступить к выбору вариантов конструкции и технологического процесса, конструктору и технологу нужно получить информацию по существу дела. Для этой цели служат соответствующие информационно-поисковые системы.

Аналогичным образом инженерный синтез связан с использованием ЕСКД, ЕСТД и ЕСТПП и других систем документов. Использование этих систем позволяет сократить сроки разработки и производства изделий.

В настоящее время оптимизация конструкции и технологии – процессы весьма сложные, требующие перебора ряда вариантов с громоздкими и трудоемкими расчетами, которые необходимо выполнять с помощью ЭВМ.

В общем случае решение проблемы создания системы – одна из задач последовательного приближения. Так, варианты конструкции и технологического процесса вначале сравниваются по выбранным критериям и модифицируются так, чтобы в большей степени удовлетворить исходным требованиям.

Отобранный модифицированный (квазиоптимальный) вариант вновь анализируется и совершенствуется до тех пор, пока не будут решены поставленные задачи и удовлетворены заданные показатели функционирования. При этом для оценки качества промежуточных решений используют математический и экспериментальный методы анализа.

ОСНОВЫ ТЕОРИИ ТОЧНОСТИ ЭЛЕКТРОННЫХ СРЕДСТВ

Лекция 5

ПОГРЕШНОСТИ И ИХ МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ

Понятие о погрешностях. Классификация погрешностей

Под *погрешностью* понимается разность между действительным и заданным значениями параметра. Иногда её называют ошибкой, или отклонением.

По характеру погрешности делят:

- на *систематические* – входят в результаты опыта в качестве постоянной величины и с постоянным знаком; они могут быть учтены и исключены из результатов опыта;

- *случайные* – входят в результаты опыта в качестве случайной величины со случайным знаком; они могут быть уменьшены, но полностью их исключить невозможно;

грубые – резко выделяются на фоне других погрешностей.

Погрешности можно также разделить по стадиям их появления:

- *при проектировании*: погрешности выбранных закономерностей, формул, законов, вычислений, округлений, технической документации и др.;

- *при производстве*: погрешности параметров исходных материалов, оборудования, приспособлений, обслуживающего персонала, погрешности измерений, испытаний и пр.;

- *при эксплуатации, хранении и транспортировании*: *объективные погрешности*, обусловленные действием внешних факторов (повышенной или пониженной температуры, давления, вибраций, процессов старения компонентов и пр.); *субъективные погрешности*, обусловленные действием обслуживающего персонала.

Таким образом, погрешности зависят от многих случайных факторов, поэтому являются случайными величинами.

Математическое описание погрешностей

Случайной величиной называется такая переменная величина, которая в результате опыта может принимать одно из возможных, заранее неизвестных значений.

Случайные величины делятся:

- на *дискретные*, которые в результате опыта могут принимать конечное количество значений; свойством дискретной случайной величины является равенство суммы вероятностей всех возможных значений случайной величины единице, то есть

$$\sum_{i=1}^n p_i = 1;$$

- *непрерывные* – количество значений бесконечно, поэтому непрерывные величины описываются интегральной функцией

$$\int_{-\infty}^{+\infty} f(x) dx = 1,$$

где $f(x)dx$ имеет, как будет показано дальше, смысл вероятности.

Наиболее полной характеристикой случайной величины является ее *закон распределения*, который устанавливает связь между всеми возможными значениями случайной величины и соответствующими им вероятностями.

Законы распределения могут иметь различную форму, но наиболее часто они представляются в виде ряда, функции или плотности распределения.

Ряд распределения применяется только для дискретных случайных величин и обычно приводится в виде таблицы (табл. 5.1), но может быть представлена и графически.

Таблица 5.1

X	x_1	x_2	x_3	...	x_n
P	p_1	p_2	p_3	...	p_n

Функция распределения – это вероятность того, что случайная величина X не превысит некоторого заданного значения x , то есть

$$F(x) = P(X < x).$$

Суть функции распределения можно пояснить следующим образом – это вероятность попадания случайной величины X на участок от $-\infty$ до x (рис. 5.1).

Свойства функции распределения: 1) при $x = \infty F(x) = F(+\infty) = 1$; 2) при $x = -\infty F(x) = F(-\infty) = 0$; 3) функция распределения – это неубывающая функция своего аргумента, если $a_1 > a_2$, то $F(a_1) \geq F(a_2)$.

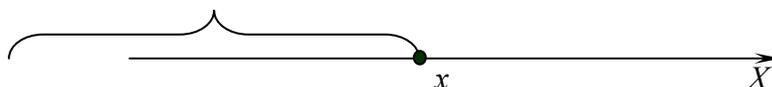


Рис. 5.1

Графически функцию распределения можно представить в виде, показанном на рис. 5.2: а – непрерывная случайная величина; б – дискретная случайная величина.

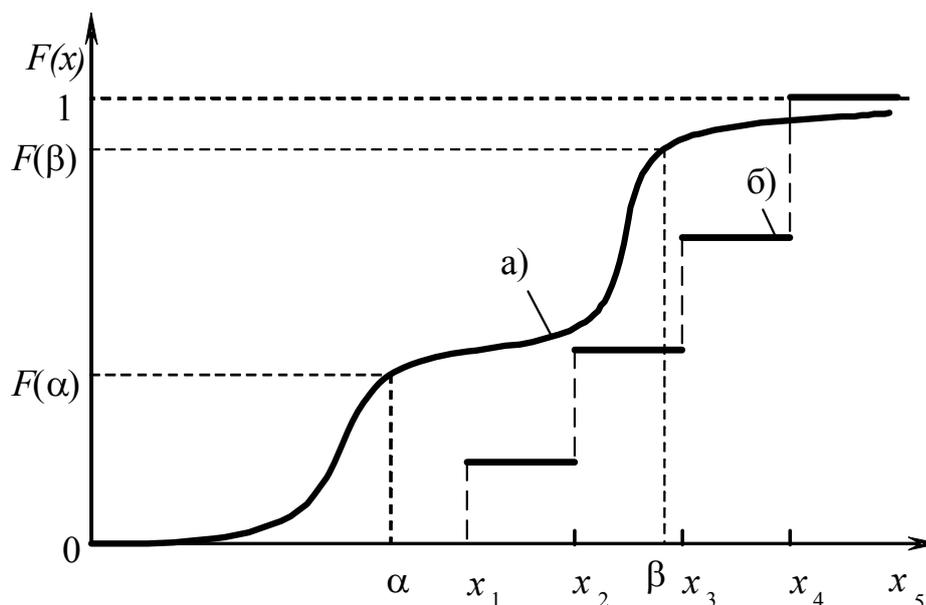


Рис. 5.2

Зная функцию распределения, легко определить вероятность попадания случайной величины в заданный интервал (от α до β):

$$P(\alpha \leq X \leq \beta) = F(\beta) - F(\alpha).$$

Плотность распределения $f(x)$ связана с функцией распределения следующим соотношением:

$$f(x) = F'(x) = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{P(\Delta x)}{\Delta x}.$$

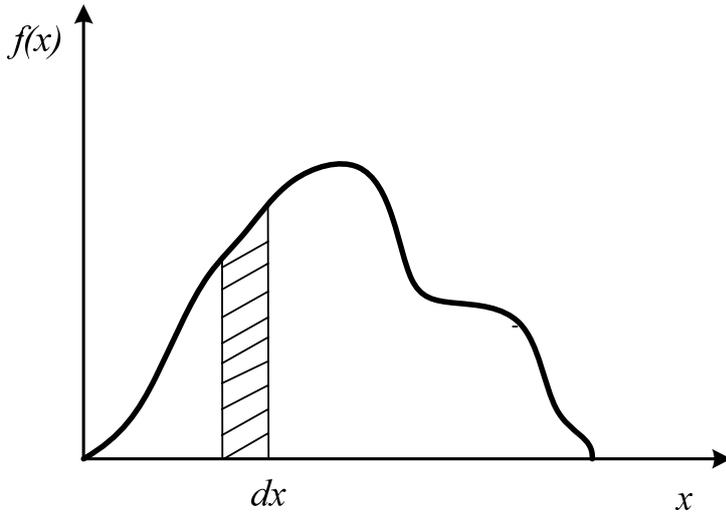


Рис. 5.3

Таким образом, плотность распределения случайной величины есть плотность распределения ее вероятности. Примерный вид плотности распределения показан на рис. 5.3.

Из предыдущего соотношения получим два важных выражения:

$$F(x) = \int_{-\infty}^x f(x) dx;$$

$$P(dx) = f(x) dx .$$

Последнее выражение получило название *элемента вероятности*. Из рис. 5.3 видно, что он численно равен площади участка dx под кривой плотности распределения.

Плотность распределения имеет два свойства: 1) $f(x) \geq 0$;
 2) $\int_{-\infty}^{+\infty} f(x) dx = 1$.

На практике часто применяют типовые плотности распределения, называемые законами распределения: закон равномерной плотности, нормальный закон распределения, законы Пуассона, Вейбулла, Релея и т.д.

Закон равномерной плотности. Здесь плотность одинакова на всем участке (график плотности распределения показан на рис. 5.4.):

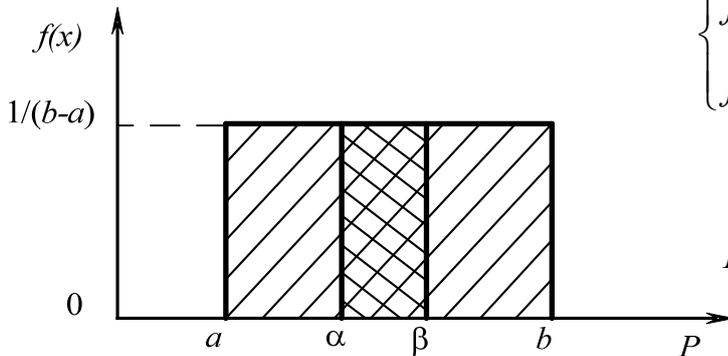


Рис. 5.4

$$\begin{cases} f(x) = \frac{1}{b-a} & \text{при } a \leq x < b, \\ f(x) = 0 & \text{при } x < a \text{ или } x > b, \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{при } x < a, \\ \frac{x-a}{b-a} & \text{при } a \leq x < b, \\ 1 & \text{при } x > b. \end{cases}$$

$$\text{Отсюда } P(\alpha \leq x \leq \beta) = \frac{\beta - \alpha}{b - a}.$$

Нормальный закон распределения. График плотности распределения показан на рис. 5.5.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}},$$

где m и σ – параметры закона, называемые *математическим ожиданием* и *средним квадратическим отклонением*, или *стандартом*. Для нормального закона распределения справедливо соотношение $P(\alpha \leq x < \beta) = \Phi\left(\frac{\beta - m}{\sigma}\right) - \Phi\left(\frac{\alpha - m}{\sigma}\right)$,

где $\Phi(\dots)$ – табулированная функция Лапласа.

С нормальным законом распределения связано правило «трех сигм», говорящее о том, что вероятность попадания X в интервал $(m - 3\sigma ; m + 3\sigma)$ равна 0,9973 (99,73 %).

На практике действия с законами распределения часто бывают громоздкими, поэтому используются некоторые числовые характеристики законов (параметры), отражающие наиболее характерные черты этих законов: математическое ожидание, дисперсия, асимметрия, эксцесс.

Математическое ожидание определяет центр группирования случайной величины.

Математическое ожидание дискретной случайной величины есть сумма произведений всех возможных значений случайной величины на соответствующие им вероятности:

$$m_x = M[X] = \sum_{i=1}^n x_i p_i.$$

Для непрерывной случайной величины

$$m_x = \int_{-\infty}^{+\infty} x f(x) dx,$$

где $f(x)dx$ – элемент вероятности, определяющий вероятность попадания непрерывной случайной величины бесконечно малый интервал dx .

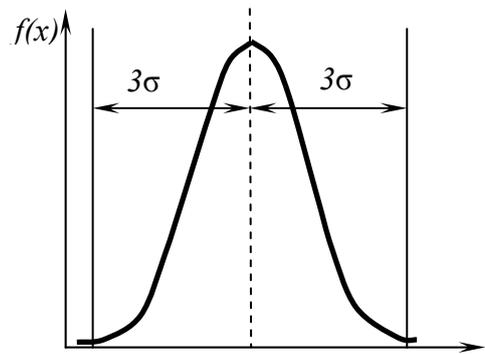


Рис. 5.5

Для характеристики разброса случайной величины пользуются понятиями дисперсии или среднего квадратического отклонения.

Дисперсия – математическое ожидание квадрата центрированной случайной величины

$$D_x = D[X] = M[\overset{\circ}{X}]^2,$$

где $\overset{\circ}{X}$ – *центрированная случайная величина* – разность между значением случайной величины и ее математическим ожиданием

$$\overset{\circ}{X} = X - m_x.$$

Дисперсия дискретной случайной величины

$$D_x = \sum_{i=1}^n (x_i - m)^2 p_i.$$

Дисперсия непрерывной случайной величины

$$D_x = \int_{-\infty}^{+\infty} (x - m)^2 f(x) dx.$$

Использование дисперсии на практике неудобно, так как она имеет размерность квадрата случайной величины, поэтому вместо дисперсии часто применяют *среднее квадратическое отклонение* (СКО), называемое также *стандартом*:

$$\sigma_x = \sqrt{D_x}.$$

Для характеристики асимметрии распределения применяют *коэффициент асимметрии* А (рис. 5.6), а для характеристики острровершинности – *коэффициент эксцесса* Е (рис. 5.7).

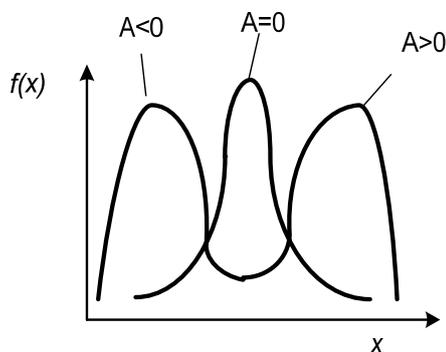


Рис. 5.6

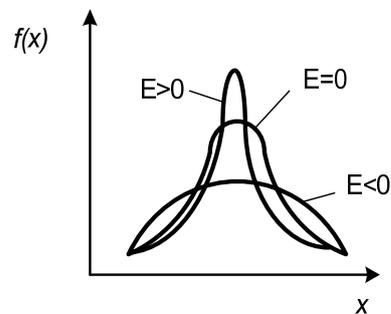


Рис. 5.7

У нормального закона распределения $A=0$ и $E=0$.

Математическое ожидание, дисперсия, стандарт, коэффициенты асимметрии, островершинности – общие (интегральные) характеристики распределения. Применяется также группа параметров, характеризующая отдельные значения случайной величины. Это квантиль, мода и медиана.

Квантилем x_p распределения случайной величины X называется значение случайной величины, для которого выполняется условие

$$P(X < x_p) = p.$$

Медиана (Me) – это значение случайной величины, для которого выполняется условие

$$P(X < Me) = 0,5,$$

а следовательно, и

$$P(X > Me) = 0,5.$$

Медиана равна квантилю $x_{0,5}$.

Мода – это значение случайной величины, которому соответствует наибольшая ордината плотности распределения.

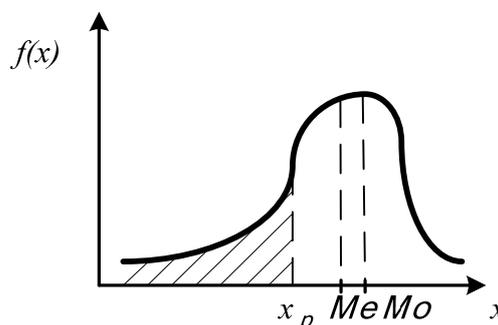


Рис. 5.8

Случайные величины могут

быть зависимыми и независимыми друг от друга. Эта зависимость (связь) может быть линейной и нелинейной.

Линейную связь обычно называют *корреляционной*. Она характеризуется *корреляционным моментом*, называемым также *ковариацией*. Его часто обозначают K_{xy} или $K[xy]$. Это есть математическое ожидание произведения центрированных случайных величин $\overset{\circ}{X}$ и $\overset{\circ}{Y}$, то есть $K_{xy} = M[\overset{\circ}{X}\overset{\circ}{Y}]$, где $\overset{\circ}{X} = X - m_x$; $\overset{\circ}{Y} = Y - m_y$.

В соответствии с формулами для математических ожиданий получим для дискретных случайных величин

$$K_{xy} = \sum_{i=1}^n \sum_{j=1}^m (x_i - m_x)(y_j - m_y)p_i p_j$$

и для непрерывных случайных величин

$$K_{xy} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_x)(y - m_y)f(x)f(y)dx dy.$$

Коэффициент корреляции между двумя случайными величинами $r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y}$ может принимать значения в диапазоне от -1 до 1 , то есть $-1 \leq r_{xy} \leq 1$ (рис. 5.9).

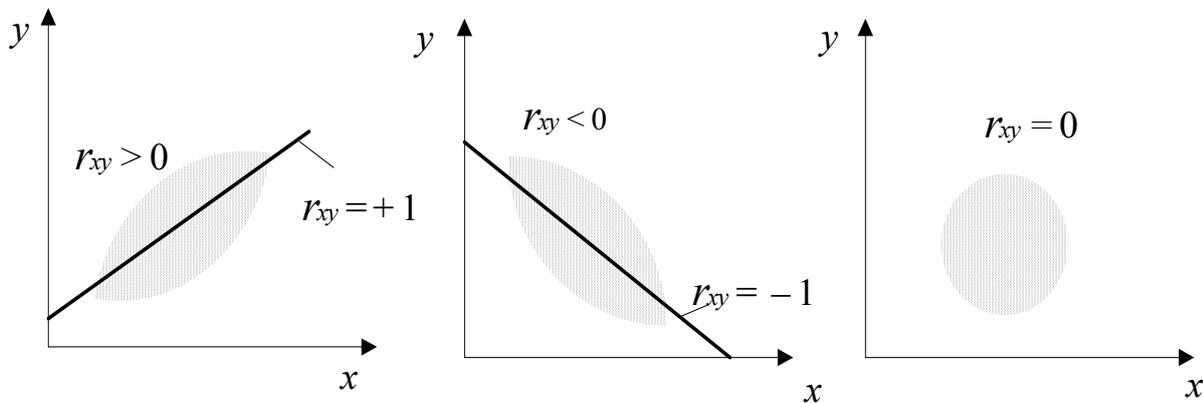


Рис 5.9

Действия с числовыми характеристиками случайных величин

1. Математическое ожидание неслучайной величины C равно этой величине

$$M[C] = C.$$

2. Дисперсия неслучайной величины C равна 0

$$D[C] = 0.$$

3. Математическое ожидание произведения неслучайной величины на случайную имеет вид

$$M[CX] = CM[X].$$

4. Дисперсия произведения неслучайной величины на случайную имеет вид

$$D[CX] = C^2 D[X].$$

5. Математическое ожидание суммы двух случайных величин равно сумме математических ожиданий этих величин

$$M[X + Y] = M[X] + M[Y].$$

6. Дисперсия суммы двух случайных величин есть сумма дисперсий с удвоенным корреляционным моментом

$$D[X + Y] = D[X] + D[Y] + 2K_{xy}.$$

Статистические оценки

На практике вид закона распределения случайной величины и их числовые характеристики оцениваются на основе экспериментальных (статистических) данных.

Пусть общее число статистических данных равно n . Весь интервал полученных значений от x_{\min} до x_{\max} разбивается на некоторое количество подынтервалов, обычно не менее 10, и подсчитывается количество m_i попаданий параметра в каждый из них.

Результаты обработки экспериментальных данных в качестве примера приведены в табл. 5.2.

Таблица 5.2

Интервал	Середина интервала, \bar{x}_i	Частота, m_i	Относительная частота, P_i^*	Накопленная частота, $\sum m_i$	Относительная накопленная частота, $F^*(x)$
970 – 980	975	1	0,012	1	0,012
980 – 990	985	3	0,036	4	0,048
...

Относительная частота определяется по формуле

$$P_i^* = \frac{m_i}{n},$$

где m_i – абсолютная частота попадания параметра x в интервал; n – общее число статистических данных.

Примерные виды гистограммы, накопленного полигона и кумулятивной кривой, которые являются статистическими аналогами плотности и функции распределения, изображены на рис. 5.10 и 5.11, где обозначено: 1 – накопленный полигон, 2 – кумулятивная кривая.

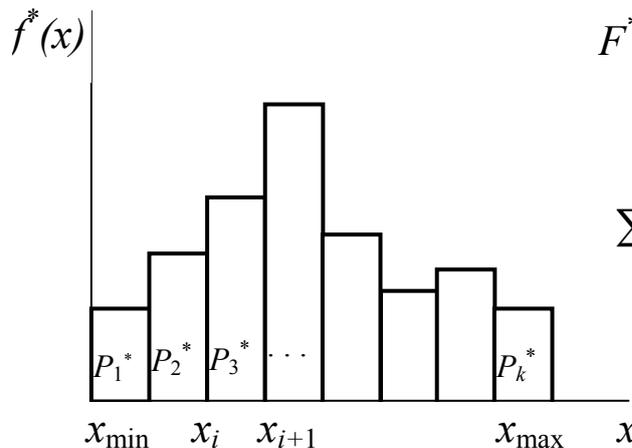


Рис. 5.10

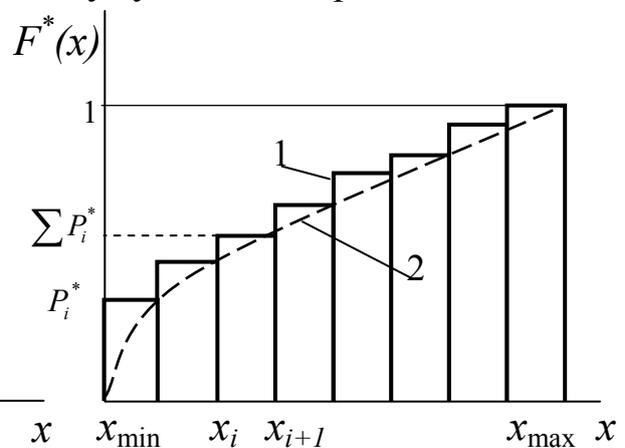


Рис. 5.11

Высота прямоугольника гистограммы находится по формуле

$$f_i^*(x) = \frac{P_i^*}{x_{i+1} - x_i}.$$

Высота прямоугольника кумулятивной кривой равна

$$F^*(x_i) = \sum_{i=1}^m P_i^*,$$

где m – число суммируемых частот до $x=x_i$.

Оценки числовых характеристик должны быть состоятельными, несмещенными и эффективными [2].

Среднее арифметическое значение параметра (статистическое математическое ожидание) рассчитывается по формуле

$$m_x^* = \frac{\sum_{j=1}^m x_j}{n}$$

или приближенно

$$m_x^* = \sum_{i=1}^k \bar{x}_i P_i^*,$$

где k – число разрядов, \bar{x}_i – середина i -го интервала.

Статистическая дисперсия имеет вид

$$D_x^* = \frac{\sum_{i=1}^n (x_i - m_x^*)^2}{(n-1)}$$

или приближенно

$$D_x^* = \sum_{i=1}^k (\bar{x}_i - m_x^*)^2 P_i^*.$$

Лекция 6

УРАВНЕНИЯ ПОГРЕШНОСТЕЙ. КОЭФФИЦИЕНТЫ ВЛИЯНИЯ

Уравнения погрешностей

Каждая система (электронное средство, технологический процесс) характеризуется своими выходными параметрами Y (быстродействие, мощность, скорость передачи информации и т.д.), которые зависят от первичных параметров x и под которыми могут пониматься параметры электрорадиоэлементов (ЭРЭ), размеры деталей и т.д.

То есть выходной параметр

$$Y = f(x_1, x_2, \dots, x_n). \quad (6.1)$$

Все первичные параметры имеют погрешности $\Delta x_1, \Delta x_2, \dots, \Delta x_n$, которые возникают в силу различных причин. Тогда и каждый выходной параметр Y имеет погрешность ΔY .

Следовательно, можно записать

$$Y + \Delta Y = f(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_i + \Delta x_i).$$

Если функция непрерывна и дифференцируема на каком-то интервале, то она может быть разложена в ряд Тейлора. Это разложение имеет вид

$$Y + \Delta Y = f(x_1, x_2, \dots, x_n) + \frac{\partial y}{\partial x_1} \Delta x_1 + \frac{\partial y}{\partial x_2} \Delta x_2 + \dots + \frac{\partial y}{\partial x_n} \Delta x_n + \frac{1}{2} \frac{\partial^2 y}{\partial x_i^2} \Delta x_i^2 + \dots$$

Если погрешности Δx_i ($i = 1, 2, \dots, n$) малые, то всеми слагаемыми, где Δx в квадрате или более высокой степени, можно пренебречь. Погрешности часто считаются малыми, если их величина не превышает 5 % от номинального значения. Тогда, учитывая (6.1), можно получить

$$\Delta Y = \frac{\partial y}{\partial x_1} \Delta x_1 + \frac{\partial y}{\partial x_2} \Delta x_2 + \dots + \frac{\partial y}{\partial x_n} \Delta x_n.$$

Или более компактно

$$\Delta Y = \sum_{i=1}^n A_i \Delta x_i. \quad (6.2)$$

Коэффициент влияния A_i показывает, как влияет погрешность i -го параметра на погрешность выходного параметра. Уравнение (6.2) называется *уравнением абсолютных погрешностей*, а коэффициент

$$A_i = \frac{\partial y}{\partial x_i} \quad (6.3)$$

– *коэффициентом влияния*.

Уравнение относительных погрешностей получим после несложных преобразований уравнения (6.2):

$$\frac{\Delta Y}{Y} = \sum_{i=1}^n \left(\frac{\partial y}{\partial x_i} \frac{x_i}{y} \right) \frac{\Delta x_i}{x_i}.$$

Его можно записать в компактной форме

$$\frac{\Delta Y}{Y} = \sum_{i=1}^n B_i \frac{\Delta x_i}{x_i}, \quad (6.4)$$

где коэффициенты влияния B_i определяются по формуле

$$B_i = \frac{\partial y}{\partial x_i} \frac{x_i}{Y}. \quad (6.5)$$

Расчет по формулам (6.2) или (6.4) не представляет труда, если известны погрешности первичных параметров (как правило, из технической документации – чертежей, перечня элементов и т.д.) и коэффициенты влияния.

Методы определения коэффициентов влияния

Наибольшее применение нашли два метода: расчетно-аналитический и экспериментальный, или метод малых приращений.

Расчетно-аналитический метод применяется, когда имеется аналитическое выражение, связывающее выходные и первичные параметры.

Для определения коэффициентов влияния используются формулы (6.3) и (6.5).

Пример. Найти коэффициенты влияния B_x и B_z , если имеем функцию $y = x^2 z^3$.

Применяя формулу (6.5), найдем B_x и B_z :

$$B_x = \frac{\partial y}{\partial x} \frac{x}{y} = 2xz^3 \frac{x}{x^2 z^3} = 2;$$

$$B_z = \frac{\partial y}{\partial z} \frac{z}{y} = x^2 3z^2 \frac{z}{x^2 z^3} = 3.$$

Аналитический метод имеет два недостатка.

1. Не всегда имеется необходимое аналитическое выражение, связывающее выходной и первичный параметры.

2. В аналитических выражениях число первичных параметров может быть большим, и они могут находиться в сложной взаимосвязи. Определение частных производных в этом случае затруднительно и используется экспериментальный метод (метод малых приращений).

Метод малых приращений метод основан на линейности уравнения погрешностей

$$\frac{\Delta y}{y} = B_1 \frac{\Delta x_1}{x_1} + \dots + B_i \frac{\Delta x_i}{x_i} + \dots + B_n \frac{\Delta x_n}{x_n}.$$

Так как это уравнение линейно, оно обладает свойством суперпозиции, которое заключается в независимости влияния погрешности каждого первичного параметра на погрешность выходного параметра.

Считая погрешности всех параметров, кроме одного, равными нулю, можно получить уравнение

$$\frac{\Delta y}{y} = B_i \frac{\Delta x_i}{x_i}.$$

Из него найдем

$$B_i = \frac{\Delta y / y_i}{\Delta x_i / x_i} \quad (i = 1, 2, 3, \dots, n).$$

Лекция 7

МЕТОДЫ РАСЧЕТА ТОЧНОСТИ

Понятие о допусках

Допуски устанавливают границы, в которых должны находиться параметры. Их подразделяют на производственные, лабораторные, эксплуатационные. Производственные допуски ограничивают погрешности, возникающие при производстве. Лабораторные допуски ограничивают производственные погрешности и отклонения, возникающие при испытаниях (лабораторный допуск должен быть шире производственного). Эксплуатационные допуски ограничивают погрешности, возникающие при эксплуатации (испытание на повышенных температурах).

Допуски подразделяются также на механические и электрические в зависимости от типа параметров, которые они ограничивают.

Предельные значения параметров $Y_{\text{пр}}$ рассчитываются по формуле

$$Y_{\text{пр}} = Y_{\text{н}} + E\left(\frac{\Delta y}{y}\right) \pm \delta\left(\frac{\Delta y}{y}\right), \quad (7.1)$$

где $Y_{\text{н}}$ – номинальное значение; $E\left(\frac{\Delta y}{y}\right)$ – отклонение середины поля допуска от номинального значения (координаты середины поля допуска); $\delta\left(\frac{\Delta y}{y}\right)$ – половина поля допуска.

На использовании уравнений погрешностей основываются методы предельных отклонений и вероятностей.

Метод предельных отклонений

Если использовать уравнение погрешностей $\frac{\Delta y}{y} = \sum_{i=1}^n B_i \frac{\Delta x_i}{x_i}$, то

можем записать:

$$E\left(\frac{\Delta y}{y}\right) = \sum_{i=1}^n B_i E\left(\frac{\Delta x_i}{x_i}\right); \quad (7.2)$$

$$\delta\left(\frac{\Delta y}{y}\right) = \sum_{i=1}^n |B_i| \delta\left(\frac{\Delta x_i}{x_i}\right). \quad (7.3)$$

Эти выражения – основные при расчете допусков методом предельных отклонений.

Вероятностный метод расчета электрических допусков

Необходимо определить $E\left(\frac{\Delta y}{y}\right)$ и $\delta\left(\frac{\Delta y}{y}\right)$.

Начнем с решения первой задачи. Используя уравнения погрешностей и применяя правило сложения математических ожиданий, запишем

$$M\left(\frac{\Delta y}{y}\right) = \sum_{i=1}^n B_i M\left(\frac{\Delta x_i}{x_i}\right). \quad (7.4)$$

Связь между математическим ожиданием и отклонением E устанавливается известным уравнением Бородачева:

$$M\left(\frac{\Delta y}{y}\right) = E\left(\frac{\Delta y}{y}\right) + a_y \delta\left(\frac{\Delta y}{y}\right)$$

или

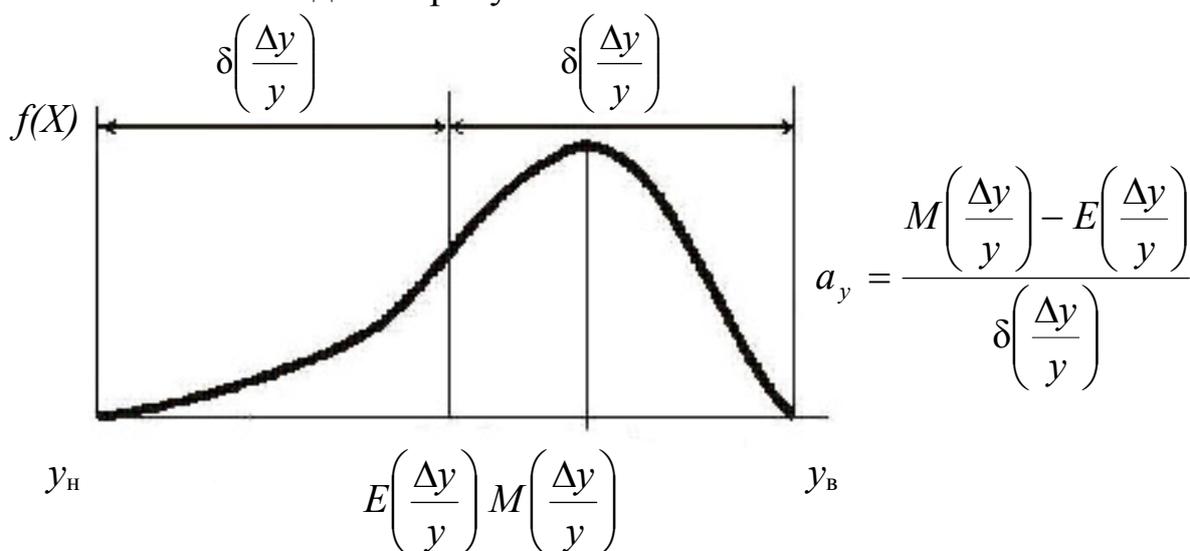
$$M\left(\frac{\Delta x}{x}\right) = E\left(\frac{\Delta x}{x}\right) + a_x \delta\left(\frac{\Delta x}{x}\right).$$

Подставляя эти выражения в (7.4), получим

$$E\left(\frac{\Delta y}{y}\right) + a_y \delta\left(\frac{\Delta y}{y}\right) = \sum_{i=1}^n B_i \left[E_i\left(\frac{\Delta x}{x}\right) + a_i \delta_i\left(\frac{\Delta x}{x}\right) \right], \quad (7.5)$$

где a – коэффициент относительной асимметрии.

Его смысл виден из рисунка.



Так как радиоэлементы изготавливаются в крупносерийном и массовом производстве, то их погрешности подчиняются, как правило, нормальному закону распределения. В этом случае коэффициенты относительной асимметрии равны нулю. Выражение (7.5) принимает вид

$$E\left(\frac{\Delta y}{y}\right) = \sum_{i=1}^n B_i E\left(\frac{\Delta x_i}{x_i}\right).$$

Как видим, это уравнение полностью совпадает с уравнением (7.2), применяемым при методе предельных отклонений.

Определим теперь выражение для половины поля допуска $\delta\left(\frac{\Delta y}{y}\right)$ выходного параметра.

Если погрешности независимы, то, используя правило сложения дисперсий, можно записать следующие выражения:

$$D\left(\frac{\Delta y}{y}\right) = \sum_{i=1}^n B_i^2 D\left(\frac{\Delta x_i}{x_i}\right)$$

или

$$\sigma^2\left(\frac{\Delta y}{y}\right) = \sum_{i=1}^n B_i^2 \sigma^2\left(\frac{\Delta x_i}{x_i}\right). \quad (7.6)$$

Введем величину относительного рассеивания

$$\lambda = \frac{\sigma}{\delta}, \quad (7.7)$$

где σ – среднеквадратичное значение погрешностей; δ – половина поля допуска.

Применим также величину K – коэффициент относительного рассеивания

$$K = \frac{\lambda}{\lambda_3}, \quad (7.8)$$

где λ_3 – относительное рассеивание эталонного распределения, в качестве которого обычно принимается нормальное распределение. Используя выражения (7.7) и (7.8), получаем $\sigma^2 = \lambda^2 \delta^2 = K^2 \lambda_3^2 \delta^2$.

Подставляя это выражение в (7.6), запишем

$$K^2 \lambda_3^2 \delta_y^2 = \sum_{i=1}^n B_i^2 K_i^2 \lambda_3^2 \delta_i^2,$$

где

$$\delta_y = \delta\left(\frac{\Delta y}{y}\right); \quad \delta_i = \delta\left(\frac{\Delta x_i}{x_i}\right),$$

или

$$\delta_y = \frac{1}{K_y} \sqrt{\sum_{i=1}^n B_i^2 K_i^2 \delta_i^2}, \quad (7.9)$$

где B_i – коэффициент влияния; K_i – коэффициент относительного рассеивания i -го первичного параметра; K_y – коэффициент относительного рассеивания выходного параметра.

Если законы распределения первичных параметров нормальные, то коэффициенты относительного рассеивания K_i и K_y равны единице.

Пример. Определить предельные отклонения параметра $y = x^2 z^3$, если $x = 5^{+0,1}$; $z = 2 \pm 0,05$.

Решение. Применяя метод предельных отклонений получим

$$E\left(\frac{\Delta y}{y}\right) = B_x E\left(\frac{\Delta x}{x}\right) + B_z E\left(\frac{\Delta z}{z}\right).$$

$$B_x = 2;$$

$$B_z = 3;$$

$$E\left(\frac{\Delta x}{x}\right) = +0,05;$$

$$E\left(\frac{\Delta z}{z}\right) = 0;$$

$$E\left(\frac{\Delta y}{y}\right) = 2 \cdot 0,05 + 3 \cdot 0 = 0,1;$$

$$\delta\left(\frac{\Delta y}{y}\right) = B_x \delta\left(\frac{\Delta x}{x}\right) + B_z \delta\left(\frac{\Delta z}{z}\right);$$

$$\delta\left(\frac{\Delta y}{y}\right) = 2 \cdot 0,05 + 3 \cdot 0,05 = 0,1 + 0,15 = 0,25;$$

$$Y_{\text{пред}} = 200 + 0,1 \pm 0,25 = 200^{+0,35}_{-0,15}.$$

Лекция 8

РАСЧЁТ РАЗМЕРНЫХ ЦЕПЕЙ

Понятие о размерных цепях

Под *размерной цепью* понимается совокупность расположенных по замкнутому контуру размеров, определяющих расположение поверхностей в детали или деталей в узле (рис. 8.1, 8.2).

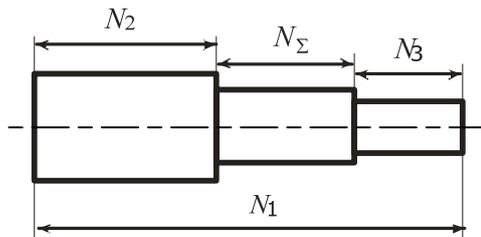


Рис. 8.1

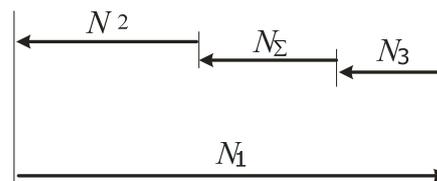


Рис. 8.2

Размер детали, который получается последним в процессе изготовления, называется *замыкающим* (N_{Σ}).

Размеры в размерной цепи называются *звеньями*.

Все звенья, входящие в размерную цепь, кроме замыкающего, называются *составляющими звеньями*. Они могут быть *увеличивающими* (N_1) и *уменьшающими* (N_2, N_3).

Уменьшающее звено – это звено, с увеличением которого уменьшается замыкающее звено.

Увеличивающее звено – это звено, с увеличением которого увеличивается замыкающее звено.

Размерные цепи могут быть *линейными* и *угловыми*. Если хотя бы одно звено расположено под углом, то это угловая размерная цепь.

Размерные цепи могут быть *независимыми* и *взаимосвязанными*. Если все звенья цепи не зависят от звеньев других размерных цепей, то такая размерная цепь называется *независимой*.

Метод предельных отклонений (максимума – минимума)

При этом методе расчёта используется уравнение погрешностей

$$\Delta N_{\Sigma} = \sum_{k=1}^N A_k \Delta N_k,$$
$$N = n + m,$$

где N – количество составляющих звеньев;

n – количество увеличивающих звеньев;

m – количество уменьшающих звеньев.

При расчёте линейных размерных цепей коэффициенты влияния $A_k = \pm 1$: «+» – для увеличивающих звеньев, «-» – для уменьшающих звеньев. Следовательно, можно написать

$$\Delta N_{\Sigma} = \sum_{i=1}^n \Delta N_i - \sum_{j=1}^m \Delta N_j. \quad (8.1)$$

С учётом этого

$$E(\Delta N_{\Sigma}) = \sum_{i=1}^n \vec{E}(\Delta N_i) - \sum_{j=1}^m \overleftarrow{E}(\Delta N_j). \quad (8.2)$$

Стрелка «→» означает, что звено увеличивающее, а стрелка «←» – звено уменьшающее. Половина поля допуска замыкающего звена

$$\delta(\Delta N_{\Sigma}) = \sum_{k=1}^{n+m} \delta(\Delta N_k).$$

Вероятностный метод

Используя формулу Бородачёва, можем записать

$$M(\Delta N) = E(\Delta N) + a\delta(\Delta N). \quad (8.3)$$

Применяя правило сложения математических ожиданий и выражение (8.1), получим

$$M(\Delta N_{\Sigma}) = \sum_{i=1}^n \vec{M}(\Delta N_i) + \sum_{j=1}^m \overleftarrow{M}(\Delta N_j).$$

Используя формулу (8.3), можем получить

$$E(\Delta N_{\Sigma}) = \sum_{i=1}^n \left[\vec{E}(\Delta N_i) + a_i \vec{\delta}(\Delta N_i) \right] - \sum_{j=1}^m \left[\overleftarrow{E}(\Delta N_j) + a_j \overleftarrow{\delta}(\Delta N_j) \right] - a_{\Sigma} \delta_{\Sigma}(\Delta N),$$

где a_i и a_j – коэффициенты относительной асимметрии.

Половина поля допуска имеет вид

$$\delta(N_{\Sigma}) = \frac{1}{K_{\Sigma}} \sqrt{\sum_{k=1}^{n+m} K_k^2 \delta_k^2(\Delta N)},$$

где K – коэффициент относительного рассеивания.

При расчётах коэффициенты асимметрии a_k и относительного рассеивания K_k берутся из справочной литературы.

Если a_k и K_k неизвестны, то рекомендуется брать $K_k = 1,2 - 1,3$ и $a_k = 0,1 - 0,2$.

Кроме того, $K_{\Sigma} = 1$, $a_{\Sigma} = 0$ в следующих случаях:

1) когда законы распределения погрешностей составляющих звеньев нормальные (гауссовские);

2) законы распределения одновершинные и число составляющих звеньев не менее 5;

3) законы распределения имеют несколько вершин, но не резко различаются друг от друга, и число составляющих звеньев не менее 8.

Если эти условия не выполняются, то расчёты K_{Σ} и a_{Σ} проводятся по эмпирическим формулам [6].

$$a_{\Sigma} = 0,59 \left[\sum_{i=1}^n a_i \overset{\rightarrow}{\delta}_i - \sum_{j=1}^m a_j \overset{\leftarrow}{\delta}_j \right] / \sum_{k=1}^{n+m} \delta_k ;$$

$$K_{\Sigma} = 1 + \frac{0,55}{\sum_{k=1}^{n+m} \delta_k} \left[\sqrt{\sum_{k=1}^{n+m} K_k^2 \delta_k^2} - \sqrt{\sum_{k=1}^{n+m} \delta_k^2} \right].$$

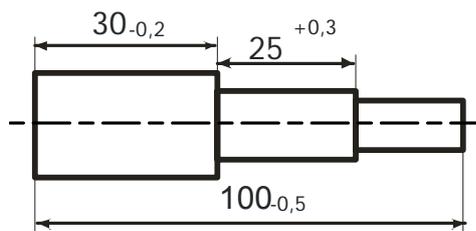


Рис. 8.3

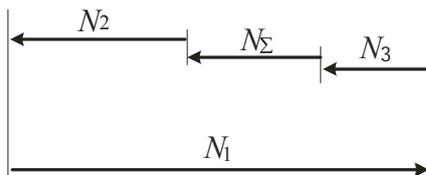


Рис. 8.4

Пример. Определить предельные отклонения замыкающего звена детали, показанной на рис. 8.3, методом предельных отклонений и вероятностным методом ($a_i = 0$; $K_i = 1$).

Составим размерную цепь (рис. 8.4).

$$N_{\Sigma \text{пред}} = N_n + E_{\Sigma}(\Delta N) = \delta_{\Sigma}(\Delta N).$$

Определяем:

$$N_n = 100 - 30 - 25 = 45.$$

Согласно методу предельных отклонений

$$E_{\Sigma} = \sum_{i=1}^n \vec{E}_i - \sum_{j=1}^m \overset{\leftarrow}{E}_j = E_1 - (E_2 + E_3) = -0,25 - (0,15 - 0,1) = -0,3;$$

$$\delta_{\Sigma} = \sum_{k=1}^{n+m} \delta_k = \delta_1 + \delta_2 + \delta_3 = 0,25 + 0,15 + 0,1 = 0,5;$$

$$N_{\Sigma \text{пред}} = 45 - 0,3 \pm 0,5 = 45_{-0,8}^{+0,2}.$$

Применяя вероятностный метод, примем $a_i = a_j = 0$;
 $K_i = K_j = 1$.

E_Σ определяется, как и в методе предельных отклонений, поэтому $E_\Sigma = -0,3$.

$$\delta_\Sigma = \sqrt{\sum_{k=1}^{n+m} \delta_k^2} = \sqrt{\sum_{k=1}^3 \delta_k^2} = \sqrt{\delta_1^2 + \delta_2^2 + \delta_3^2} = 0,3.$$

Различия полученных результатов объясняются следующим: при методе предельных отклонений мы считаем, что размеры деталей имеют предельные отклонения, что не соответствует действительности.

Метод предельных отклонений применяется при расчёте точности единичных или особо важных (взрывоопасных) изделий. Этот метод гарантирует 100%-е нахождение допуска замыкающего звена в рассчитанных пределах. При массовом производстве целесообразно применять вероятностный метод расчёта. Необходимо учитывать, что он гарантирует нахождение погрешности выходного звена в пределах расчётного допуска с вероятностью, равной 0,9973.

Лекция 9

МЕТОДЫ МАТРИЧНЫХ ИСПЫТАНИЙ И СТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ

Метод матричных испытаний

Метод матричных испытаний – метод физического моделирования. Он позволяет определить рабочую область радиоустройства при всех возможных значениях *первичных параметров*, находящихся в пределах допусков, оптимизировать его работоспособность.

Первичные параметры можно разбить на следующие группы:

- параметры деталей, из которых состоит устройство;
- параметры источников питания;
- параметры входных сигналов;
- параметры нагрузки;
- параметры, характеризующие условия окружающей среды.

Суть метода матричных испытаний и задачи оценки и оптимизации работоспособности по результатам этих испытаний поясним на примере устройства, работоспособность которого существенным образом зависит только от двух первичных параметров X_1 и X_2 . Такие параметры будем называть *определяющими первичными параметрами*.

Пусть диапазон возможного изменения этих параметров задан интервалами $[x_{1\min}, x_{1\max}]$ и $[x_{2\min}, x_{2\max}]$. Разобьем интервалы на $l^{(1)}$ и $l^{(2)}$ равных квантов соответственно (рис. 9.1).

В качестве *представителей квантов* выберем значения параметров, соответствующие серединам квантов. Если устройство оказывается неработоспособным (работоспособным) при значениях параметров, соответствующих представителю кванта, то будем считать, что оно не работает (работает) при всех значениях параметров, лежащих в этом кванте.

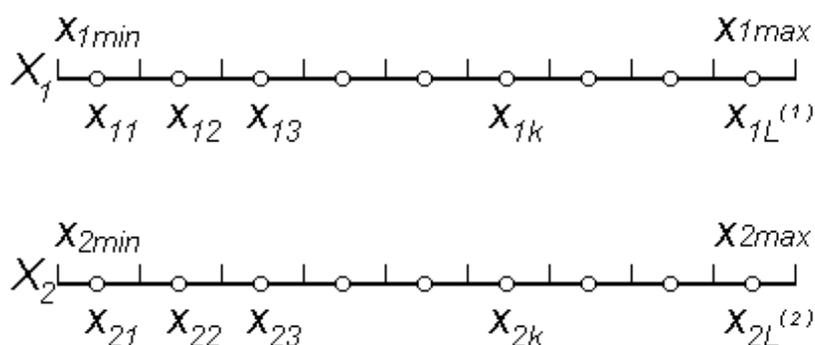


Рис. 9.1

Введем понятие *ситуации*, под которым будем понимать такое состояние устройства, когда каждый из его двух определяющих параметров (двумерный случай) принимает значение, соответствующее представителю определенного кванта. Определение легко распространить на n -мерный случай. Число всех возможных ситуаций устройства в двумерном случае, очевидно, равно

$$N_2 = l^{(1)}l^{(2)}.$$

Индекс 2 при N означает, что рассматривается двумерный случай, т.е. имеются два определяющих параметра. В общем случае при n определяющих параметрах число возможных ситуаций равно

$$N = \prod_{i=1}^n l^{(i)}.$$

Перебрав в какой-либо последовательности все несовместные ситуации устройства, составим матрицу ситуаций. Для двумерного случая она будет иметь вид

$$\|\alpha_{n_2}\| = \begin{vmatrix} x_{11} & x_{21} \\ x_{11} & x_{22} \\ \dots & \dots \\ x_{11} & x_{2l^{(2)}} \\ \dots & \dots \\ x_{12} & x_{21} \\ x_{12} & x_{22} \\ \dots & \dots \\ x_{1l^{(1)}} & x_{2l^{(2)}} \end{vmatrix}.$$

В матрице вторая строка $x_{11} x_{22}$, например, означает, что параметр X_1 принял значение в первом, а параметр X_2 – во втором кванте.

Упорядоченную таким образом последовательность всех возможных ситуаций устройства обозначим $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_{N_2}$.

Среди N_2 возможных ситуаций устройства при матричных испытаниях будет обнаружено некоторое количество Q отказных ситуаций с точки зрения заданного *критерия отказа* для рассматриваемого устройства. Отказные ситуации будем отмечать чертой сверху. Они могут быть определены натурным физическим моделированием, поскольку такое моделирование позволяет довольно точно воспроизвести работу устройства и автоматически учесть все факторы, влияющие на его работоспособность, в том числе и второстепенные, которые при всех других видах моделирования, как правило, не учитываются.

Результаты матричных испытаний позволяют найти область работоспособности устройства (рис. 9.2), определить вероятность нахождения выходного параметра в пределах поля допуска и оптимизировать работоспособность.

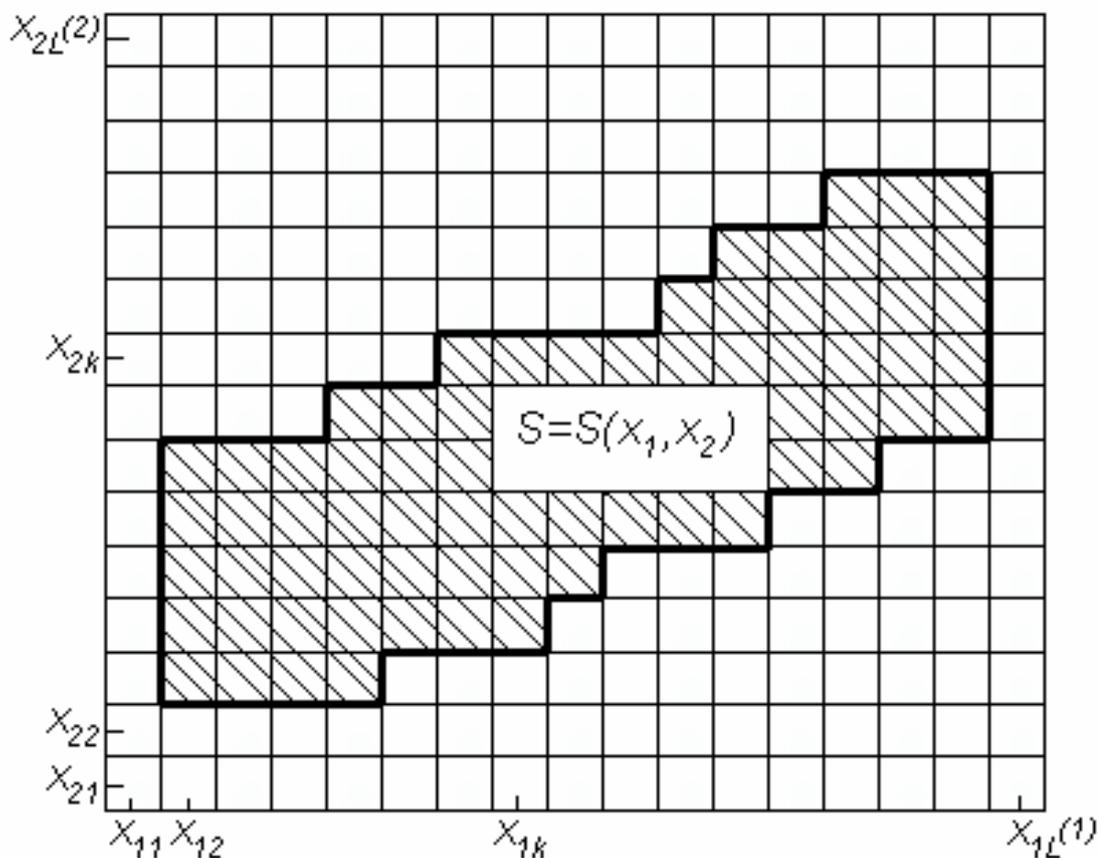


Рис. 9.2

Вероятность того, что схема окажется неработоспособной, выразится так:

$$P = \sum_{i=1}^Q P(\bar{\alpha}_i),$$

где Q – число отказных ситуаций.

Вероятность отказных ситуаций, если определяющие первичные параметры независимы, вычисляют как произведение вероятностей нахождения параметров в соответствующих квантах. Например, для ситуации α_1

$$P(\alpha_1) = P(x_{11})P(x_{21}).$$

Вероятность нахождения параметров X_1 и X_2 в квантах x_{11} и x_{21} можно вычислить, если известны законы распределения $f(x)$ этих параметров, по выражению

$$P(\alpha < X \leq \beta) = \int_{\alpha}^{\beta} f(x) dx.$$

Вероятность будет численно равна площади фигуры под кривой распределения, построенной на основании, равном длине кванта (рис. 9.3). При нормальном законе распределения эти вероятности могут быть найдены по формуле

$$P(\alpha < X \leq \beta) = \Phi\left(\frac{\beta - m}{\sqrt{D}}\right) - \Phi\left(\frac{\alpha - m}{\sqrt{D}}\right),$$

где $\Phi(z)$ – функция Лапласа (из таблицы в [2]); α, β – границы кванта; m, D – математическое ожидание и дисперсия соответственно.

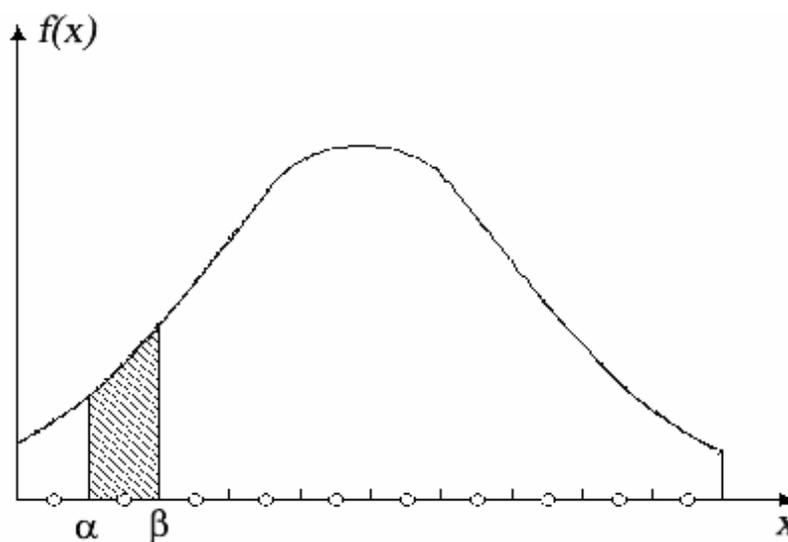


Рис. 9.3

Для сравнительной оценки вероятность работоспособности исследуемого устройства может быть найдена по формуле

$$P = 1 - \frac{Q}{N}.$$

Вероятность здесь будет несколько заниженной из-за принятия веса каждой отказной ситуации равным единице. Если в некоторый фиксированный момент времени t известна плотность распределения параметров X_1 и X_2 , то задача оптимизации работоспособности иссле-

дуемого устройства для этого момента времени геометрически может рассматриваться как задача отыскания такого положения функции $f(x_1, x_2)$, при котором объем, вырезаемый из $f(x_1, x_2)$, восстановленный цилиндром с основанием $S(x_1, x_2)$, был бы максимальным. Аналитически это записывается в виде

$$P_{\max} = \max_{S(x_1; x_2)} \iint f(x_1 - m_{x_1}; x_2 - m_{x_2}) dx_1 dx_2,$$

где m_{x_1}, m_{x_2} – математическое ожидание функции $f(x_1, x_2)$; P_{\max} – максимальная вероятность нахождения выходного параметра в пределах поля допуска.

Практически задача оптимизации может быть сведена к расчету новых номинальных значений определяющих первичных параметров.

Эти значения могут быть установлены простой процедурой нахождения координат центра тяжести n -мерной области работоспособности (в рассматриваемом случае – двумерной области) исследуемого устройства. Координату центра тяжести по j -му параметру x_{jl} находим по формуле

$$x_{jl} = x_{j1} + \frac{N(l^{(j)} - 1)\Delta x_j}{2(N - Q)} \left[1 - \frac{2}{N(l^{(j)} - 1)} \sum_{\chi=1}^{l^{(j)}-1} \chi Q_{j,\chi+1} \right],$$

где N – общее количество ситуаций; Q – количество отказных ситуаций; $Q_{j, \chi+1}$ – количество отказных ситуаций, соответствующих $\chi+1$ кванту.

Достоинства метода матричных испытаний – возможность количественной оценки работоспособности. Недостатки – значительная трудоемкость и сложность использования для устройств, в которых моделирование первичных параметров затруднено.

Метод статистического моделирования

Будем считать, что известна функциональная зависимость

$$N = f(X_1, X_2, \dots, X_n), \quad (9.1)$$

где N – выходной параметр изделия; X_1, X_2, \dots, X_n – начальные параметры, влияющие на N .

Под начальными параметрами понимаются параметры элементов изделия, внешние воздействия, паразитные связи и т. п. Зависимость (9.1) может быть задана аналитически или графически. Под-

ставляя в эту зависимость по одному случайному значению x_i , получим одно значение N_i . Если многократно повторять расчёт, пользуясь случайными значениями X_1, X_2, \dots, X_n , то получим множество значений N . Очевидно, если начальные параметры выбираются в соответствии с теми законами распределения, которые их характеризуют, то распределение параметра N соответствует искомой плотности распределения.

Согласно теореме Бернулли частота $p^* = \frac{m}{n}$ попадания параметра N в пределы поля допуска α, β при неограниченном увеличении числа опытов (расчётов) приблизится к вероятности $P(\alpha \leq N < \beta)$. Здесь m – количество попаданий параметра N в пределы от α до β , а n – общее количество испытаний. Поэтому при достаточно большом количестве расчётов можно считать, что $p^* \approx P(\alpha \leq N < \beta)$.

Алгоритм вычисления вероятности на ЭВМ следующий:

1. Параметрам X_1, X_2, \dots, X_n задаются случайные значения согласно их законам распределения $f(x_1), f(x_2), \dots, f(x_n)$.
2. В соответствии с функциональной зависимостью (9.1) определяется значение N_i .
3. Найденное значение N_i сравнивается с граничными значениями интервала (α, β) ; выполнение неравенства $\alpha \leq N_i < \beta$ отмечается знаком $\omega = 1$, невыполнение – знаком $\omega = 0$.
4. Полученное значение признака ω прибавляется к количеству случайных чисел, попавших внутрь интервала (α, β) .
5. К содержанию счётчика количества испытаний прибавляется единица.
6. Управление передаётся снова первой операции.
7. После проведения n таких испытаний вычисляется значение частоты $\frac{m}{n}$.

При решении задачи методом статистических испытаний большое значение имеет оценка необходимого объёма испытаний для заданной погрешности ε или же оценка ε по известному числу испытаний n . В тех случаях, когда мы имеем дело с обработкой статистического материала, нельзя ограничиваться лишь только заданной точностью решения, необходимо ещё задавать так называемую степень надёжности (достоверности) в виде доверительной вероятности γ .

Моделирование непрерывных случайных величин. Решить эту

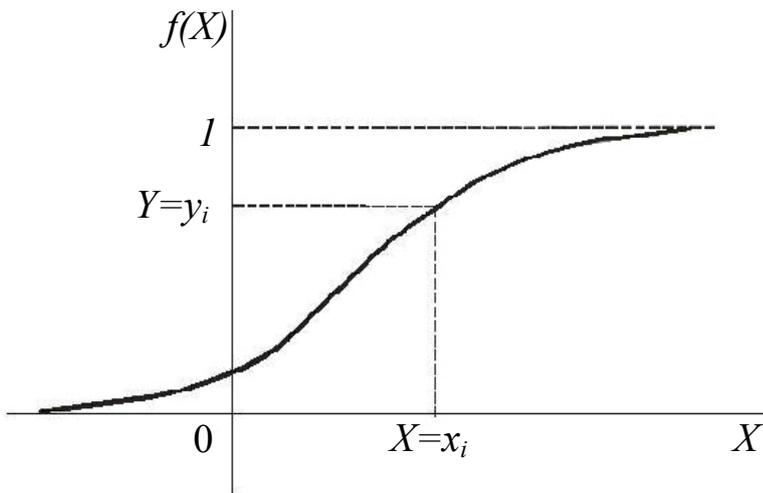


Рис. 9.4

задачу можно различными способами. Один из них основан на известной теореме: если случайная величина имеет плотность распределения $f(x)$, то распределение случайной величины $Y = F(x_i)$ (рис. 9.4) может интерпретироваться как последовательность случайных чисел, равномерно распределённых в

интервале $[0, 1]$. Следовательно, справедлив и обратный переход: если мы располагаем совокупностью равномерно распределённых в интервале $[0, 1]$ случайных чисел y_i , то каждому из них соответствует определённое для данного вида функции $F(x)$ число x_i , значение которого находится обратным преобразованием $x_i = F^{-1}(y_i)$. Случайная величина X , полученная таким способом, имеет функцию распределения $F(x)$.

Отсюда следует, что процесс получения последовательности случайных чисел x_i с заданным законом распределения сводится к решению относительно x_i уравнения

$$F(x_i) = \int_{-\infty}^{x_i} f(x) dx. \quad (9.2)$$

Посмотрим, как делается такое преобразование для нескольких законов распределения, широко применяемых на практике.

Э к с п о н е н ц и а л ь н ы й з а к о н. Функция распределения при экспоненциальном законе определяется выражением

$$F(x_i) = 1 - e^{-\lambda x_i}.$$

Учитывая (9.2), можно записать

$$y_i = F(x_i) = 1 - e^{-\lambda x_i}.$$

Решив это уравнение относительно x_i , получим

$$x_i = -\frac{1}{\lambda} \ln(1 - y_i). \quad (9.3)$$

Р а в н о м е р н ы й з а к о н. Функция распределения случайной величины X , равномерно распределённой в интервале $[a, b]$, определяется выражением

$$F(x) = \frac{x - a}{b - a}.$$

Учитывая, что $y_i = F(x_i)$, из последнего выражения получим

$$x_i = a + y_i (b - a). \quad (9.4)$$

З а к о н Р е л е я. Для закона Релея функция распределения может быть записана в виде

$$F(x) = 1 - e^{-\frac{x^2}{2\sigma^2}}.$$

Приравнивая $y_i = F(x_i)$, из последнего выражения найдём

$$x_i = \sqrt{-2\sigma^2 \ln y_i}. \quad (9.5)$$

З а к о н В е й б у л л а. Функцию распределения случайной величины, имеющей распределение Вейбулла, можно записать в виде

$$F(x) = 1 - e^{-\frac{x^b}{a}}.$$

Как и в предыдущих случаях, учитывая (9.2), из последнего выражения получим

$$x_i = [-a \ln(1 - y_i)]^{\frac{1}{b}}.$$

К сожалению, не всегда существуют элементарные преобразования для получения случайных чисел с заданным законом распределения из равномерно распределённых случайных чисел. В частности, у случайных величин с нормальным распределением функции обратная функция распределения не выражается в замкнутом виде через элементарные функции. В этих случаях для формирования случайных величин с заданным законом распределения применяются приближённые методы.

Рассмотрим один из таких методов, который используется для моделирования случайной величины X , распределённой по нормальному закону с параметрами $M[x]$ и $\sigma[x]$. Известно [2], что такую величину можно представить в виде

$$X = M[x] + \sigma[x]Z, \quad (9.6)$$

где Z – случайная величина, имеющая нормальный закон распределения с параметрами $m = 0$ и $\sigma = 1$.

Воспользуемся центральной предельной теоремой теории вероятностей: если Z_1, Z_2, \dots, Z_n – независимые случайные величины, имеющие один и тот же закон распределения с математическим ожи-

данием m и дисперсией σ^2 , то при неограниченном увеличении n закон распределения суммы неограниченно приближается к нормальному:

$$Z_{\Sigma} = \sum_{i=1}^n Z_i. \quad (9.7)$$

Математическое ожидание m_{Σ} и дисперсия σ_{Σ}^2 суммарной величины будут равны

$$m_{\Sigma} = nm, \quad \sigma_{\Sigma}^2 = n\sigma^2.$$

Если суммируемые величины распределены равномерно на интервале $[0, 1]$, то есть имеют $m = 0,5$ и $\sigma = \frac{1}{2}\sqrt{3}$, то соответствующие параметры для суммарной величины можно найти по формулам

$$m_{\Sigma} = \frac{n}{2}, \quad \sigma_{\Sigma} = \sqrt{\frac{n}{12}}. \quad (9.8)$$

Согласно выражениям (9.6) и (9.7) можно получить

$$Z = \frac{\sum_{i=1}^n Z_i - m_{\Sigma}}{\sigma_{\Sigma}}.$$

С учётом (9.8) последнее выражение представим в виде

$$Z = \sqrt{\frac{12}{n}} \left(\sum_{i=1}^n Z_i - \frac{n}{2} \right)$$

и, подставив его в (9.6), получим

$$X = M[x] + \sigma[x] \sqrt{\frac{12}{n}} \left(\sum_{i=1}^n Z_i - \frac{n}{2} \right). \quad (9.9)$$

Формула (9.8) связывает последовательность случайных чисел X , распределённых по нормальному закону с параметрами $M[x]$ и $\sigma[x]$, с последовательностью случайных чисел Z_i , распределённых по равномерному закону в интервале $[0, 1]$.

Точность моделирования X возрастает с увеличением n . На практике обычно бывает достаточно $n = 4 \dots 12$. Если $n = 12$, то формула (9.8) принимает более простой вид

$$X = M[x] + \sigma[x] \left(\sum_{i=1}^{12} Z_i - 6 \right). \quad (9.10)$$

В некоторых случаях необходимо иметь случайные числа X_i , распределённые по нормальному закону с параметрами $m = 0$ и $\sigma = 1$. Их можно получить, используя выражение

$$X_i = \sum_{i=1}^{12} Z_i - 6, \quad (9.11)$$

полученное из формулы (9.9).

Определение погрешности ε . Погрешность оценки вероятности p по частоте p^* находится по выражению

$$\varepsilon = t_\gamma \sqrt{\frac{p^*(1-p^*)}{n}},$$

где t_γ – квантиль распределения Стьюдента, определяемый из таблиц [11] по величине доверительной вероятности γ и числу степеней свободы $f=n-1$.

ОСНОВЫ ТЕОРИИ НАДЕЖНОСТИ ЭЛЕКТРОННЫХ СРЕДСТВ

Лекция 10

ВВЕДЕНИЕ

Причины возникновения проблемы надежности

Можно выделить пять причин возникновения проблемы надежности, которые привели к быстрому развитию теории надежности в 40 – 50-е годы прошлого столетия. К ним относятся:

- 1) усложнение ЭС и систем автоматического управления;
- 2) отставание качества и надежности ЭС от их количественного применения;
- 3) усложнение условий эксплуатации ЭС;
- 4) возрастание цены отказа ЭС;
- 5) невозможность восстановления после отказа во время работы в ряде случаев (атомные реакторы, агрессивные среды и т.д.).

Методы повышения, обеспечения и сохранения надежности

Вопросами надежности ЭС необходимо заниматься на всех этапах «жизни» изделий: при проектировании, производстве и эксплуатации.

При проектировании – это выбор надежных элементов, облегчение режимов работы элементов, резервирование, современные методы проектирования изделий и др.

При производстве – применение современных технологий, более качественных материалов, современных методов контроля и испытаний и др.

При эксплуатации – повышение квалификации персонала и своевременное проведение профилактических работ, обеспечение запасными частями и др.

Основные понятия и определения в теории надежности

В соответствии с ГОСТ 27.002-89 «Надежность в технике. Термины и определения» *надежность* – свойство объекта сохранять во времени в установленных пределах значения всех параметров, характеризующих способность выполнять требуемые функции в заданных режимах и условиях применения, технического обслуживания, ремонта, хранения и транспортирования.

Надежность определяется безотказностью, долговечностью, ремонтпригодностью, сохраняемостью.

Безотказность – свойство объекта непрерывно сохранять работоспособное состояние в течение некоторого времени или некоторой наработки.

Долговечность – свойство объекта сохранять работоспособность до наступления предельного состояния при установленной системе технического обслуживания и ремонта.

Ремонтпригодность – свойство объекта, заключающееся в приспособленности к предупреждению и обнаружению причин возникновения отказов, повреждений, поддержанию и восстановлению работоспособного состояния путем проведения технического обслуживания и ремонтов.

Сохраняемость – свойство объекта сохранять значения показателей безотказности, долговечности и ремонтпригодности в течение и после хранения и (или) транспортирования.

Исправное состояние – состояние объекта, при котором он соответствует всем требованиям нормативно-технической и (или) конструкторской документации.

Работоспособное состояние – состояние объекта, при котором значения всех параметров, характеризующих способность выполнять заданные функции, соответствуют требованиям нормативно-технической и (или) конструкторской документации. Например, если повреждено покрытие корпуса прибора, то он является работоспособным, но неисправным.

Отказ – событие, заключающееся в нарушении работоспособного состояния объекта.

Восстанавливаемый объект – объект, для которого в рассматриваемой ситуации проведение восстановления работоспособного состояния предусмотрено в нормативно-технической и (или) конструкторской документации.

Ремонтируемый объект – объект, для которого проведение ремонтов предусмотрено в нормативно-технической и (или) конструкторской документации.

Единичный показатель надежности – показатель надежности, характеризующий одно из свойств, составляющих надежность объекта.

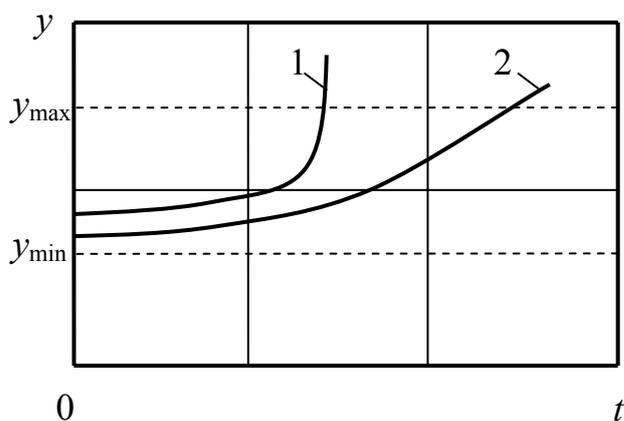
Комплексный показатель надежности – показатель надежности, характеризующий несколько свойств, составляющих надежность объекта.

Независимый отказ – отказ объекта, не обусловленный отказом другого объекта.

Зависимый отказ – отказ объекта, обусловленный отказом другого объекта. Например, перегорание резистора часто обусловлено пробоем конденсатора.

Внезапный отказ – отказ, характеризующийся скачкообразным изменением значений одного или нескольких заданных параметров объекта (см. рисунок, 1). Например, пробой конденсатора приводит к перегоранию резистора.

Постепенный отказ – отказ, характеризующийся постепенным изменением значений одного или нескольких заданных параметров объекта (см. рисунок, 2). Обычно постепенный отказ происходит в результате старения или износа материалов.



Перебегающий отказ – многократно возникающий, самоустраняющийся отказ объекта одного и того же характера.

Конструкционный отказ – отказ, возникающий в результате несовершенства или нарушения установленных правил и (или) норм конструирования объекта.

Производственный отказ – отказ, возникающий в результате несовершенства или нарушения установленного процесса изготовления или ремонта объекта, выполнявшегося на ремонтном предприятии.

Эксплуатационный отказ – отказ, возникший в результате нарушения установленных правил и (или) условий эксплуатации объекта.

Лекция 11

ПОКАЗАТЕЛИ НАДЕЖНОСТИ

Показатели, характеризующие одну из сторон надежности, называются единичными, две или более – комплексными.

Показатели надежности должны учитывать максимальное число факторов, влияющих на надежность, быть удобными для записи в техническую документацию и давать возможность их экспериментального определения.

Показатели надежности неремонтируемых изделий

К ним относятся вероятность безотказной работы, средняя наработка до отказа, интенсивность отказов и другие.

Вероятность безотказной работы $P(t)$ – это вероятность того, что в пределах заданного времени или заданной наработки не произойдет отказа изделия.

В соответствии с определением

$$P(t) = P(T \geq t), \quad (11.1)$$

где T – время безотказной работы, являющееся случайной величиной, t – заданное время.

Если записать $P(t) = P(t \leq T < \infty)$ и вспомнить из теории вероятностей, что $P(\alpha \leq X < \beta) = \int_{\alpha}^{\beta} f(x) dx$, то можем записать

$$P(t) = \int_t^{\infty} f(\tau) d\tau \quad (11.2)$$

Вероятность безотказной работы имеет следующие свойства (рис. 11.1):

1) $P(t) = 1$ при $t = 0$;

2) $P(t) = 0$ при $t = +\infty$;

3) если $t_1 > t_2$, то $P(t_1) \leq P(t_2)$, то есть вероятность безотказной работы – невозрастающая функция своего аргумента.

Экспериментально $P(t)$ можно оценивать по выражению

$$P^*(t) = \frac{N(t)}{N}, \quad (11.3)$$

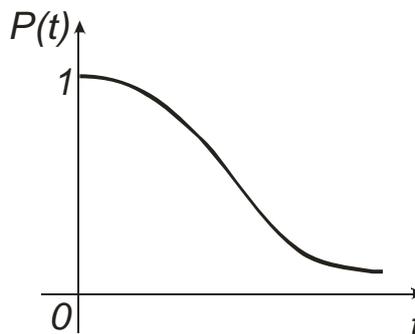


Рис. 11.1

где N – число испытываемых изделий; $N(t)$ – число работоспособных изделий к моменту времени t .

Вместо $P(t)$ часто используется *вероятность отказа* $Q(t)$, то есть вероятность того, что в пределах заданного времени (наработки) произойдет отказ изделия.

Можно записать

$$Q(t) = P(T < t) = P(0 < T < t), \quad (11.4)$$

а следовательно,

$$Q(t) = \int_0^t f(\tau) d\tau. \quad (11.5)$$

Складывая $P(t)$ и $Q(t)$, получим

$$P(t) + Q(t) = \int_0^t f(t) dt + \int_t^{\infty} f(t) dt = 1$$

или

$$P(t) + Q(t) = 1. \quad (11.6)$$

Экспериментально $Q(t)$ оценивается по выражению

$$Q^*(t) = \frac{n(t)}{N}, \text{ где } n(t) \text{ – число отказавших изделий за время } t.$$

Из выражения (11.4) следует, что вероятность отказа равна функции распределения времени безотказной работы изделия, то есть $Q(t) = F(t)$.

Средняя наработка до отказа t_{cp} – это математическое ожидание наработки изделия до первого отказа, то есть

$$t_{cp} = \int_0^{\infty} t f(t) dt. \quad (11.7)$$

Учитывая, что

$$f(t) = F'(t) = Q'(t) = [1 - P(t)]' = -P'(t) = -\frac{dP(t)}{dt}, \quad (11.8)$$

и подставляя это выражение в (11.7), получим

$$t_{cp} = - \int_0^{\infty} t dP(t).$$

Используя правило интегрирования по частям

$$\int u dv = uv - \int v du,$$

получим

$$t_{cp} = -tP(t)|_0^{\infty} + \int_0^{\infty} P(t) dt.$$

При увеличении t вероятность $P(t)$ значительно быстрее стремится к нулю, чем t – к бесконечности. Поэтому можно считать, что первое слагаемое равно нулю и

$$t_{cp} = \int_0^{\infty} P(t) dt. \quad (11.9)$$

Очевидно, что численно $P(t)$ равняется площади под кривой $P(t)=\varphi(t)$, показанной на рис. 11.1.

Статистически $t_{cp} = \frac{\sum_{i=1}^N t_i}{N}$, где t_i – время безотказной работы i -го изделия, N – число изделий, поставленных на испытания.

Интенсивность отказов $\lambda(t)$ – условная плотность вероятности возникновения отказа невозстанавливаемого объекта, определяемая для рассматриваемого момента времени при условии, что до этого момента отказ не возник. Статистически определяется по формуле

$$\lambda^*(t) = \frac{n(\Delta t)}{N(t)\Delta t}, \quad (11.10)$$

где $n(\Delta t)$ – число изделий, отказавших на интервале времени Δt ; $N(t)$ – число изделий, не отказывавших к моменту t (рис. 11.2).

Вероятность отказа изделий за время Δt находится по формуле

$$Q^*(\Delta t) = \frac{n(\Delta t)}{N(t)}.$$

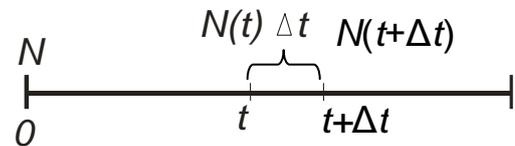


Рис. 11.2

Поэтому интенсивность можно трактовать как плотность вероятности отказов в интервале времени Δt .

Учитывая, что $n(\Delta t) = N(t) - N(t + \Delta t)$,

$$P^*(t) = \frac{N(t)}{N}$$

и

$$P^*(t + \Delta t) = \frac{N(t + \Delta t)}{N},$$

получим

$$n(\Delta t) = N[P^*(t) - P^*(t + \Delta t)].$$

Подставляя последнее выражение в формулу (11.10), найдем

$$\lambda^*(t) = \frac{P^*(t) - P^*(t + \Delta t)}{P^*(t)\Delta t}.$$

Интенсивность отказов

$$\lambda(t) = \lim_{\substack{\Delta t \rightarrow 0 \\ n \rightarrow \infty}} \lambda^*(t) = -\frac{1}{P(t)} \lim_{\substack{\Delta t \rightarrow 0 \\ n \rightarrow \infty}} \frac{P^*(t + \Delta t) - P^*(t)}{\Delta t}.$$

Следовательно,

$$\lambda(t) = -\frac{P'(t)}{P(t)}. \quad (11.11)$$

Учитывая (11.8), получим

$$\lambda(t) = \frac{f(t)}{P(t)}. \quad (11.12)$$

Логарифмируя и затем интегрируя полученное выражение в пределах от 0 до t , найдем

$$P(t) = e^{-\int_0^t \lambda(\tau) d\tau}. \quad (11.13)$$

Это одна из основных формул теории надежности.

На рис. 11.3 показана типичная зависимость $\lambda(t)$.

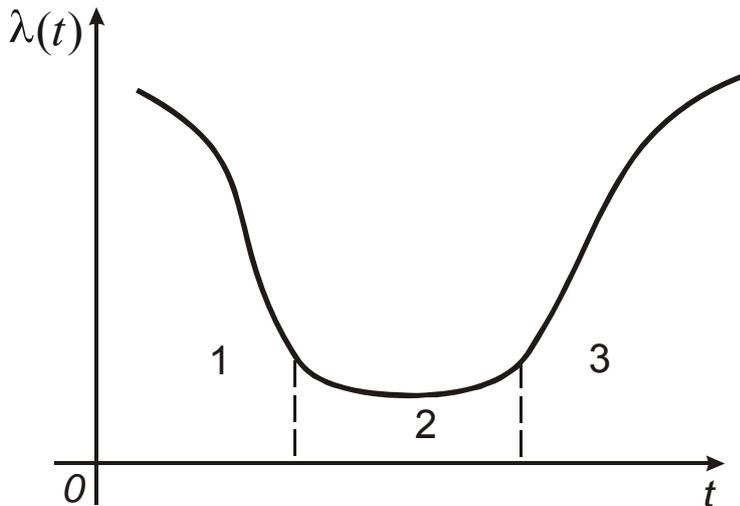


Рис. 11.3

Видно, что можно выделить три периода:
1 – период приработки изделия, характеризующийся высоким значением $\lambda(t)$, уменьшающимся к концу изгиба. Такой характер зависимости объясняется наличием скрытых дефектов, не выясненных при контроле, поэтому период приработки стремятся проводить на предприятии, «выжигая» элементы со скрытыми дефектами (электротренировка и т.д.);

2 – период нормальной эксплуатации изделия, характеризующийся наиболее низкой интенсивностью отказов, $\lambda \approx \text{const}$;

3 – период старения и износа, идет снижение надежности, увеличение интенсивности отказов.

Интенсивность отказов – основная характеристика надежности ЭС.

Лекция 12

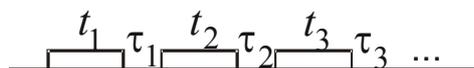
ПОКАЗАТЕЛИ НАДЕЖНОСТИ РЕМОНТИРУЕМЫХ ИЗДЕЛИЙ

К ним относятся: вероятность безотказной работы, средняя наработка на отказ (СНО), параметр потока отказов (ППО), вероятность восстановления и др.

Средняя наработка на отказ статистически определяется по формуле

$$T^* = \frac{\sum_{i=1}^n t_i}{n},$$

где t_i – время безотказной работы; n – число циклов безотказной работы (см. рисунок).



Параметр потока отказов статистически определяется по формуле

$$\omega^*(t) = \frac{n(\Delta t)}{N\Delta t},$$

где N – число изделий, поставленных на испытания. При определении параметра потока отказов отказавшие изделия заменяются новыми.

Вероятность восстановления $P_B(t) = P(t_B < t)$ – это функция распределения случайной величины t_B – времени восстановления изделия.

Среднее время восстановления статистически определяется по формуле

$$\bar{t}_B^* = \frac{\sum_{i=1}^n \tau_i}{n},$$

где τ_i – время восстановления изделия после i -го отказа (см. рисунок); n – общее число отказов.

Комплексные показатели надежности

Коэффициент готовности K_T экспериментально определяется по формуле

$$K_{\Gamma}^* = \frac{\sum_{i=1}^n t_i}{\sum_{i=1}^n t_i + \sum_{i=1}^n \tau_i},$$

где t_i – время безотказной работы; τ_i – время восстановления после i -го отказа.

Если поделить числитель и знаменатель на n (число отказов), то получим

$$K_{\Gamma}^* = \frac{\sum_{i=1}^n \frac{t_i}{n}}{\sum_{i=1}^n \frac{t_i}{n} + \sum_{i=1}^n \frac{\tau_i}{n}}$$

или

$$K_{\Gamma}^* = \frac{T^*}{T^* + t_{\text{в}}^*}.$$

При n больше 50 – 100 K_{Γ} приближается к

$$K_{\Gamma} = \frac{T}{T + t_{\text{в}}}.$$

Видно, что K_{Γ} – вероятность работоспособного состояния изделия в любой момент времени.

Часто применяется коэффициент оперативной готовности

$$K_{\text{ог}} = K_{\Gamma} P(t).$$

Коэффициент технического использования

$$K_{\text{т.и}}^* = \frac{\sum_{i=1}^n t_i}{\sum_{i=1}^n t_i + \sum_{i=1}^n \tau_i + \sum_{j=1}^m t_{\text{то}j}},$$

где $t_{\text{то}j}$ – время технического обслуживания; m – количество технических обслуживаний изделия.

Надежность типовых элементов

Надежность типовых элементов характеризуется интенсивностью отказов. Например, интенсивность отказов $\lambda_{i0} = 5 \cdot 10^{-6}$ 1/ч. Эта интенсивность относится к нормальным условиям эксплуатации. Усложнение условий эксплуатации приводит к снижению надежности радиоаппаратуры. Для учета дестабилизирующих факторов (повы-

шенная электрическая нагрузка, повышенная или пониженная температура, вибрации и т.д.) вводятся поправочные коэффициенты.

Электрическая нагрузка характеризуется коэффициентом нагрузки K_n – отношением значения некоторого параметра, характеризующего работу элемента в реальном режиме, к номинальному значению этого параметра, установленному ТУ.

Интенсивность отказов ЭРЭ при отклонении режимов их работы от нормальных можно представить в виде

$$\lambda(v) = \lambda_{i0} a_1 a_2 \dots a_n,$$

где a_1, a_2, \dots, a_n – коэффициенты, учитывающие эти отклонения, а λ_{i0} – интенсивность отказов при нормальных условиях ($K_n = 1$, $t = +25$ °С, механические нагрузки отсутствуют, относительная влажность 65 %).

Значения λ_{i0} берутся из таблиц, а коэффициенты a_i – из таблиц или графиков [5].

Лекция 13

ЗАКОНЫ РАСПРЕДЕЛЕНИЯ ВРЕМЕНИ БЕЗОТКАЗНОЙ РАБОТЫ ЭС

Знание этих законов распределения $f(t)$ необходимо, так как все основные показатели безотказности являются функцией этих законов, например $P(t) = \int_t^{\infty} f(t)dt$, $t_{cp} = \int_0^{\infty} tf(t)dt$. Они важны также при создании экономичных методов испытаний на надежность.

Потоки отказов

Под потоком отказов понимают такую последовательность отказов, которые происходят один за другим в случайные моменты времени (рис. 13.1).

Наибольшее применение находит простейший поток отказов.

Простейший поток отказов – это поток, который обладает свойствами стационарности, ординарности и отсутствия последствия.



Рис. 13.1

Поток стационарный, если вероятность некоторого количества отказов в интервале времени Δt зависит от длины этого интервала и не зависит от положения его на оси времени.

У потока *отсутствует последствие*, если вероятность появления некоторого количества отказов в интервале времени Δt не зависит от того, сколько отказов произошло до начала этого интервала времени.

Поток ординарный, если вероятность появления одного отказа в малом интервале времени Δt значительно превышает вероятность появления двух или большего количества отказов.

Поток отказов сложных электронных схем в период их нормальной эксплуатации соответствует простейшему потоку. Простейший поток отказов хорошо изучен в теории вероятностей и для него получена формула, называемая законом Пуассона:

$$P_m(t) = \frac{a^m}{m!} e^{-a},$$

где m – количество отказов; t – время; a – среднее количество отказов за время t , равное $a = \lambda t$, λ – интенсивность отказов (среднее количе-

ство отказов в единицу времени). Простейший поток часто называют пуассоновским.

Можем записать

$$P_m(t) = \frac{(\lambda t)^m}{m!} e^{-\lambda t}.$$

Если, например, $m = 3$, то

$$P_3(t) = \frac{(\lambda t)^3}{6} e^{-\lambda t},$$

а если $m > 1$, то

$$P_{m>1}(t) = P_2(t) + P_3(t) + \dots$$

Пример. На испытания поставлено 100 изделий. Необходимо определить вероятность того, что откажет не менее двух изделий.

Можем записать

$$P_{m \geq 2}(t) = P_2(t) + P_3(t) + \dots + P_{100}(t),$$

но учитывая, что $\sum_{i=1}^n P_i = 1$, легче рассчитать по выражению

$$P_{m \geq 2}(t) = 1 - P_0 + P_1.$$

Экспоненциальный закон надежности

Если рассматривать случай $m = 0$, то $P_0(t) = \frac{(\lambda t)^0}{0!} e^{-\lambda t} = e^{-\lambda t}$,

или

$$P(t) = e^{-\lambda t}. \quad (13.1)$$

Эта формула получила название экспоненциального закона надежности (рис. 13.2).

Можно вывести эту формулу и исходя из других соображений. Вероятность безотказной работы находится по формуле

$$(11.13) \quad P(t) = e^{-\int_0^t \lambda(\tau) d\tau}.$$

Для периода нормальной эксплуатации $\lambda(\tau) = \lambda - \text{const}$. Тогда получим

$$P(t) = e^{-\lambda \int_0^t d\tau} = e^{-\lambda \tau} \Big|_0^t = e^{-\lambda(t-0)} = e^{-\lambda t}. \quad (13.2)$$

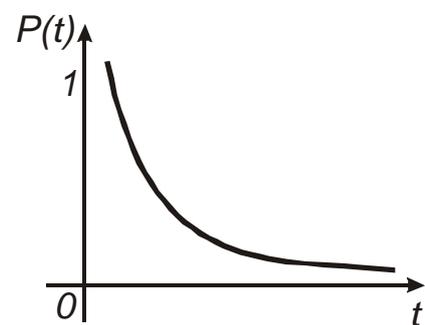


Рис. 13.2

Средняя наработка до отказа

$$t_{\text{cp}} = \int_0^{\infty} P(t) dt = \int_0^{\infty} e^{-\lambda t} dt = -\frac{1}{\lambda} e^{-\lambda t} \Big|_0^{\infty} = -\frac{1}{\lambda} (0 - 1) = \frac{1}{\lambda}.$$

Эта формула

$$t_{\text{cp}} = \frac{1}{\lambda}, \quad (13.3)$$

как и формула (13.2), имеет большое значение в теории надежности.

Учитывая, что $\lambda(t) = \frac{f(t)}{P(t)}$, получим плотность распределения

при экспоненциальном законе в виде $f(t) = \lambda e^{-\lambda t}$.

Экспоненциальный закон надежности – основной закон для расчета сложных электрических схем в период нормальной эксплуатации изделий.

Нормальный закон распределения

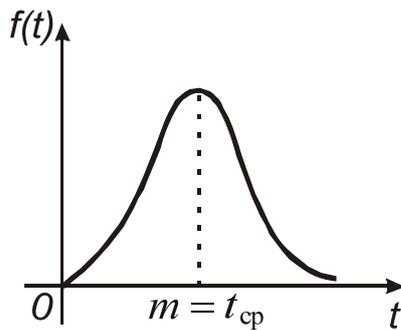


Рис. 13.3

Плотность распределения

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-t_{\text{cp}})^2}{2\sigma^2}}$$

графически показана на рис. 13.3.

Учитывая, что $f(t) = F'(t) = Q'(t)$, получим

$$Q(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^t e^{-\frac{(t-t_{\text{cp}})^2}{2\sigma^2}} d\tau.$$

Введя обозначения

$$u = \frac{\tau - t_{\text{cp}}}{\sigma}; \quad \tau = u\sigma + t_{\text{cp}},$$

найдем

$$Q(t) = \frac{1}{\sqrt{2\pi}} \int_0^{\frac{t-t_{\text{cp}}}{\sigma}} e^{-\frac{u^2}{2}} du.$$

Это выражение табулировано и называется интегралом вероятностей, или функцией Лапласа

$$Q(t) = \Phi\left(\frac{t-t_{\text{cp}}}{\sigma}\right).$$

Вероятность безотказной работы

$$P(t) = 1 - \Phi\left(\frac{t - t_{\text{ср}}}{\sigma}\right).$$

Нормальный закон распределения является моделью отказов в период старения и износа.

Закон Вейбулла

Плотность распределения (рис. 13.4)

$$f(t) = \frac{b}{a} \left(\frac{t}{a}\right)^{b-1} e^{-\left(\frac{t}{a}\right)^b},$$

где a характеризует масштаб распределения, b – форму распределения.

Интенсивность отказов

$$\lambda(t) = \frac{b}{a} \left(\frac{t}{a}\right)^{b-1};$$

вероятность безотказной работы

$$P(t) = e^{-\left(\frac{t}{a}\right)^b}.$$

Пусть $b = 1$, тогда $\lambda(t) = \frac{1}{a}$; $a = t_{\text{ср}}$.

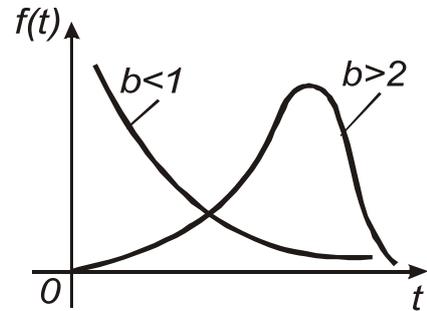


Рис. 13.4

Получим $P(t) = e^{-\frac{t}{a}}$, то есть при $b = 1$ закон Вейбулла совпадает с экспоненциальным законом распределения. При $b = 3$ закон Вейбулла практически совпадает с нормальным законом распределения.

Закон распределения Вейбулла широко применяется в теории надежности, так как он более универсален, чем ранее рассмотренные законы распределения. Наиболее часто используется как модель отказов для различных механических и электромеханических изделий.

Лекция 14

МЕТОДЫ РАСЧЕТА НАДЕЖНОСТИ

Классификация методов расчета надежности

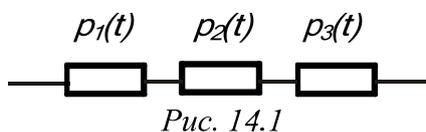
Методы расчета надежности в зависимости от вида соединения элементов можно разделить на методы при последовательном, параллельном и смешанном соединениях.

Последовательное соединение – когда отказ хотя бы одного элемента приводит к отказу всей системы. *Параллельное соединение* – это когда отказ системы наступает при отказе всех составляющих её элементов.

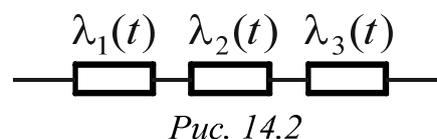
Методы расчета надежности также можно классифицировать в зависимости от видов отказов. Они делятся на расчеты при внезапных, постепенных и перемежающихся отказах.

В зависимости от этапа работы и исходных данных выделяют приближенные методы расчета надежности, проводящиеся на начальных стадиях проектирования, и точные методы.

Методы расчета надежности при последовательном соединении элементов



Пример. Необходимо определить вероятность безотказной работы $P(t)$ системы, показанной на рис. 14.1. Событие, заключающееся в безотказной работе, будет наблюдаться, если безотказно будут работать элементы 1, 2 и 3. Тогда вероятность безотказной работы системы



При последовательном соединении элементов вероятность безотказной работы системы равняется произведению вероятностей безотказной работы элементов (рис. 14.2).

Пусть Λ – интенсивность отказа системы, законы экспоненциальные. Можно записать

$$P(t) = e^{-\Lambda t} = e^{-\lambda_1 t} e^{-\lambda_2 t} e^{-\lambda_3 t} = e^{-(\lambda_1 + \lambda_2 + \lambda_3)t}.$$

Следовательно,

$$\Lambda = \lambda_1 + \lambda_2 + \lambda_3.$$

При последовательном соединении элементов и экспоненциальном законе интенсивность отказов системы определяется как сумма интенсивностей отказов элементов, составляющих данную систему.

Приближенные методы расчета надежности

Прикидочный метод. Применяется на самых ранних стадиях проектирования изделий (на стадиях технического задания и технического предложения) и основан на следующих допущениях:

- законы распределения времени безотказной работы экспоненциальные;
- соединение элементов последовательное;
- имеется аналог, характеристики надежности которого известны;
- все составляющие проектируемого изделия и аналога считаются равнонадежными.

Порядок расчета:

1. Определяется усредненная интенсивность отказов элементов аналога по формуле

$$\lambda_a = \frac{\Lambda_a}{N_a},$$

где Λ_a – интенсивность отказа аналога (как системы); N_a – количество элементов в аналоге.

2. Интенсивность отказов проектируемого изделия принимается равной интенсивности отказов элементов аналога $\lambda_n = \lambda_a$.

3. Определяется число элементов проектируемого изделия N_n .

4. Находится интенсивность отказов проектируемого изделия

$$\Lambda_n = \lambda_n N_n.$$

5. Рассчитываются $P(t) = e^{-\Lambda_n t}$, $t_{cp n} = \frac{1}{\lambda_n}$.

Ориентировочный метод. Проводится на стадиях эскизного или технического проектирования, т.е. когда известна принципиальная электрическая схема изделия.

Допущения:

- законы экспоненциальные;
- соединение элементов последовательное;
- режимы работы элементов нормальные.

Порядок расчета:

1. Определяется количество элементов каждого типа N_i .
2. Находится интенсивность отказов элементов каждого типа (по таблицам, например в [5]).
3. Определяется интенсивность отказа системы

$$\Lambda = \lambda_1 N_1 + \lambda_2 N_2 + \dots = \sum_{i=1}^n \lambda_i N_i,$$

где n – количество типов радиоэлементов.

4. Рассчитываются $P(t) = e^{-\lambda t}$ и $t_{cp} = \frac{1}{\Lambda}$.

Расчет надежности с учетом режимов работы элементов. Последовательность расчета:

1. Определяют режимы работы элементов \mathcal{G}_i .
2. Устанавливают количество элементов каждого типа, работающих в одинаковых режимах $N_i(\mathcal{G})$.

3. Определяют интенсивность отказов элементов с учетом режимов их работы $\lambda_i(\mathcal{G}) = \lambda_{i0} a_1 a_2 \dots$, где λ_{i0} – интенсивность отказов при нормальных условиях; a_1, a_2, \dots – коэффициенты, учитывающие режимы работы.

4. Находят $\Lambda(\mathcal{G}) = \sum_{i=1}^m \lambda_i(\mathcal{G}) N_i(\mathcal{G})$, где m – количество одинаковых элементов, работающих в идентичных условиях.

5. Рассчитывают $P_n(t) = e^{-\Lambda(\mathcal{G})t}$; $t_{cp}(\mathcal{G}) = \frac{1}{\Lambda(\mathcal{G})}$.

Расчет надежности корпусированных полупроводниковых микросхем. Интенсивность отказов в микросхеме определяется при рассмотрении ее как функционального узла, в который входят транзисторы, диоды, соединения внешние и внутренние, то есть

$$\Lambda = N_T a_T \lambda_T + N_D a_D \lambda_D + (3N_T + 2N_D + N_B) \lambda_{соед},$$

где N_T и N_D – число транзисторных и диодных переходов;

λ_T и λ_D – интенсивность отказов;

a_T и a_D – коэффициент, учитывающий режим работы;

$\lambda_{соед}$ – интенсивность отказов соединений;

N_B – количество внешних выводов.

Ориентировочно

$$\lambda_T = 1 \cdot 10^{-8} \text{ ч}^{-1}; \lambda_{соед} = 0,1 \cdot 10^{-8} \text{ ч}^{-1}; \lambda_D = 0,6 \cdot 10^{-8} \text{ ч}^{-1}.$$

Лекция 15

МЕТОДЫ РАСЧЕТА НАДЕЖНОСТИ ПРИ ПАРАЛЛЕЛЬНОМ И СМЕШАННОМ СОЕДИНЕНИЯХ ЭЛЕМЕНТОВ

Понятия о резервировании

Резервирование – это способ повышения надежности путем введения дополнительных средств и возможностей с целью получения избыточности (структурной, информационной, временной и т.д.).

Структурное резервирование реализуется путем параллельного соединения элементов, как показано на рис. 15.1.

Под кратностью резервирования понимают отношение количества резервных цепей к количеству основных цепей.

Резервирование подразделяется на постоянное и динамическое.

При постоянном резервировании резервные цепи постоянно находятся в работе, как и основные цепи. К моменту отказа основных цепей ресурс резервных цепей снижается.

При динамическом резервировании основные цепи замещают резервными только после отказа основных цепей.

При динамическом резервировании элемент может находиться в ненагруженном состоянии (холодный резерв), нагруженном состоянии (горячий резерв) и среднем состоянии (теплый резерв).

Резервирование делится на общее и отдельное. При общем резервировании резервируется система в целом. При отдельном – отдельные элементы.

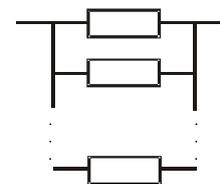


Рис. 15.1

Расчет надежности при параллельном соединении элементов

Для системы (рис. 15.2) $P(t) = 1 - Q(t)$,
где $Q(t) = q_1(t)q_2(t) \dots q_{m+1}(t)$ – вероятность отказа.

$$P(t) = 1 - \{[1 - p_1(t)] \dots [1 - p_{m+1}(t)]\} =$$

$$= 1 - \prod_{j=1}^{m+1} [1 - p_j(t)].$$

Если $p_1(t) = p_2(t) = \dots = p(t)$, то вероятность безотказной работы

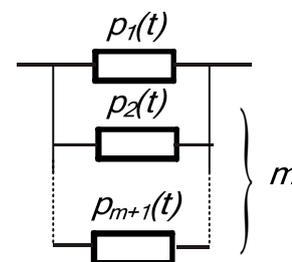


Рис. 15.2

$$P(t) = 1 - [1 - p(t)]^{m+1},$$

где m – число резервных элементов.

Пример. Если $m = 2$, $p(t) = 0,8$ (вероятность безотказной работы каждого элемента), то вероятность безотказной работы системы $P(t) = 0,992$.

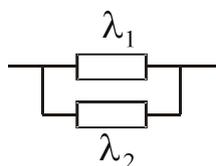


Рис.15.3

Пусть имеется система, показанная на рис. 15.3.

Получим

$$P(t) = 1 - [1 - p_1(t)][1 - p_2(t)] = p_1(t) + p_2(t) - p_1(t)p_2(t). \text{ Если законы экспоненциальные, то}$$

$$P(t) = e^{-\lambda_1 t} + e^{-\lambda_2 t} - e^{-(\lambda_1 + \lambda_2)t}$$

и

$$t_{\text{ср}} = \int_0^{\infty} P(t) dt = \int_0^{\infty} [e^{-\lambda_1 t} + e^{-\lambda_2 t} - e^{-(\lambda_1 + \lambda_2)t}] dt.$$

Получим

$$t_{\text{ср}} = \frac{1}{\lambda_1} + \frac{1}{\lambda_2} - \frac{1}{\lambda_1 + \lambda_2}.$$

Если $\lambda_1 = \lambda_2 = \lambda$, то

$$t_{\text{ср}} = \frac{2}{\lambda} - \frac{1}{2\lambda} = \frac{3}{2\lambda} = 1,5 \frac{1}{\lambda} = 1,5 t_{\text{ср.э}},$$

где $t_{\text{ср.э}}$ – средняя наработка до отказа одного элемента.

При постоянном резервировании ресурс резервного элемента является значительно исчерпанным, поэтому увеличивая количество элементов в два раза, только в 1,5 раза повысим надежность.

Расчет надежности при общем резервировании

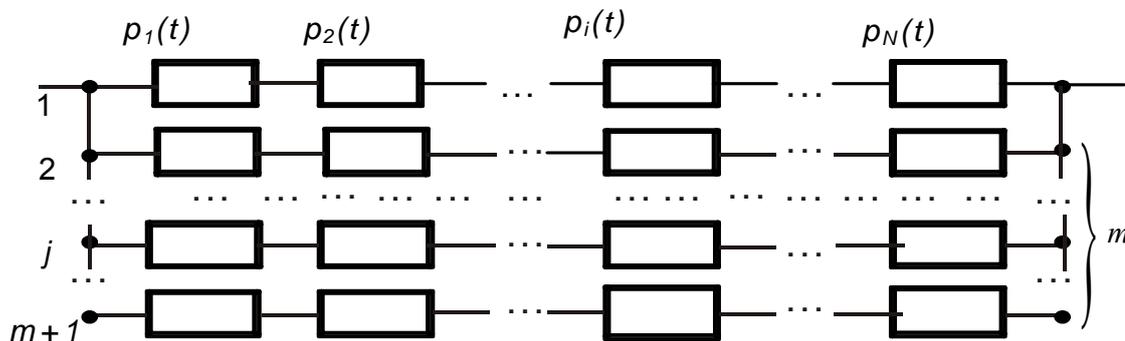


Рис. 15.4

Необходимо определить вероятность безотказной работы системы (рис. 15.4), если одна цепь основная, а остальные m резервные.

Данная система откажет, когда откажут все цепи, составляющие систему. Тогда вероятность отказа системы будет иметь вид

$$Q(t) = \prod_{i=1}^{m+1} q_{ци}(t), \text{ где } q_{ци}(t) \text{ – вероятность отказа } i\text{-й цепи.}$$

Вероятность безотказной работы цепи

$$P_{ци}(t) = \prod_{j=1}^N p_j(t).$$

Отсюда
$$Q(t) = \prod_{i=1}^{m+1} [1 - p_{ци}(t)] = \prod_{i=1}^{m+1} [1 - \prod_{j=1}^N p_j(t)].$$

Окончательно получим

$$P(t) = 1 - \prod_{i=1}^{m+1} \left[1 - \prod_{j=1}^N \overline{P}_j(t) \right].$$

Расчет надежности при раздельном резервировании

При раздельном резервировании (рис. 15.5) резервируется не вся цепь, а отдельные элементы.

Вероятность безотказной работы

$$P(t) = \prod_{j=1}^N p_{звj}(t);$$

$$Q_{звj}(t) = \prod_{i=1}^{m+1} q_i(t);$$

$$P_{звi}(t) = 1 - \prod_{i=1}^{m+1} q_i(t) = 1 - \prod_{i=1}^{m+1} [1 - p_i(t)].$$

Вероятность безотказной работы системы

$$P(t) = \prod_{j=1}^N P_{звj}(t),$$

и окончательно получим

$$P(t) = \prod_{j=1}^N \left(1 - \prod_{i=1}^{m+1} [1 - p_i(t)] \right).$$

Будем считать, что все элементы равнонадежны, то есть $p_1(t) = p_2(t) = \dots = p(t)$, тогда $p_1(t) = 1 - [1 - p(t)]^{m+1}$ при общем резервировании и $P(t) = \{1 - [1 - p(t)]^{m+1}\}^N$ при раздельном резервировании.

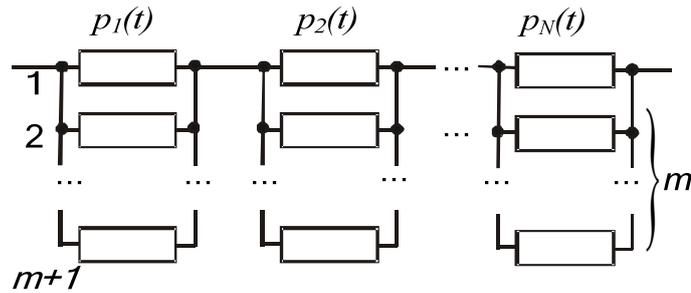


Рис. 15.5

Предположим, что $m = 1$, $N = 2$, $P(t) = 0,8$, тогда $p_1(t) = 1 - [1 - 0,8^2]^2 = 0,87$ при общем резервировании; $P(t) = \{1 - (1 - 0,8)^2\}^2 = 0,92$ при раздельном резервировании.

Видно, что раздельное резервирование эффективнее общего. Но этот вывод справедлив только для постоянного резервирования.

Если резервирование динамическое, т.е. осуществляется замещением отказавших элементов, то при раздельном резервировании требуется значительно большее количество переключателей, которые являются наименее надежными элементами. Поэтому раздельное резервирование в этом случае часто оказывается менее эффективным, чем общее резервирование.

Принципы расчета надежности при динамическом резервировании

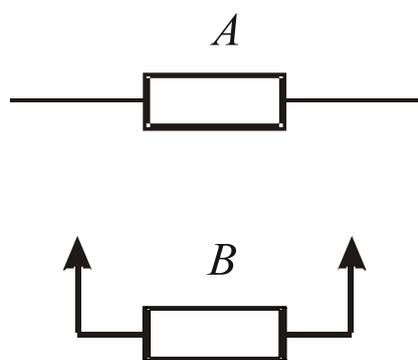


Рис. 15.6

Возможны гипотезы H о работоспособности системы, показанной на рис. 15.6.

H_1 – элементы A и B проработали безотказно в течение заданного времени t .

H_2 – элемент A отказал в момент времени $\tau < t$, включился элемент B и проработал до окончания заданного времени t .

H_3 – элемент A проработал безотказно в течение заданного времени t , а B отказал в момент времени $\tau < t$.

H_4 – элемент A отказал в момент времени $\tau < t$, включился элемент B и тоже отказал до наступления момента времени t .

H_5 – элемент A отказал в момент времени τ , а элемент B еще раньше.

Тогда вероятность работоспособного состояния будет равняться сумме вероятностей первых трех гипотез, то есть

$$P(t) = \sum_{i=1}^3 P(H_i).$$

Вероятность первой гипотезы

$$P(H_1) = P_A(t)P_B(t) = e^{-\lambda_A t} e^{-\lambda_B t} = e^{-(\lambda_A + \lambda_B)t}.$$

Более подробно эти методы изложены в литературе [5].

МЕТОДЫ ОПТИМИЗАЦИИ ПРИ ПРОЕКТИРОВАНИИ ЭЛЕКТРОННОЙ АППАРАТУРЫ

Л е к ц и я 16

КЛАССИЧЕСКИЕ МЕТОДЫ ОПТИМИЗАЦИИ

Введение

Под *оптимизацией* понимают нахождение наилучшего варианта изделия (технологического процесса) из какого-либо множества вариантов.

Если, например,

$$y = f(x_1, x_2, \dots, x_n),$$

то необходимо найти решение, которое обеспечило бы наилучшее значение параметра y .

Можно выделить оптимизацию параметров и оптимизацию структуры.

При *оптимизации параметров* необходимо найти такие значения параметров x_i , при которых функция y , называемая *целевой функцией*, имеет при *постоянной структуре системы* экстремальное значение.

При *оптимизации структуры* ищут такую структуру системы, т.е. такое преобразование f , при которой целевая функция y при постоянных параметрах x_i ($i = 1, 2, 3, \dots, n$) принимает экстремальное значение.

Задача оптимизации может решаться при ограничениях типа равенств, неравенств, дискретности:

$$\left. \begin{aligned} \Phi(x_1, x_2, \dots, x_n) &= 0, \\ \Phi(x_1, x_2, \dots, x_n) &< b, \\ x_1 &= x_{11}, x_{12}, \dots, x_{1n}, \\ x_2 &= x_{21}, x_{22}, \dots, x_{2m}. \end{aligned} \right\} \quad (16.1)$$

Выражения (16.1) называются *функциями ограничений*, или *функциями связи*.

Задача оптимизации может быть решена экспериментальными или расчетно-аналитическими методами.

Расчетно-аналитические методы применяют в том случае, когда целевая функция или функция связи могут быть представлены в виде математического выражения. Задача решается в два этапа.

1. Находятся математические выражения для целевой функции и функции связи.

2. Производится непосредственно оптимизация.

Методы оптимизации можно разделить:

- на классические;
- методы регулярного и случайного поиска оптимума;
- методы математического программирования.

Классические методы оптимизации

К классическим методам оптимизации относят метод производных, метод множителей Лагранжа, метод оптимизации на основе вариационного исчисления.

Оптимизация методом производных. Пусть имеется целевая функция $y = f(x_1, x_2, \dots, x_n)$, у которой нет ограничений, заданная в аналитическом виде, непрерывная и дифференцируемая на рассматриваемом интервале.

Находятся производные от функции y по каждому из параметров и приравниваются к нулю:

$$\left. \begin{array}{l} \frac{\partial y}{\partial x_1} = 0, \\ \frac{\partial y}{\partial x_2} = 0, \\ \dots \\ \frac{\partial y}{\partial x_n} = 0. \end{array} \right\}$$

Получена система из n уравнений, решая которую, находят значения параметров x_1, x_2, \dots, x_n , при которых целевая функция будет иметь экстремальное значение.

Пример. Необходимо найти оптимальные параметры и экстремальные значения целевой функции $y = 6 + 3x_1 + 4x_2 - 0,5x_1^2 - 0,4x_2^2 + 0,8x_1x_2$ и определить характер экстремума (максимум или минимум).

Находим первые производные и приравниваем их к нулю:

$$\left. \begin{array}{l} \frac{\partial y}{\partial x_1} = 3 - x_1 + 0,8x_2 = 0, \\ \frac{\partial y}{\partial x_2} = 4 - 0,8x_2 + 0,8x_1 = 0. \end{array} \right\}$$

Решая эту систему, найдем оптимальные решения:

$$x_1=35; \quad x_2=40.$$

Экстремум целевой функции

$$y=6+3\cdot 25+4\cdot 40-0,5\cdot 35^2-0,4\cdot 40^2+0,8\cdot 35\cdot 40=138,5.$$

Вторые производные равны

$$\frac{\partial^2 y}{\partial x_1^2} = -1; \quad \frac{\partial^2 y}{\partial x_2^2} = -0,8.$$

Следовательно, $y = 138,5$ – максимум.

Оптимизация методом множителей Лагранжа. Пусть $y = f(x_1, x_2, \dots, x_n)$ – целевая функция.

Имеются ограничения

$$\left. \begin{aligned} \Phi_1(x_1, x_2, \dots, x_n) &= c_1, \\ \Phi_2(x_1, x_2, \dots, x_n) &= c_2, \\ &\dots \\ \Phi_m(x_1, x_2, \dots, x_n) &= c_m, \end{aligned} \right\} \quad (16.2)$$

где c_i ($i = 1, 2, \dots, n$) – постоянные величины.

Параметры x_i – независимые величины.

Функции связи и целевая функция непрерывны и дифференцируемы.

В точке экстремума полный дифференциал

$$dy = \sum_{i=1}^n \frac{\partial y}{\partial x_i} dx_i = 0. \quad (16.3)$$

Полные дифференциалы функций связи

$$\left. \begin{aligned} d\Phi_1 &= \sum_{i=1}^n \frac{\partial \Phi_1}{\partial x_i} dx_i = 0, \\ d\Phi_2 &= \sum_{i=1}^n \frac{\partial \Phi_2}{\partial x_i} dx_i = 0, \\ &\dots \\ d\Phi_m &= \sum_{i=1}^n \frac{\partial \Phi_m}{\partial x_i} dx_i = 0. \end{aligned} \right\}$$

Каждое из уравнений этой системы умножим на постоянный множитель λ (для каждого уравнения он будет свой). Множители λ называются *неопределенными множителями Лагранжа*.

$$\left. \begin{aligned} \lambda_1 d\Phi_1 &= \sum_{i=1}^n \lambda_1 \frac{\partial \Phi_1}{\partial x_i} dx_i = 0, \\ \lambda_2 d\Phi_2 &= \sum_{i=1}^n \lambda_2 \frac{\partial \Phi_2}{\partial x_i} dx_i = 0, \\ &\dots \\ \lambda_m d\Phi_m &= \sum_{i=1}^n \lambda_m \frac{\partial \Phi_m}{\partial x_i} dx_i = 0. \end{aligned} \right\}$$

Складывая левые и правые части этих уравнений и уравнение (16.3), получим

$$\sum_{i=1}^n \left(\frac{\partial y}{\partial x_i} + \lambda_1 \frac{\partial \Phi_1}{\partial x_i} + \dots + \lambda_m \frac{\partial \Phi_m}{\partial x_i} \right) dx_i = 0.$$

Так как параметры x независимы, это уравнение обратится в нуль, когда каждое выражение в скобках равно нулю.

Запишем

$$\left. \begin{aligned} \frac{\partial y}{\partial x_1} + \lambda_1 \frac{\partial \Phi_1}{\partial x_1} + \dots + \lambda_m \frac{\partial \Phi_m}{\partial x_1} &= 0, \\ \frac{\partial y}{\partial x_2} + \lambda_1 \frac{\partial \Phi_1}{\partial x_2} + \dots + \lambda_m \frac{\partial \Phi_m}{\partial x_2} &= 0, \\ &\dots \\ \frac{\partial y}{\partial x_n} + \lambda_1 \frac{\partial \Phi_1}{\partial x_n} + \dots + \lambda_m \frac{\partial \Phi_m}{\partial x_n} &= 0. \end{aligned} \right\} \quad (16.4)$$

Решая совместно n уравнений системы (16.4) и m уравнений функций связи (16.2), можно определить n неизвестных параметров x и m неизвестных множителей Лагранжа λ .

Пример. Требуется спроектировать цилиндрический сосуд заданного объема V , который бы имел наименьшую площадь поверхности S (рис. 16.1).

Площадь поверхности цилиндра

$$S = 2\pi r^2 + 2\pi r l. \quad (16.5)$$

Здесь S – целевая функция.

Так как имеется ограничение $\pi r^2 l = V$, то функцию связи можно записать в виде

$$\Phi = V - \pi r^2 l. \quad (16.6)$$

На основе (16.4) запишем

$$\begin{aligned} \frac{\partial S}{\partial r} + \lambda_1 \frac{\partial \Phi}{\partial r} &= 0, \\ \frac{\partial S}{\partial l} + \lambda_1 \frac{\partial \Phi}{\partial l} &= 0. \end{aligned}$$

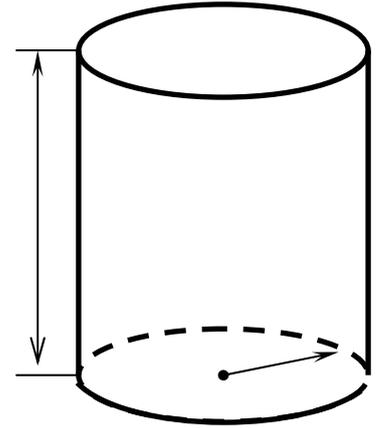


Рис. 16.1

Подставляя в эти уравнения выражение (16.5), находим

$$4\pi r + 2\pi l + \lambda_1(-2\pi l) = 0,$$

$$2\pi r + \lambda_1(-\pi l^2) = 0.$$

Решая эти уравнения совместно с уравнением (16.6), найдем значения параметров r и l , при которых площадь S будет минимальной.

Оптимизация на основе вариационного исчисления. Функционал – это когда параметру ставится в соответствие функция. Например, функционал

$$J = \int_{x_1}^{x_2} F\left(x, y, \frac{dy}{dx}\right) dx,$$

где F – функция.

В вариационном исчислении доказано, что если функция F удовлетворяет уравнению Эйлера

$$\frac{\partial F}{\partial y} - \frac{\partial}{\partial x} \left(\frac{\partial F}{\partial y'} \right) = 0,$$

то функционал J имеет экстремальное значение.

Пример. Найти зависимость $y = f(x)$, при которой расстояние между точками A и B будет минимально (рис. 16.2).

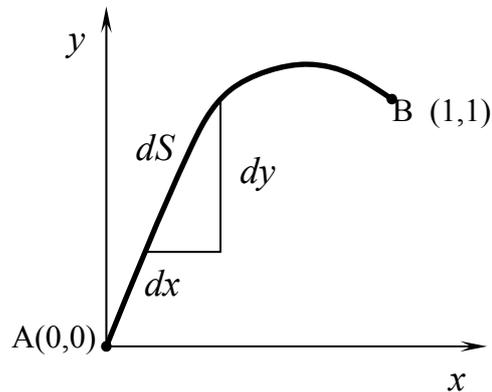


Рис. 16.2

Очевидно, что $dS = \sqrt{dy^2 + dx^2}$.

Обозначая расстояние между точками J , получим

$$J = \int_A^B dS = \int_A^B \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx.$$

Обозначим $F = \sqrt{1 + y'^2}$.

Решая последовательно уравнение Эйлера, найдем

$$\frac{\partial F}{\partial y} = \frac{\partial(\sqrt{1 + y'^2})}{\partial y} = 0,$$

$$\frac{\partial(\sqrt{1 + y'^2})}{\partial y'} = \frac{1}{2}(1 + y'^2)^{-1/2} \cdot 2y' = \frac{y'}{\sqrt{1 + y'^2}},$$

$$\frac{\partial}{\partial x} \left(\frac{y'}{\sqrt{1 + y'^2}} \right) = 0.$$

Отсюда $\frac{y'}{\sqrt{1 + y'^2}} = c$, где c – постоянная величина. Очевидно,

что $y'^2 = c^2 + c^2 y'^2$ или $y' = \frac{c}{\sqrt{1 - c^2}}$. Интегрируя это выражение, по-

лучим $y = \frac{c}{\sqrt{1 + c^2}} x + k$.

Неизвестные постоянные k и c найдем, используя граничные условия:

– в точке A : $x = 0$; $y = 0$; $k = 0$;

– в точке B : $x = 1$, $y = 1$, $\frac{c}{\sqrt{1 - c^2}} = 1$.

Следовательно, получим $y = x$.

Оптимизация методом вариационного исчисления применяется при оптимизации систем с распределенными параметрами.

Лекция 17

МЕТОДЫ РЕГУЛЯРНОГО ПОИСКА ЭКСТРЕМУМА [14]

Экстремумов функции может быть несколько (рис. 17.1). Один из них – глобальный экстремум, остальные – частные экстремумы.

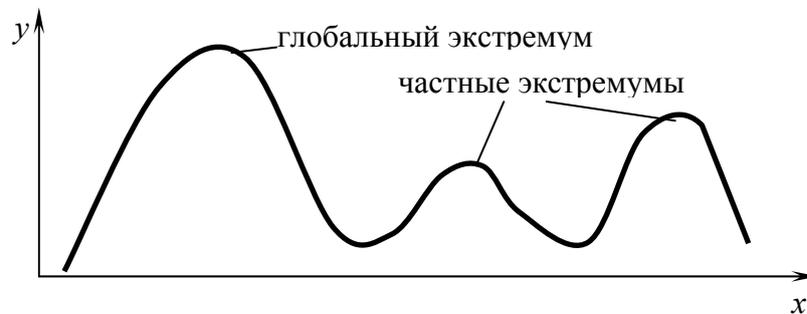


Рис. 17.1

Для объяснения ниже рассматриваемых методов целесообразно использовать топографические графики (рис. 17.2).

Для наглядности будем искать экстремум функции $y = f(x_1, x_2)$.

Будем считать, что экстремум функции равен ее максимальному значению.

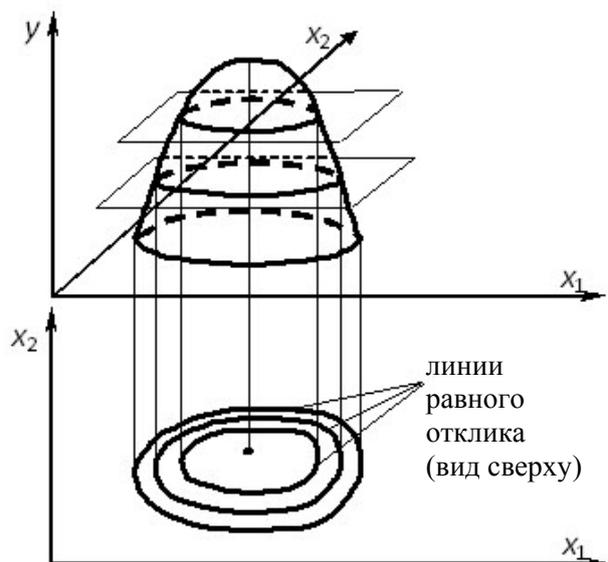


Рис. 17.2

Метод слепого поиска (метод сканирования)

Дано $y = f(x_1, x_2)$, и x_1 принимает значения $x_{11}, x_{12}, x_{13}, \dots, x_{1n}$, $x_2 - x_{21}, x_{22}, x_{23}, \dots, x_{2m}$.

В этом методе значение y определяется для всех возможных сочетаний x :

$$\begin{aligned} x_{11}, x_{21} &\rightarrow y_1 \\ x_{11}, x_{22} &\rightarrow y_2 \\ &\dots \\ x_{11}, x_{2n} &\rightarrow y_i \\ x_{1n}, x_{2m} &\rightarrow y_k \end{aligned}$$

Находится максимальное значение y_{\max} , а соответствующие ему параметры x_{10}, x_{20} и будут оптимальными.

Если x_1 и x_2 непрерывные, т.е. имеют бесчисленное множество значений, то вначале переходят к дискретным параметрам (как в методе матричных испытаний), а затем ищут экстремум.

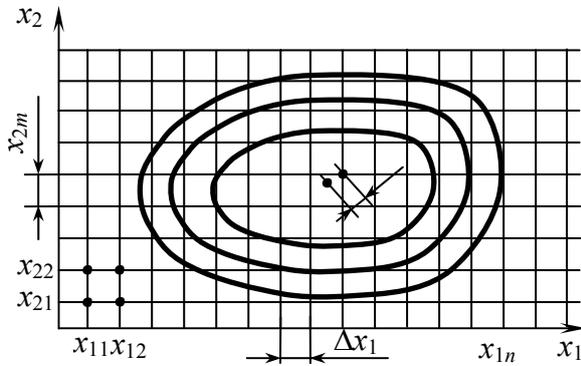


Рис. 17.3

Ошибка (ε) в определении нахождения экстремума (рис. 17.3) зависит от величины шагов ($\Delta x_1, \Delta x_2$) – чем меньше шаг, тем выше точность.

Метод довольно трудоемкий, однако позволяет находить как глобальные, так и частные экстремумы.

Метод Гаусса – Зейделя

Метод Гаусса – Зейделя предусматривает поочередное нахождение частных экстремумов целевой функции по каждому фактору x_i ($i = 1, 2, \dots, n$). При этом на каждом i -м этапе стабилизируют $(n - 1)$ факторов и варьируют только один i -й фактор. Графическая интерпретация метода дана на рис. 17.4, где на плоскости двух факторов $x_1,$

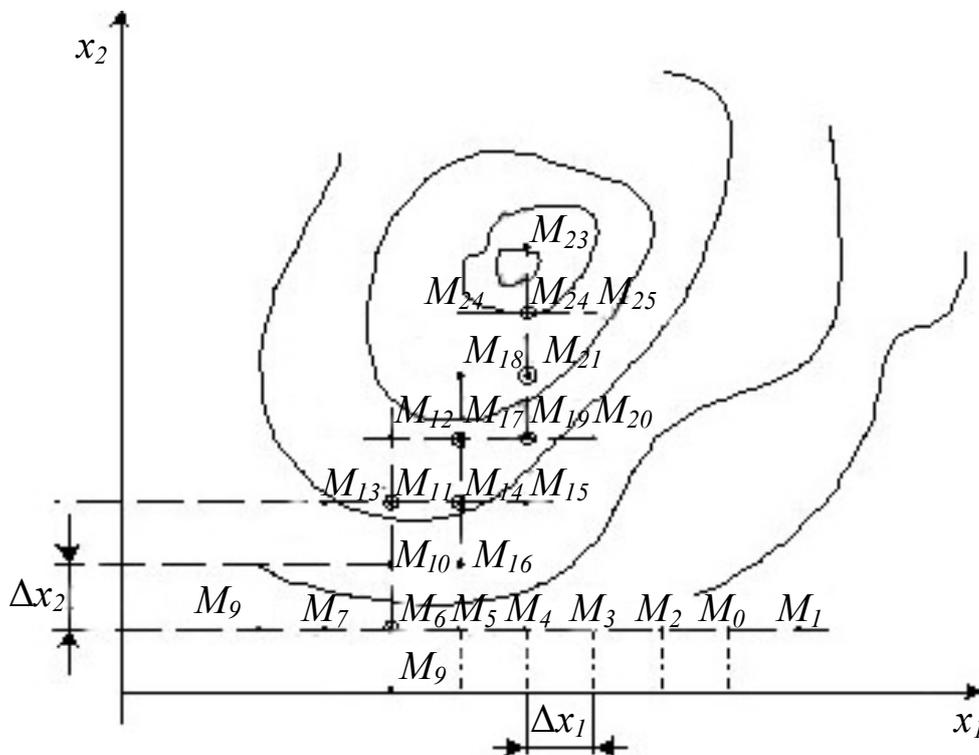


Рис. 17.4

x_2 изображена функция отклика y топографическим способом с помощью замкнутых линий постоянного уровня этой оптимизируемой выходной функции. Эти линии (см. рис. 17.4) соответствуют некоторым относительным величинам, однако, как указывалось выше, форма функции отклика до начала исследования обычно неизвестна. Путь движения обозначен точками M . Задачу поиска максимума методом Гаусса – Зейделя решают в несколько этапов, объединенных в циклы. Рассмотрим процедуру метода с иллюстрацией двухфакторного пространства.

I этап. 1. Выбирают основную (начальную, базовую) точку (на рис. 17.4 это точка M_0), обычно она соответствует номинальному режиму ведения технологического процесса $\vec{x}_0 = (x_{10}; x_{20}; \dots; x_{n0})$.

Иногда эту точку выбирают в центре области, которую желательно исследовать, либо в центре области ограничений, если они имеются. При таком выборе базовой точки все направления оказываются равноправными, а это важно в случае, когда заведомо ничего неизвестно о том, где, хотя бы примерно, расположен экстремум.

2. Выбирают интервал (ступень) варьирования Δx_1 (см. рис. 17.4) по фактору x_1 . Очевидно, что ступень варьирования не должна быть слишком малой, иначе движение к экстремуму окажется замедленным. Кроме того, на интервале варьирования Δx_i ($i = 1, 2, \dots, n$) изменение целевой функции Δy должно быть существенно большим, чем погрешность ее измерения δy (не менее чем в 5 – 10 раз).

3. Определяют координаты пробных точек M_1 и M_2 :

$$\begin{aligned} \vec{x}(M_1) &= (x_{10} + \Delta x_1; x_{20}; \dots; x_{n0}), \\ \vec{x}(M_2) &= (x_{10} - \Delta x_1; x_{20}; \dots; x_{n0}). \end{aligned} \tag{17.1}$$

4. В точках M_1 и M_2 ставят пробные опыты (для повышения точности результатов могут выполняться параллельные опыты), измеряют отклики $y(M_1)$ и $y(M_2)$.

5. Сравнивают полученные отклики, и если

$$y(M_2) > y(M_1) \tag{17.2}$$

(как на рис. 17.4), то совершают рабочее движение на один рабочий шаг Δx_1 по направлению $\overline{M_0 M_2}$ в точку M_3 .

6. Аналогичные шаги продолжают в том же направлении до тех пор, пока на каком-то k -м шаге не окажется, что

$$y(M_k) < y(M_{k-1}), \tag{17.3}$$

т.е. значение отклика в очередной k -й рабочей точке станет уменьшаться. Это послужит признаком достижения частного экстремума, за который принимают $(k - 1)$ -ю точку с откликом $y(M_{k-1})$. На рис. 17.4 это точка M_6 .

II этап. Его проводят в том же порядке, что и I этап, с той лишь разницей, что стабилизируют все факторы, кроме x_2 . За новую базовую точку принимают точку с координатами

$$\bar{x}(M_{k-1}) = (x_{10} \pm \Delta x_1(k - 2); x_{20}; \dots; x_{n0}), \quad (17.4)$$

а x_2 варьируют на выбранную по аналогичным условиям величину интервала варьирования Δx_2 . По достижении частного экстремума по фактору x_2 точку нового частного экстремума принимают за новую базовую точку. На рис. 17.4 это точка M . Если начало движения из точки M_0 сразу совпало с возрастанием y , то в равенстве (17.4) вместо $(k - 1)$ берут $(k - 2)$.

Первый цикл продвижения к экстремуму заканчивается n -м этапом, на котором стабилизируют все факторы, кроме x . Для него выбирают ступень варьирования Δx_n и совершают пробное, а затем рабочее движение до достижения частного экстремума по фактору x_n . Если экстремум не достигнут, то выполняют второй цикл поиска.

Второй цикл, как и первый, начинается с I этапа, на котором варьируют фактор x_1 при стабилизации остальных x ($i \neq 1$), затем последовательно выполняют n этапов по каждому из n факторов.

Поисковое шаговое движение к экстремуму заканчивают по достижении той точки факторного пространства, при движении из которой в любую сторону по всем n -факторным осям x_i в положительном или отрицательном направлениях значения отклика оказываются меньшими. Такую точку принимают за экстремум (максимум).

Достоинства метода Гаусса – Зейделя: 1) очевидная простота стратегии и наглядность; 2) высокая помехозащищенность в смысле выбора направления движения.

Недостатки: 1) путь к главному экстремуму оказывается обычно долгим, особенно при большом числе n факторов; 2) в условиях крупного промышленного производства трудно застабилизировать $n - 1$ фактор на длительное время; 3) если поверхность отклика имеет сложную форму (узкие гребни, овраг и т. п.), то использование метода может привести к ложному ответу на вопрос о месте расположения экстремума; 4) метод не дает информации о взаимодействиях факторов.

Исторически метод Гаусса–Зейделя известен как первый из рассматриваемых. В настоящее время он иногда применяется при машинном эксперименте.

Градиентные методы

Градиентные методы имеют несколько разновидностей, различающихся правилами выбора ступеней варьирования и рабочих шагов на каждом этапе движения к экстремуму. Сущность стратегии всех этих разновидностей состоит в том, что на каждом этапе вокруг очередной базовой точки организуют пробные эксперименты, по результатам которых оценивают новое направление градиента, после чего в этом направлении совершают один рабочий шаг. Напомним, что вектор-градиент в n -факторном пространстве определяется соотношением

$$\text{grad } y = \frac{\partial y}{\partial x_1} \bar{x}_1^0 + \frac{\partial y}{\partial x_2} \bar{x}_2^0 + \dots + \frac{\partial y}{\partial x_n} \bar{x}_n^0, \quad (17.5)$$

где \bar{x}_i^0 ($i = 1, 2, \dots, n$) – единичные направляющие векторы (орты, расположенные вдоль факторных осей); $\frac{\partial y}{\partial x_i}$ – частная про-

изводная целевой функции по i -му фактору. Пробные опыты (по два в точках, расположенных на прямых, параллельных каждой факторной оси и проходящих через базовую точку) проводят с целью получить приближенные оценки частных производных. Рассмотрим две разновидности градиентных методов.

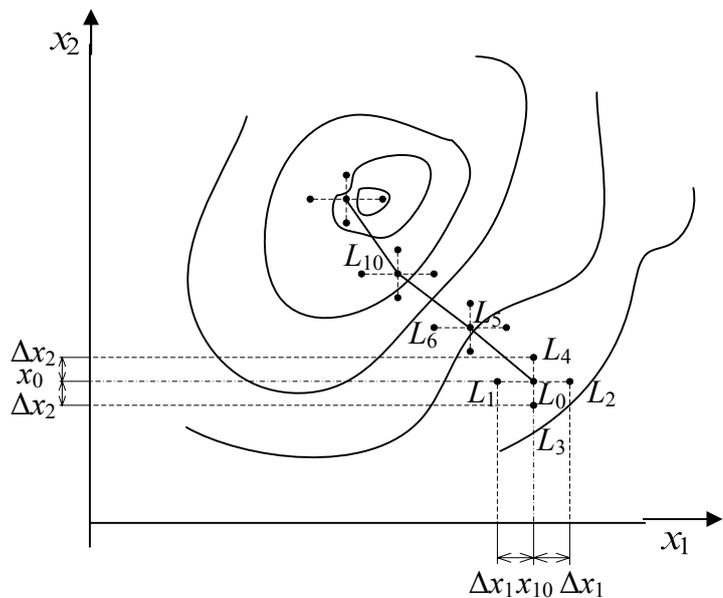


Рис. 17.5

Метод градиента (обычный) осуществляется по следующей процедуре.

1. Выбирают начальную (базовую) точку $\bar{x}_0 = (x_{10}; x_{20}; \dots; x_{n0})$. На рис. 17.5 это точка L_0 .

2. Выбирают интервал варьирования Δx_i по каждому из факторов x_i ($i = 1, 2, \dots, n$), пользуясь уже известными правилами.

3. Определяют координаты пробных точек (см. рис. 17.5). Например, вдоль направления, параллельного факторной оси x_1 , ими являются точки L_1, L_2 с координатами

$$\begin{aligned}\vec{x}(L_1) &= (x_{10} - \Delta x_1; x_{20}; \dots; x_{n0}), \\ \vec{x}(L_2) &= (x_{10} + \Delta x_2; x_{20}; \dots; x_{n0}),\end{aligned}\tag{17.6}$$

т. е. варьируют один фактор x_1 при стабилизации остальных факторов на базовом уровне. Аналогично вычисляют координаты пробных точек вдоль направлений, параллельных остальным факторным осям x_2, x_3, \dots, x_n . В пробных точках ставят опыты и получают значения целевой функции y .

4. По результатам пробных опытов вычисляют оценки составляющих вектор-градиента в точке L_0 для каждого i -го фактора

$$\text{grad } y(L_0)|_{x_i} = \frac{\partial y}{\partial x_i} \approx \frac{\Delta y|_{x_i}}{2\Delta x_i} = \hat{a}_i.\tag{17.7}$$

В частности, для фактора x_1 по результатам опытов в точках L_1 и L_2 вычисление выполняют по формуле

$$\text{grad } y(L_0)|_{x_1} \approx \frac{\Delta y|_{x_1}}{2\Delta x_1} = \frac{y(L_2) - y(L_1)}{x_1(L_2) - x_1(L_1)} = \hat{a}_1.\tag{17.8}$$

Как известно, частные производные являются коэффициентами a_i ($i = 1, 2, \dots, n; i \neq 0$) уравнения плоскости, касательной к поверхности отклика в точке L_0 :

$$y = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_n x_n.\tag{17.9}$$

Оценки \hat{a}_i коэффициентов получают по формуле (17.8).

5. Находят координаты рабочей точки на направлении градиента. Для этого выбирают параметр рабочего шага $\rho_{\text{гр}}$ и вычисляют координаты первой рабочей точки по всем факторным осям x_i ($i = 1, 2, \dots, n$):

$$x_{i1} = x_{i0} + \rho_{\text{гр}} \hat{a}_{i0}.\tag{17.10}$$

На рис. 17.5 первой рабочей точкой является точка L_5 . Чтобы из основной точки L_0 попасть в точку L_5 , от L_0 откладывают в масштабе отрезки, равные $\rho_{\text{гр}} \hat{a}_1$ и $\rho_{\text{гр}} \hat{a}_2$, причем если $\hat{a}_i < 0$, то по соответствующему фактору отрезок откладывают в отрицательном направлении от точки L_0 , т.е. для фактора x_1 – влево от точки L_0 , а для фактора x_2 – вниз от точки L_0 . Если $\hat{a}_i > 0$, то отрезки $\rho_{\text{гр}} \hat{a}_i$ откладывают в положительном направлении от основной точки.

6. Первую рабочую точку принимают за новую базовую точку и вокруг нее организуют новые пробные опыты для оценивания нового направления градиента, после чего совершают новый рабочий шаг (на рис. 17.5 – в точку L_{10}). В общем случае вокруг каждой k -й рабочей точки по результатам пробных опытов получают оценки составляющих градиента \hat{a}_i и совершают $(k + 1)$ -й рабочий шаг ($k = 1, 2, \dots$) в точку с координатами

$$x_{i,k+1} = x_{ik} + \rho_{гр} \hat{a}_{ik}. \quad (17.11)$$

7. Рабочее движение производят до тех пор, пока на очередном шаге все составляющие градиента не станут пренебрежимо малыми, т. е. все $\hat{a}_{i,k+1} \approx 0$ ($i = 1, 2, \dots, n$). Для этого достаточно, чтобы выполнялось неравенство

$$\rho_{гр} \hat{a}_{i,k+1} < 1. \quad (17.12)$$

Если по результатам пробных опытов в $(k + 1)$ -й рабочей точке выполняется условие (17.12), то движение к экстремуму прекращают и эту рабочую точку принимают за точку экстремума.

Достоинства метода градиента: 1) достаточная простота стратегии; 2) повышенная по сравнению с методом Гаусса – Зейделя скорость движения к экстремуму (эффективность).

Недостатки: 1) большая чуткость к помехам ϵ в отношении выбора направления рабочего движения; 2) в случаях, когда поверхность отклика имеет сложную форму, метод градиента может не привести к истинному экстремуму; 3) если поверхность отклика достаточно пологая, то в условиях помех метод мало эффективен в смысле точности выхода к экстремуму; 4) как и метод Гаусса – Зейделя, метод градиента не дает информации о взаимодействиях факторов, которые характеризуют степень кривизны поверхности отклика.

Метод Кифера – Вольфовица отличается от описанного выше обычного метода градиента тем, что если в первом из них размеры интервалов варьирования Δx_i при постановке пробных опытов и параметр $\rho_{гр}$ рабочего шага остаются неизменными на любом рабочем шаге, то в рассматриваемом методе Δx_{ik} и $\rho_{грk}$ выбирают в зависимости от номера k рабочего шага:

$$\Delta x_{ik} = \frac{\Delta x_{i0}}{k\gamma}, \quad \rho_{грk} = \frac{\rho_{гр0}}{k}, \quad (17.13)$$

где Δx_{i0} – начальный интервал варьирования в основной точке L_0 ; $\rho_{гр0}$ – начальное значение параметра рабочего шага; k – номер рабоче-

го шага ($k = 1, 2, \dots$); γ – постоянная степень, обычно выбираемая в пределах $0 < \gamma < 0,5$. Чаще всего полагают $\gamma = 0,25$.

Если в методе градиента фактический размер k -го рабочего шага уменьшается только из-за уменьшения градиента, т. е. крутизны наклона поверхности отклика, при приближении к области экстремума, то в методе Кифера – Вольфовица фактический размер рабочего шага уменьшается, как мы видели, и в прямой зависимости от номера этого шага.

Достоинство метода Кифера – Вольфовица по сравнению с немодифицированным методом – его повышенная точность отыскания экстремальной точки, если поверхность отклика достаточно крутая, а экстремум находится от базовой точки не слишком далеко. *Недостаток* этого метода – его низкая эффективность в условиях пологих поверхностей отклика. При очень пологих поверхностях отклика метод Кифера – Вольфовица вообще не приводит к цели: рабочие шаги становятся сравнимыми с погрешностями измерения до достижения экстремума. Остальные достоинства и недостатки, а также вся процедура работы такие же, как и в методе градиента.

Симплексный метод

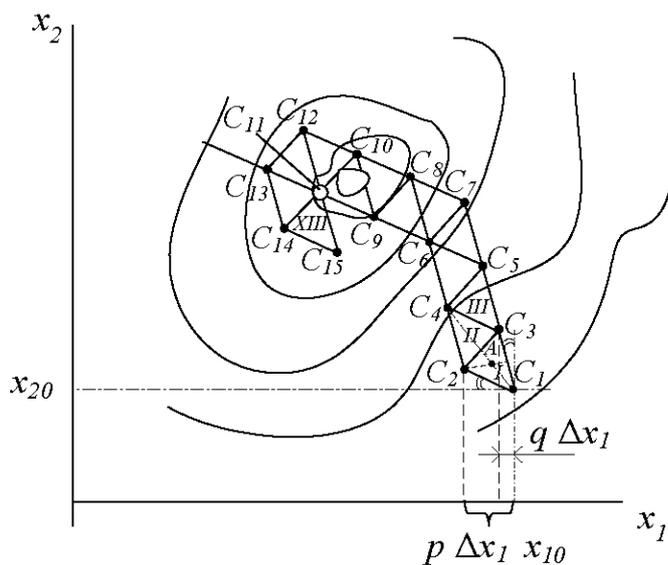


Рис. 17.6

Симплексом называют выпуклую фигуру (или тело), образованную $n + 1$ вершинами в пространстве n факторов, причем эти $n + 1$ вершин не принадлежат одновременно ни одному из подпространств из $n - 1$ факторов. В пространстве одного фактора ($n = 1$) симплексом служит отрезок установленного размера,

при $n = 2$ – треугольник, при $n = 3$ – тетраэдр. При $n \geq 4$ привычным образом интерпретировать симплекс невозможно.

Симплексный метод позволяет совмещать пробные опыты для определения направления движения с рабочим движением по поверхности отклика к области оптимума. Основная идея симплексного метода состоит в следующем. Если во всех $p + 1$ вершинах симплекса поставить опыты и измерить отклик, то (при не слишком большом уровне шумов) по величине отклика в вершинах можно судить, в каком направлении следует двигаться, чтобы приблизиться к экстремуму.

Рассмотрим это на примере двухфакторного пространства (рис. 17.6). Допустим, что на основе некоторых соображений (о них говорится ниже) построен начальный симплекс I с вершинами C_1, C_2, C_3 и в них измерен отклик y . Если уровень шума не слишком велик, то, очевидно, отклик в точке C_1 является наименьшим по сравнению с откликами в вершинах C_2 и C_3 . Тогда можно полагать, что максимум лежит приблизительно в направлении луча, проведенного из вершины C_1 через центр A симплекса. В соответствии с этим предположением при применении симплексного метода продвижение к экстремуму совершается путем зеркального отражения вершины с минимальным значением отклика через противоположащую сторону (или грань) симплекса.

Таким образом, новый симплекс II образуется путем постановки опыта всего лишь в одной новой точке C_4 (см. рис. 17.6). После получения наблюдаемого значения отклика в точке C_4 снова сравнивают величины откликов в вершинах симплекса II, выбирают новую вершину с минимальным откликом и вновь отражают ее относительно противоположащей стороны, образуют симплекс III и так далее, пока симплекс не совершит полный оборот относительно одной из вершин.

Порядок работы при использовании симплексного метода состоит в следующем.

1. По уже известным правилам выбирают начальную точку C_1 , а также интервалы варьирования Δx_i для всех факторов ($i = 1, 2, \dots, n$).

2. Выбирают безразмерную величину $\rho_{\text{сим}}$ стороны (или ребра) симплекса в относительных единицах по отношению к интервалам варьирования Δx_i , наиболее просто выбрать $\rho_{\text{сим}} = 1$. Стремятся, чтобы в безразмерных единицах стороны симплекса были равны (регулярный симплекс).

3. Вычисляют координаты остальных вершин начального симплекса. Обычно для этого используют следующее правило. Через начальную точку C_1 проводят осевые линии, параллельные координатным осям, и выбирают квадрант, в котором по предположениям должен располагаться экстремум целевой функции. В начальную точку помещают вершину симплекса C_1 , а сам симплекс I располагают так, чтобы его стороны образовали с осевыми линиями равные углы, отмеченные на рис. 17.6 двойными дужками. При таком расположении начального симплекса координаты его вершин определяют с помощью матрицы, приведенной в таблице ниже, в которой даны координаты вершин $(n + 1)$ -мерного симплекса в n -факторном пространстве.

Безразмерные относительные величины p и q при таком расположении симплекса определяют по формулам

$$p = \frac{1}{n\sqrt{2}}(\sqrt{n+1} + n - 1), \quad q = \frac{1}{n\sqrt{2}}(\sqrt{n+1} - 1). \quad (17.14)$$

На рис. 17.6 показаны размеры $p\Delta x_i$ и $q\Delta x_i$ для случая $\rho_{\text{сим}} = 1$. Если принимают $\rho_{\text{сим}} \neq 1$, то Δx_i умножают еще на $\rho_{\text{сим}}$. Знаки Δx_i зависят от номера квадранта, в котором расположен начальный симплекс. Для $n = 2$ имеем $p \approx 0,966$, $q \approx 0,259$.

Факторы x_i	x_1	x_2	x_3	...	x_i	...	x_n
Вершина симплекса C	x_1	x_2	x_3	...	x_i	...	x_n
C_1	x_{10}	x_{20}	x_{30}	...	x_{i0}	...	x_{n0}
C_2	$x_{10} + p\Delta x_1$	$x_{10} + q\Delta x_2$	$x_{10} + q\Delta x_3$...	$x_{i0} + q\Delta x_i$...	$x_{n0} + q\Delta x_n$
C_3	$x_{10} + q\Delta x_1$	$x_{10} + p\Delta x_2$	$x_{10} + q\Delta x_3$...	$x_{i0} + q\Delta x_i$...	$x_{n0} + q\Delta x_n$
...
...
C_{i+1}	$x_{10} + q\Delta x_1$	$x_{10} + q\Delta x_2$	$x_{10} + q\Delta x_3$...	$x_{i0} + q\Delta x_i$...	$x_{n0} + q\Delta x_n$
...
C_{n+1}	$x_{10} + q\Delta x_1$	$x_{10} + q\Delta x_2$	$x_{10} + q\Delta x_3$...	$x_{i0} + q\Delta x_i$...	$x_{n0} + q\Delta x_n$

4. В вершинах симплекса выполняют наблюдения отклика и сравнивают по величине; выбирают вершину с минимальным откликом и отражают ее относительно противоположной стороны или грани; находят вершину следующего симплекса II, n вершин которого одновременно являются и вершинами предыдущего симплекса I. Координаты отраженной вершины вычисляют по формуле

$$x_{il,k+1} = \frac{2}{n}(x_{i1k} + x_{i2k}) \pm x_{ilk}, \quad (17.15)$$

где i – номер фактора ($i = 1, 2, \dots, n$); l – номер вершины k -го симплекса, где обнаружен минимальный отклик; $k + 1$ – номер последующего симплекса, содержащего отраженную вершину (ей условно присваивают тот же номер l); n – число факторов.

Если минимальный отклик оказался сразу в двух вершинах, то вопрос, какую из них отражать, решают произвольно, например с помощью подбрасывания монеты.

5. Ставят эксперимент в отраженной вершине нового симплекса и отклик в ней сравнивают с откликами в остальных n вершинах, а затем снова выбирают вершину с минимальным откликом и отражают ее через противоположащую сторону (или грань) симплекса. Если в новой вершине $(k + 1)$ -го симплекса отклик оказался опять минимальным, то возвращаются к k -му симплексу и отражают вторую по минимальности вершину. Если это явление повторяется, то отражают третью по минимальности вершину и т. д.

6. Эксперимент продолжают до тех пор, пока симплекс не совершит полный оборот вокруг одной из вершин. На рис. 17.6 это вершина C_{11} . Очевидно, что точность нахождения точки экстремума зависит от двух причин: размера симплекса и влияния помех. Для уточнения положения экстремальной точки статического объекта в последних симплексах рекомендуется ставить параллельные опыты, чтобы снизить влияние помех, а также выполнить опыт в середине того симплекса, в вершинах которого отклик оказался максимальным по сравнению с остальными симплексами.

Достоинства симплексного метода: 1) достаточно высокая помехоустойчивость в смысле выбора направления движения к экстремуму; 2) изучение поверхности отклика сочетается с одновременным рабочим движением к экстремуму; 3) при оптимально выбранном размере симплекса обеспечивается высокая скорость выхода к области экстремума; 4) высокая оперативность, особенно для непрерывной оптимизации объектов с дрейфующим экстремумом.

Недостатки: 1) относительно высокая сложность вычисления координат вершин симплекса, однако при наличии электронно-цифровой вычислительной машины (ЭЦВМ) этот недостаток легко преодолевается; 2) невозможность непосредственного получения математического описания изучаемого участка поверхности отклика, как, например, в методе Бокса – Уилсона; 3) в условиях пологих по-

верхностей отклика симплексный метод дает менее точное решение, чем метод крутого восхождения.

Метод случайного поиска

Основная идея метода случайного поиска заключается в том, что точку каждого пробного опыта для изучения поверхности отклика в районе базовой (начальной) точки выбирают случайным образом,

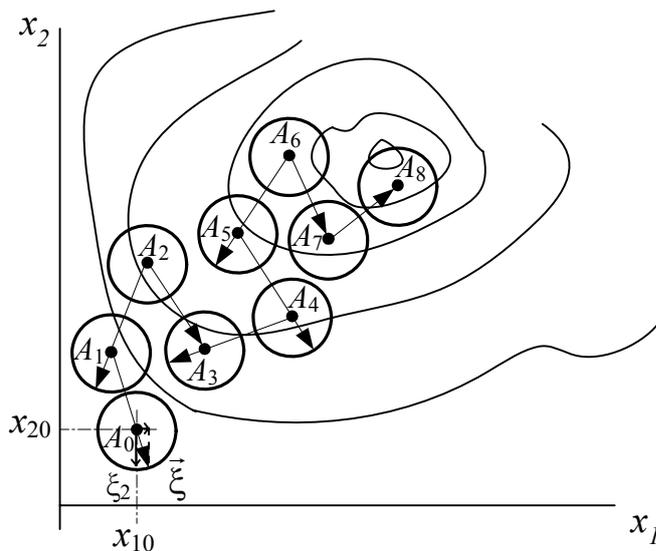


Рис. 17.7

отсюда и название метода (рис. 17.7). Несмотря на произвольность выбора пробной точки, алгоритм случайного поиска позволяет последовательно приближаться к экстремальной области. Опыты производят в исходной (начальной) и случайно выбранной пробной точках, измерения отклика в них сравнивают и, если ищется максимум, совершают рабочий шаг в направлении возрастания целевой функции.

Новую рабочую точку принимают за новую начальную и снова выбирают пробную точку случайным образом. Обычно длина рабочего шага превышает интервал варьирования между нулевой и пробной точкой.

Порядок работы при использовании метода случайного поиска состоит в следующем.

1. Выбирают начальную точку и размер случайного вектора E такой, что $|\xi| = \rho_{\text{сл.п}}$. Так как $\rho_{\text{сл.п}}$ должен быть безразмерной величиной, то сначала переходят к нормированному факторному пространству, причем за нормированные единицы варьирования принимают некоторые условные интервалы варьирования Δx_i по каждому фактору x_i ($i = 1, 2, \dots, n$). За такие условные интервалы варьирования могут быть приняты, например, абсолютные погрешности измерения δx_i по каждому i -му фактору или более крупные отрезки.

2. Определяют все n составляющих случайного вектора $\vec{\xi}$, началом которого служит нулевая точка (на рис. 17.7 точка A_0), а конец

вектора $\vec{\xi}$ равномерно распределен по окружности (или сфере) с радиусом $\rho_{\text{сл.п}}$ и центром в нулевой точке. Для этого используют таблицу равномерно распределенных случайных чисел. Пусть, например, принято $\rho_{\text{сл.п}} = 15$. В таблице случайных чисел [14] случайным образом выбирают начало отсчета и находят первое попавшееся число из интервала $(0,15)$, скажем 5 (оно находится в 6-м столбце и 9-й строке). Это и есть первая составляющая вектора $\vec{\xi}$, т.е. ξ_1 . Знак перед ξ_1 устанавливают также с помощью таблицы случайных чисел. Если перед числом 5 в столбце стоит число четное (в данном случае это 02), то перед ξ_1 ставят плюс, т. е. $\xi_1 = +5$. Если бы перед числом 5 было нечетное число, то перед ξ_1 стоял бы минус.

В том случае, когда число факторов $n = 2$, вторая составляющая оценивается однозначно по теореме Пифагора:

$$\xi_2 = \pm \sqrt{\rho_{\text{сл.п}}^2 - \xi_1^2}. \quad (17.16)$$

Знак перед ξ_2 устанавливают также по таблице случайных чисел: если после числа 5 в столбце стоит нечетное число (в данном случае это 03), то ставят минус, т.е. $\xi_2 = \sqrt{15^2 - 5^2} = -14$. Если бы после числа 5 было четное число, то перед ξ_2 стоял бы плюс.

Пусть в общем случае имеется n факторов, тогда после выбора ξ_1 продолжают выбирать из той же таблицы $\xi_2, \xi_3, \dots, \xi_{n-1}$, используя условие

$$|\xi_{k+1}| \leq \sqrt{\rho_{\text{сл.п}}^2 - \xi_1^2 - \xi_2^2 - \dots - \xi_k^2}$$

(где $k = 1, 2, \dots, n - 2$) и устанавливая их знаки по предыдущему случайному числу. Последнюю составляющую ξ_n определяют однозначно по теореме Пифагора:

$$\xi_n = \pm \sqrt{\rho_{\text{сл.п}}^2 - \xi_1^2 - \xi_2^2 - \dots - \xi_{n-1}^2}; \quad (17.17)$$

знак ξ_n устанавливают по числу, следующему в таблице за ξ_{n-1} . Составляющую ξ_i откладывают от начальной точки A_0 (в соответствии с присвоенным ей знаком) в отрицательном или положительном направлении от начальной точки параллельно своей факторной оси x_i . В данном примере $\xi_1 = +5$, $\xi_2 = -14$, поэтому на рис. 17.7 составляющая ξ_1 отложена вправо, а ξ_2 – вниз.

3. В начальной точке A_0 и в точке, служащей концом вектора $\vec{\xi}$, выполняют пробные опыты, полученные значения отклика сравнивают. Если отклик в исходной точке меньше, чем в конце вектора $\vec{\xi}$, то совершают рабочий шаг по направлению этого вектора, а если, на-

оборот, в начальной точке A_n отклик больше, чем в конце вектора $\vec{\xi}$, то рабочий шаг делают в противоположном направлении. Размер рабочего шага $a_{\text{сл.п}}$ обычно выбирают из условия

$$a_{\text{сл.п}} \geq \rho_{\text{сл.п}}. \quad (17.18)$$

В данном примере рабочий шаг совершен в точку A_1 (см. рис. 17.7).

4. Точку A_1 принимают за новую начальную точку и снова определяют составляющие нового случайного вектора (по уже описанной процедуре).

5. Если после k -го рабочего шага пробные опыты в очередной базовой точке и в конце пробного вектора $\vec{\xi}_k$ дадут равные значения отклика, то направление рабочего шага выбирают случайно – вдоль или против направления вектора $\vec{\xi}_k$. Если на k -м шаге отклики в новой базовой точке и в пробной точке конца вектора $\vec{\xi}$ меньше, чем отклик в старой базовой точке на $(k - 1)$ -м шаге, то возвращаются в $(k - 1)$ -ю начальную точку и выбирают другое случайное направление вектора $\vec{\xi}_{k-1}$.

6. Критерием выхода к экстремуму является тот факт, что во все стороны от очередной базовой точки, т. е. в любой точке окружности или сферы с центром в этой базовой точке, отклик оказывается меньшим, чем в достигнутой базовой точке. Тогда случайные пробные точки следует дополнить *регулярно выбранными*, чтобы окружность или сфера оказалась достаточно хорошо обследованной. В случае необходимости и при наличии возможности для этого ставят параллельные опыты.

Достоинства метода случайного поиска: 1) выбор случайного вектора $\vec{\xi}$ для выполнения пробного опыта не зависит от случайных помех ε и формы поверхности отклика; 2) простота алгоритма позволяет легко реализовать его в машинном эксперименте; 3) возможность простого введения в алгоритм поиска операции самообучения существенно повышает эффективность метода; 4) метод особенно эффективен для оптимизации многофакторных объектов в условиях большого числа ограничений, что особенно важно при проектировании новых объектов; 5) с ростом числа n факторов эффективность метода возрастает.

Недостатки: 1) в общем случае направление рабочих шагов не является оптимальным; 2) при отсутствии программ самообучения метод существенно менее эффективен, чем методы крутого восхождения и симплексный; 3) метод мало эффективен в условиях пологих поверхностей отклика.

Лекция 18

ОПТИМИЗАЦИЯ МЕТОДОМ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Метод относится к методам математического программирования (метод нелинейного программирования, динамического программирования и др.).

Пусть известна целевая функция

$$y = c_1x_1 + c_2x_2 + \dots + c_nx_n,$$

где c_1, c_2, \dots, c_n – постоянные коэффициенты; x_1, x_2, \dots, x_n – независимые параметры.

Требуется найти значения параметров x , при которых целевая функция y принимает экстремальное значение при ограничениях вида

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n > 0,$$

$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n > 0,$$

.....

$$x_1 \geq 0; x_2 \geq 0; \dots; x_n \geq 0.$$

Видим, что целевая функция и функции связи линейны.

Параметры x , которые удовлетворяют функциям ограничений, называются *допустимым решением*, а значения параметров x из области допустимых значений, при которых целевая функция принимает экстремальные значения, называются *оптимальным решением*.

Задачи линейного программирования решаются графическим или симплексным методом. Рассмотрим пример, основанный на применении графического метода. С решением задач симплексным методом можно ознакомиться, например, в учебнике [10].

Пример. Будем считать, что целевая функция y связана с параметрами x выражением

$$y = c_1x_1 + c_2x_2, \tag{18.1}$$

а функция ограничений имеет вид

$$\begin{aligned} a_1x_1 + a_2x_2 &\geq b, \\ x_1 \geq 0; x_2 &\geq 0. \end{aligned} \tag{18.2}$$

Из выражения (18.2) можно записать

$$x_1 \geq \frac{b}{a_1} - \frac{a_2}{a_1}x_2 \tag{18.3}$$

или

$$x_2 \geq \frac{b}{a_2} - \frac{a_1}{a_2} x_1. \quad (18.4)$$

Графически эти выражения для предельного случая можно представить отрезком прямой AB на рис. 18.1. Область допустимых значений показана точками.

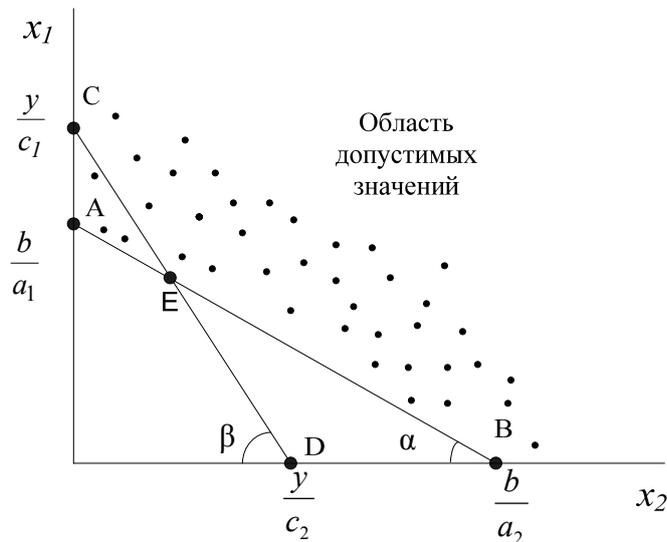


Рис. 18.1

В области допустимых значений нужно найти значения x_1, x_2 , при которых значение y будет экстремальным.

Из выражения (18.1) можем записать

$$x_1 \geq \frac{y}{c_1} - \frac{c_2}{c_1} x_2 \quad (18.5)$$

или

$$x_2 \geq \frac{y}{c_2} - \frac{c_1}{c_2} x_1. \quad (18.6)$$

Для предельного случая эти выражения можно представить отрезком CD .

Можем приравнять уравнения (18.3) и (18.5)

$$\frac{b}{a_1} - \frac{a_2}{a_1} x_2 = \frac{y}{c_1} - \frac{c_2}{c_1} x_2.$$

Запишем это уравнение в виде

$$\frac{y}{c_1} = \frac{b}{a_1} + \left(\frac{c_2}{c_1} - \frac{a_2}{a_1} \right) x_2. \quad (18.7)$$

Заметим, что

$$\frac{a_2}{a_1} = \operatorname{tg}\alpha; \quad \frac{c_2}{c_1} = \operatorname{tg}\beta.$$

Возможны три случая.

1. Угол $\beta > \alpha$, т.е. $\frac{c_2}{c_1} > \frac{a_2}{a_1}$. Если выполняется это условие, то y

будет иметь значение тем меньше, чем меньше x_2 . Следовательно, при $x_2 = 0$ y минимально

$$y_{\min} = c_1 \frac{b}{a_1};$$

$$x_1 = \frac{b}{a_1}; \quad x_2 = 0.$$

В этом случае минимальное значение будет в вершине C многоугольника ограничений.

2. Угол $\beta < \alpha$, т.е. $\frac{c_2}{c_1} < \frac{a_2}{a_1}$.

Из уравнения (18.4) и (18.6) запишем

$$\frac{b}{a_2} - \frac{a_1}{a_2} x_1 = \frac{y}{c_2} - \frac{c_1}{c_2} x_1$$

или

$$\frac{y}{c_2} = \frac{b}{a_2} + \left(\frac{c_1}{c_2} - \frac{a_1}{a_2} \right) x_1. \quad (18.8)$$

Учитывая неравенство $\frac{c_1}{c_2} > \frac{a_1}{a_2}$, y будет минимально при $x_1 = 0$:

$$y_{\min} = c_2 \frac{b}{a_2}; \quad x_1 = 0; \quad x_2 = \frac{b}{a_2}.$$

Минимальное значение y будет в вершине D многоугольника ограничений.

3. Угол $\beta = \alpha$, т.е. $\frac{c_2}{c_1} = \frac{a_2}{a_1}$.

Из уравнения (18.7) получаем

$$y_{\min} = c_1 \frac{b}{a_1}.$$

Так как

$$\frac{c_1}{c_2} = \frac{a_1}{a_2},$$

из уравнения (18.8) запишем

$$y_{\min} = c_2 \frac{b}{a_2}.$$

Минимум функции y будет находиться в вершинах многоугольника, как и в предыдущих случаях. Эти минимальные значения будут также на всем отрезке $[C; D]$, функция y постоянна.

Таким образом, независимо от значений параметров a и c целевая функция будет иметь значение в одной из вершин многоугольника ограничений. Поэтому достаточно найти значение целевой функции в этих вершинах, экстремальное значение в них будет соответствовать экстремуму целевой функции, а параметры x , соответствующие этому значению y , будут оптимальными.

Пример. Предприятия должны выпускать два вида продукции, используя при этом четыре группы производственного оборудования. Выпуск одного комплекта продукции A обеспечивает прибыль 200 000 руб., а комплекта B – 300 000 руб. Трудоемкость изготовления одного комплекта продукции вида A и B , а также фонд времени каждой группы оборудования приведен в таблице.

Группа производственного оборудования	Норма времени на выпуск одного комплекта, дни		Фонд времени в месяц, дни
	продукции A	продукции B	
I	3	3	15
II	3	6	18
III	4	0	16
IV	1	2	8

Определить, какое количество продукции видов A и B необходимо выпустить для получения максимальной прибыли.

Составим целевую функцию, обозначив x_1 – количество комплектов продукции вида A ; x_2 – количество комплектов продукции вида B :

$$\Pi = 2x_1 + 3x_2.$$

Функции ограничений будут иметь вид

$$3x_1 + 3x_2 \leq 15, \quad (1)$$

$$2x_1 + 6x_2 \leq 18, \quad (2)$$

$$4x_1 \leq 16, \quad (3)$$

$$x_1 + 2x_2 \leq 8. \quad (4)$$

Учтем также, что $x_1 \geq 0$; $x_2 \geq 0$.

Задача представляет собой линейное программирование. Нужно найти вершины многоугольника ограничений.

Неравенство (1) ограничивает область допустимых решений, лежащую ниже отрезка $С$, координатами концов $x_1=5$, $x_2=5$.

Неравенство (2) – координатами концов $x_1=9$, $x_2=3$.

Неравенство (3) – $x_1=4$, $x_2=0$.

Неравенство (4) – координатами концов $x_1=8$, $x_2=4$.

Таким образом, многоугольник ограничений имеет вершины A , B , C и D (рис. 18.2).

Рассчитаем прибыль в каждой вершине:

– для A (4; 0)

$$\Pi = 200 \cdot 4 + 300 \cdot 0 = 800 \text{ тыс. руб.};$$

– для B (4; 1)

$$\Pi = 200 \cdot 4 + 300 \cdot 1 = 1100 \text{ тыс. руб.};$$

– для C (3; 2)

$$\Pi = 200 \cdot 3 + 300 \cdot 2 = 1200 \text{ тыс. руб.};$$

– для D (0; 3)

$$\Pi = 200 \cdot 0 + 300 \cdot 3 = 900 \text{ тыс. руб.}$$

Видим, что максимальная прибыль будет при $x_1=3$ и $x_2=2$. Она составляет 1200 тыс. руб.

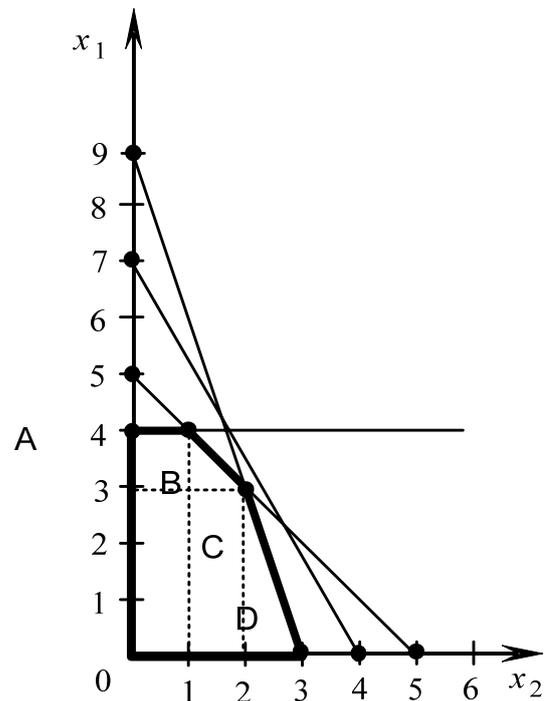


Рис. 18.2

МАТЕМАТИЧЕСКАЯ ТЕОРИЯ ЭКСПЕРИМЕНТА

Лекция 19

ВВЕДЕНИЕ

Эксперимент широко используется в различных отраслях человеческой деятельности. Как наука теория эксперимента зародилась в XIX в. Основой ее является математическая статистика (яркие представители – Фишер, Бокс, Уилсон, а из соотечественников – В.В. Налимов, Ю.П. Адлер и др.).

Для описания трудноорганизуемых (сложных) систем требовался математический аппарат, а поскольку на основе физических представлений его получить очень сложно, использовали формальный способ, основанный на эксперименте.

Можно выделить пассивный и активный эксперименты [4].

Пассивный широко применяется уже довольно давно (например наблюдение за ходом техпроцесса).

Недостатки пассивного эксперимента:

- 1) однофакторность;
- 2) пассивность (невозможность активно влиять на него).

Активный эксперимент является многофакторным, то есть в ходе него одновременно может изменяться несколько параметров, что позволяет сокращать объем эксперимента.

Для активного факторного эксперимента характерны следующие признаки:

- одновременное варьирование всеми переменными по специальным правилам – алгоритмам;
- использование математического аппарата, формализующего многие действия экспериментатора;
- выбор четкой стратегии, позволяющей принимать правильные решения после каждой серии эксперимента.

Задачи, для которых используется планирование эксперимента:

- 1) поиск оптимальных условий;
- 2) построение интерполяционных формул;
- 3) определение существенных факторов процесса;
- 4) оценка констант теоретической модели;
- 5) определение механизма явления и др.

Эксперимент реализуется следующим образом:

1) организуется полный перебор всех состояний процесса, при этом находятся значения процесса и выбирается наилучшее (количество экспериментов при этом велико);

2) определяется состояние процесса в нескольких точках, выбранных иногда случайным образом, и экстремальное значение считается наилучшим (оптимальным);

3) на основе данных эксперимента строится математическая модель, которая используется для поиска экстремума (наилучший путь).

Для математической модели используют полиномы:

- линейная функция

$$y = b_0 + b_1x_1 + b_2x_2;$$

- некоторая нелинейность

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2;$$

- существенная нелинейность

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{11}x_1^2 + b_{22}x_2^2.$$

Чем выше степень полинома, тем точнее он описывает процесс, но тем большее количество опытов необходимо провести для определения коэффициентов этой модели. Установление адекватности модели – важная задача.

На практике используется множество планов экспериментов (дисперсионного анализа, для изучения поверхности отклика, отсеивающего эксперимента, для эксперимента в условиях дрейфа, для изучения механизма явления, многофакторного анализа и др.).

Выборочный метод. Точечные и интервальные оценки параметров распределения

Точечные оценки должны быть состоятельными, несмещенными и эффективными. Оценка *состоятельная*, если при увеличении числа опытов она приближается к оцениваемому параметру, *несмещенная*, если математическое ожидание оценки равняется параметру, и *эффективная*, если обладает наименьшей дисперсией по сравнению с другими оценками.

Этими свойствами обладают выборочное математическое ожидание m_n^* для математического ожидания m_x :

$$m^* = \frac{\sum_{i=1}^n x_i}{n} \text{ – оценка математического ожидания;}$$

и выборочная дисперсия D_x^* для дисперсии D_x :

$$D^* = \frac{\sum_{i=1}^n (x_i - m_x)^2}{n-1} \quad \text{– оценка дисперсии.}$$

Тот факт, что для получения несмещенной оценки дисперсии D^* в знаменателе выборочной дисперсии пришлось n заменить на $n - 1$, непосредственно связан с тем, что величина m_n^* , относительно которой берутся отклонения, сама зависит от элементов выборки. Если бы в формуле выборочной дисперсии были две такие величины, то n нужно было бы заменить на $n - 2$ и т. д.

Каждая величина, зависящая от элементов выборки и участвующая в формуле выборочной дисперсии, называется *связью*. Оказывается (это можно строго доказать), что *знаменатель выборочной дисперсии всегда равен разности между объемом выборки и числом связей, наложенных на эту выборку*. Эта разность фактически показывает, какое количество элементов выборки можно произвольно заменять, не нарушая связей, поэтому она называется *числом степеней свободы* выборки. Число степеней свободы участвует не только в формуле выборочной дисперсии, но и в формулах всех случайных величин, так или иначе связанных с этой дисперсией.

Достоинством точечных оценок является то, что они позволяют быстро оценить значение параметров, а недостатком – невозможность определения точности, для чего используются интервальные оценки.

Интервальная оценка математического ожидания при нормальном законе распределения [11]. В соответствии с теоремой Чебышева при увеличении числа опытов среднее арифметическое m^* приближается к значению математического ожидания m . Это можно записать в виде

$$P[|m - m^*| \leq \varepsilon] = \gamma,$$

где ε – малая величина, γ – *доверительная вероятность*. Это выражение можно записать в виде

$$P(m^* - \varepsilon \leq m \leq m^* + \varepsilon) = \gamma, \tag{19.1}$$

где

$$m^* - \varepsilon \leq m \leq m^* + \varepsilon \tag{19.2}$$

является *доверительным интервалом*.

Учитывая, что закон распределения нормальный, выражение (19.1) преобразуем к виду

$$P(m - \varepsilon < m < m + \varepsilon) = \gamma = \Phi\left(\frac{m + \varepsilon - m}{\sigma_m}\right) - \Phi\left(\frac{m - \varepsilon - m}{\sigma_m}\right) = \Phi\left(\frac{\varepsilon}{\sigma_m}\right) - \Phi\left(-\frac{\varepsilon}{\sigma_m}\right),$$

где $\Phi(\dots)$ – функция Лапласа, σ_m – среднеквадратическое отклонение математического ожидания.

Учитывая, что $\Phi\left(\frac{-\varepsilon}{\sigma_m}\right) = 1 - \Phi\left(\frac{\varepsilon}{\sigma_m}\right)$, получим $\gamma = -1 + 2\Phi\left(\frac{\varepsilon}{\sigma_m}\right)$

или

$$\Phi\left(\frac{\varepsilon}{\sigma_m}\right) = \frac{\gamma + 1}{2}.$$

Так как $\sigma_m = \frac{\sigma_x}{\sqrt{n}}$ (см. выражение 19.1), получим

$$\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma_x}\right) = \frac{\gamma + 1}{2}. \quad (19.3)$$

Обозначим

$$Z_\gamma = \frac{\varepsilon\sqrt{n}}{\sigma_x}.$$

Тогда

$$\varepsilon = Z_\gamma \frac{\sigma_x}{\sqrt{n}}.$$

Подставляя ε в выражение для доверительного интервала (19.2), получим

$$m^* - Z_\gamma \frac{\sigma_x}{\sqrt{n}} \leq m \leq m^* + Z_\gamma \frac{\sigma_x}{\sqrt{n}},$$

где n – число опытов, Z_γ – аргумент функции Лапласа, определяемый по таблице функций Лапласа.

Если количество опытов $n > 30$, то $\sigma_x \approx \sigma_x^*$.

Если $n < 30$, то используется не распределение Гаусса, а распределение Стьюдента, то есть распределение случайной величины

$$t_\gamma = \frac{\varepsilon\sqrt{n}}{\sigma_x^*}.$$

Отсюда $\varepsilon = t_\gamma \frac{\sigma_x^*}{\sqrt{n}}$ и доверительный интервал $m^* - t_\gamma \frac{\sigma_x^*}{\sqrt{n}} \leq m \leq m^* + t_\gamma \frac{\sigma_x^*}{\sqrt{n}}$.

Квантиль t_γ берется из таблицы [14] в зависимости от доверительной вероятности γ и числа степеней свободы $f = n - 1$.

Лекция 20

ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ ДЛЯ ДИСПЕРСИИ НОРМАЛЬНО РАСПРЕДЕЛЕННОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Дисперсию генеральной совокупности σ_x^2 нормально распределенной случайной величины можно оценить, если известно распределение ее оценки – *выборочной дисперсии* S_x^{2*} . Распределение выборочной дисперсии можно получить при помощи распределения Пирсона (χ^2 -распределения). Если имеется выборка n независимых наблюдений x_1, x_2, \dots, x_n над нормально распределенной случайной величиной, то можно показать, что сумма

$$\chi^2 = \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma_x} \right)^2 \quad (20.1)$$

имеет распределение χ^2 с $f = n - 1$ степенями свободы.

Плотность χ^2 -распределения

$$\varphi(\chi^2) = \frac{1}{2^{f/2} \Gamma(f/2)} (\chi^2)^{\frac{f-2}{2}} e^{-\frac{\chi^2}{2}}, \quad 0 \leq \chi^2 \leq \infty,$$

где $\Gamma(f)$ – гамма-функция; χ^2 зависит только от числа степеней свободы f .

На рис. 20.1 приведены кривые плотности вероятности χ^2 -распределения при некоторых значениях f . Кривые асимметричны, степень асимметрии уменьшается с увеличением f . При доверительной вероятности $\beta = 1 - p$ двусторонняя доверительная оценка для χ^2 имеет вид

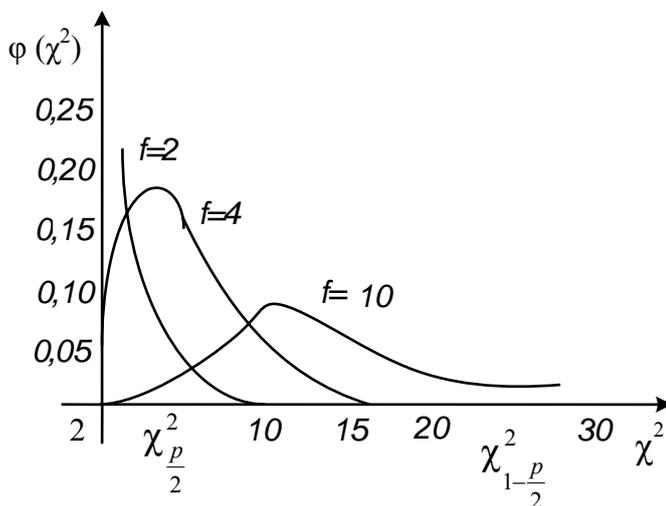


Рис. 20.1

$$\chi_{p/2}^2 \leq \chi^2 \leq \chi_{1-p/2}^2, \quad (20.2)$$

односторонние оценки имеют вид

$$\chi^2 \leq \chi_{1-p}^2, \chi^2 \geq \chi_p^2.$$

Значение квантилей χ^2_{1-p} при различных p и f приведены в таблице [14]. Так как выборочная дисперсия S_x^2 через элементы выборки определяется по формуле

$$S_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{f}, \quad (20.3)$$

сравнивая (20.1) и (20.3), получим

$$\chi^2 = \frac{fS_x^2}{\sigma_x^2}. \quad (20.4)$$

Подставляя (20.4) в (20.2), получим

$$\chi_{p/2}^2 \leq \frac{fS_x^2}{\sigma_x^2} \leq \chi_{1-p/2}^2.$$

Решая неравенство относительно σ_x^2 , получим доверительные двусторонние границы для генеральной дисперсии σ_x^2 :

$$\frac{fS_x^2}{\chi_{1-p/2}^2} \leq \sigma_x^2 \leq \frac{fS_x^2}{\chi_{p/2}^2}.$$

Аналогично получают односторонние доверительные оценки

$$\sigma_x^2 \leq \frac{fS_x^2}{\chi_{p/2}^2}, \sigma_x^2 \geq \frac{fS_x^2}{\chi_{1-p/2}^2}.$$

Сравнение двух дисперсий

При обработке наблюдений часто возникает необходимость сравнить две или несколько выборочных дисперсий. Основная гипотеза, которая при этом проверяется: можно ли считать сравниваемые выборочные дисперсии оценками одной и той же генеральной дисперсии. Рассмотрим две выборки

$$x_1, x_2, \dots, x_{n1}, \\ y_1, y_2, \dots, y_{n2},$$

средние значения которых соответственно равны \bar{x} и \bar{y} . Выборочные дисперсии

$$S_1^2 = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{n_1 - 1};$$

$$S_2^2 = \frac{\sum_{i=1}^{n_2} (y_i - \bar{y})^2}{n_2 - 1}$$

определяются со степенями свободы

$$f_1 = n_1 - 1, \quad f_2 = n_2 - 1.$$

Требуется выяснить, являются ли выборочные дисперсии S_1^2 и S_2^2 значимо различными или же полученные выборки можно рассматривать как взятые из генеральных совокупностей с равными дисперсиями. Предположим, что первая выборка сделана из генеральной совокупности с дисперсией σ_1^2 , а вторая – из генеральной совокупности с дисперсией σ_2^2 ; и проверяется гипотеза о равенстве генеральных дисперсий $\sigma_1^2 = \sigma_2^2$. Чтобы отвергнуть эту гипотезу, нужно доказать значимость различия между S_1^2 и S_2^2 при выбранном уровне значимости p . В качестве критерия значимости обычно используется критерий Фишера. *Распределением Фишера* (F -распределение) называется распределение случайной величины

$$F = \frac{(S_1^2 / \sigma_1^2)}{(S_2^2 / \sigma_2^2)}. \quad (20.5)$$

Плотность вероятности F -распределения определяется выражением

$$\varphi(F) = \frac{\Gamma\left(\frac{f_1 + f_2}{2}\right)}{\Gamma\left(\frac{f_1}{2}\right)\Gamma\left(\frac{f_2}{2}\right)} f_1^{\frac{f_2}{2}} f_2^{\frac{f_1}{2}} \frac{F^{\frac{f_1-2}{2}}}{(f_2 + f_1 F)^{\frac{f_1+f_2}{2}}},$$

где $\Gamma(f)$ – гамма-функция.

F -распределение зависит только от числа степеней свободы f_1 и f_2 . На рис. 20.2 приведены кривые плотности вероятности F -распределения для некоторых значений f_1 и f_2 . Кривые имеют асимметричную форму. В таблице [14] приведены квантили F_{1-p} для разных уровней значимости p , чисел степеней свободы f_1 и f_2 . Для определения квантилей F_p для значений p , не вошедших в таблицу, используется соотношение

$$F_p(f_1, f_2) = \frac{1}{F_{1-p}(f_2, f_1)}. \quad (20.6)$$

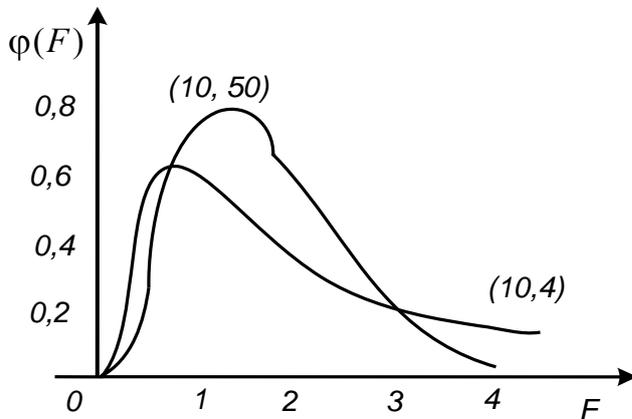


Рис. 20.2

При $\sigma_1^2 = \sigma_2^2$ и $\frac{\sigma_1^2}{\sigma_2^2} = 1$ F -распределение может быть непосредственно использовано для оценки отношения выборочных дисперсий $\frac{S_1^2}{S_2^2}$.

При доверительной вероятности $1-p$ двусторонняя доверительная оценка величины F имеет вид

$$F_{\frac{p}{2}}(f_1, f_2) \leq F \leq F_{1-\frac{p}{2}}(f_1, f_2).$$

С учетом (20.6)

$$\frac{1}{F_{1-\frac{p}{2}}(f_1, f_2)} \leq F \leq F_{1-\frac{p}{2}}(f_1, f_2).$$

При $\sigma_1^2 = \sigma_2^2$ $F = \frac{S_1^2}{S_2^2}$, следовательно, с вероятностью $1-p$ должно выполняться неравенство

$$\frac{1}{F_{1-\frac{p}{2}}(f_1, f_2)} \leq \frac{S_1^2}{S_2^2} \leq F_{1-\frac{p}{2}}(f_1, f_2).$$

или одно из односторонних неравенств

$$\frac{S_1^2}{S_2^2} \leq F_{1-p}(f_1, f_2) \quad (\text{оценка сверху}),$$

$$\frac{S_1^2}{S_2^2} \geq \frac{1}{F_{1-p}(f_1, f_2)} \quad (\text{оценка снизу}).$$

Сравнение нескольких дисперсий

При определении оценки дисперсии по текущим измерениям по формуле

$$S_y^2 = \frac{f_1 S_1^2 + f_2 S_2^2 + \dots + f_n S_n^2}{f_1 + f_2 + \dots + f_n} = \frac{\sum_{i=1}^n f_i S_i^2}{f},$$

где $f_i = n_i - 1$; была принята нулевая гипотеза равенства соответствующих генеральных дисперсий. Проверить эту гипотезу для выборок разного объема можно по критерию Бартлета. Бартлет показал, что в условиях нулевой гипотезы отношение B/C , где

$$B = 2,303 \left(f \lg S_y^2 - \sum_{i=1}^n f_i \lg S_i^2 \right),$$

$$C = 1 + \frac{1}{3(n-1)} \left(\sum_{i=1}^n \frac{1}{f_i} - \frac{1}{f} \right),$$

распределено приближенно, как χ^2 с $n - 1$ степенями свободы, если $f_i > 2$.

Гипотеза равенства генеральных дисперсий принимается, если

$$\frac{B}{C} \leq \chi_{1-p}^2$$

при выбранном уровне значимости p . Различие между выборочными дисперсиями можно считать незначимым, а сами выборочные дисперсии – однородными. Так как $C > 1$ при $B \leq \chi_{1-p}^2$, нулевую гипотезу следует принять; при $B > \chi_{1-p}^2$ критерий Бартлета вычисляют полностью.

Если выборочные дисперсии получены по выборкам одинаковых объемов $m_1 = m_2 = \dots = m_n = m$, для их сравнения используют более удобный и точный критерий Кохрена. Кохрен исследовал распределение максимальной выборочной дисперсии к сумме всех дисперсий:

$$G = \frac{S_{\max}^2}{\sum_{i=1}^n S_i^2}. \quad (20.7)$$

Распределение случайной величины G зависит только от числа суммируемых дисперсий n и числа степеней свободы f , с которым определена каждая дисперсия $f = m - 1$. Если найденное по выборочным дисперсиям значение критерия Кохрена окажется меньше табличного

$$G < G_{1-p}(n, f), \quad (20.8)$$

где n – число суммируемых дисперсий; $f = m - 1$; расхождение между дисперсиями нужно считать случайным при выбранном уровне значимости p . Если при этом определится оценка для дисперсии воспроизводимости, однородные дисперсии можно усреднить. Число степеней свободы f среднеарифметической дисперсии равно $f = n(m - 1)$.

Сравнение двух средних

Для сравнения между собой двух средних, полученных по выборкам из *нормально распределенных* генеральных совокупностей, применяется критерий Стьюдента или t -критерий. Пусть заданы две случайные выборки

$$x_1, x_2, \dots, x_{n_1};$$

$$y_1, y_2, \dots, y_{n_2}.$$

Первая выборка взята из нормально распределенной генеральной совокупности с параметрами m_x и σ_x^2 , вторая – из генеральной совокупности с параметрами m_y и σ_y^2 . По выборкам получены оценки для этих параметров: \bar{x}, S_x^2 и \bar{y}, S_y^2 . Требуется проверить нулевую гипотезу: $m_x = m_y$, при условии $\sigma_x^2 = \sigma_y^2 = \sigma^2$.

Рассмотрим случайную величину $z = \bar{x} - \bar{y}$. Вследствие линейности нормального распределения [2] z распределена нормально с параметрами

$$m_z = m_x - m_y,$$

$$\sigma_z^2 = \sigma_{\bar{x}}^2 + \sigma_{\bar{y}}^2 = \frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2} = \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right).$$

Составим нормированную случайную величину

$$\frac{z - m_z}{\sigma^2} = \frac{(\bar{x} - \bar{y}) - (m_x - m_y)}{\sigma^2 \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

которая имеет стандартное нормальное распределение.

Если генеральный стандарт σ заменить выборочным, получится величина, имеющая распределение Стьюдента:

$$t = \frac{(\bar{x} - \bar{y}) - (m_x - m_y)}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

с числом степеней свободы $f = n_1 + n_2 - 2$.

Однородность выборочных дисперсий S_1^2 и S_2^2 можно проверить по критерию Фишера. При доверительной вероятности $\beta = 1 - p$ имеем двустороннюю оценку для разности $m_x - m_y$:

$$\bar{x} - \bar{y} - t_{1-p/2} S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \leq m_x - m_y \leq \bar{x} - \bar{y} + t_{1-p/2} S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

или односторонние оценки

$$m_x - m_y \leq \bar{x} - \bar{y} + t_{1-p/2} S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \quad (20.9)$$

$$m_x - m_y \geq \bar{x} - \bar{y} - t_{1-p/2} S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}. \quad (20.10)$$

В условиях нулевой гипотезы $m_x = m_y$ неравенства (20.9) и (20.10) дают критерий проверки этой гипотезы. Нулевая гипотеза отвергается при двустороннем критерии, если

$$|\bar{x} - \bar{y}| \geq t_{1-p/2} S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}};$$

при одностороннем критерии, если

$$|\bar{x} - \bar{y}| \geq t_{1-p} S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$

Сравнение нескольких средних

При сравнении нескольких средних можно использовать t -критерий, проводя сравнение попарно. Однако для использования при сравнении полной информации о всех средних его проводят при помощи *множественного рангового критерия Дункана* [14]. Пусть по k выборкам разного объема получено k средних значений

$$\bar{x}_1, \bar{x}_2, \dots, \bar{x}_i, \dots, \bar{x}_k,$$

$$\bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ji}}{n_i}.$$

Генеральные дисперсии равны между собой, т. е.

$$\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \dots = \sigma_x^2.$$

Используя критерий Дункана, следует: 1) проранжировать k средних значений, расположив их в порядке возрастания; 2) определить ошибку воспроизводимости результатов S_x с соответствующим числом степеней свободы f_x ; 3) определить ошибку для каждого среднего

$$S_{\bar{x}_j} = \sqrt{\frac{S_x^2}{n_j}}, \quad j = 1, 2, \dots, k;$$

4) выписать из таблицы значимых рангов Дункана с выбранным уровнем значимости, числом f_x , $p = 2, 3, \dots, k(k-1)$ значений рангов; 5) умножить эти значения рангов на $S_{\bar{x}_y}$ и таким образом определить $(k-1)$ наименьших значимых рангов; 6) проверить значимость различия между средними, начиная с крайних в ранжированном ряду; разность максимального и минимального значений среднего сравнить с наименьшим значимым рангом при $p = k$, затем найти разность максимального среднего и второго среднего в ранжированном ряду и сравнить ее с наименьшим значимым рангом при $p = k-1$ и т. д. Это сравнение продолжить для второго по величине среднего, которое сравнивается с наименьшим, и так далее, пока не будут исследованы на значимость различия между всеми $k(k-1)/2$ парами.

Лекция 21

ОБРАБОТКА РЕЗУЛЬТАТОВ ПАССИВНОГО ЭКСПЕРИМЕНТА

Пусть $y = f(x_1, x_2, \dots, x_n)$ есть функция n переменных.

При пассивном эксперименте $n - 1$ независимых переменных фиксируют, а изменяется только одна, то есть находят зависимость $y = f(x_i)$, как показано на рис. 21.1.

Часто цель пассивного эксперимента – построение математической модели на основе регрессионного анализа. При этом нужно решить следующие задачи:

– проверить постулаты регрессионного анализа;

– построить математическую модель;
– проверить адекватность модели;
– проверить значимость коэффициентов модели (коэффициентов регрессии).

Регрессионный анализ основан на трех постулатах:

1) выходной параметр y имеет нормальный закон распределения;
2) факторы (первичные параметры) есть неслучайные величины (на практике должно выполняться условие $\sigma_{x_i}^2 \ll \sigma_y^2$);

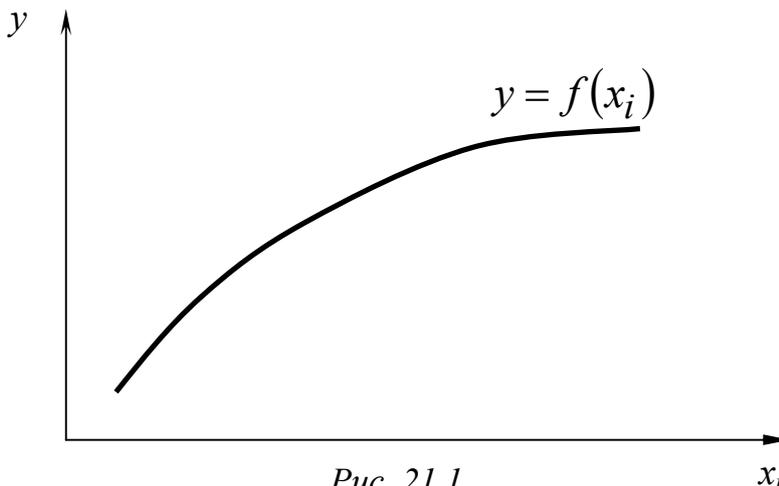
3) точность определения выходного параметра y не зависит от его величины (дисперсии y должны быть однородны в различных точках факторного пространства). Однородность дисперсии проверяется с помощью критериев Фишера, Бартлета или Кохрена.

Математические модели строятся часто в виде полиномов:

- линейных:

$$y = b_0 + b_1x_2 \text{ – однофакторное пространство;}$$

$$y = b_0 + b_1x_1 + b_2x_2 \text{ – двухфакторное пространство;}$$



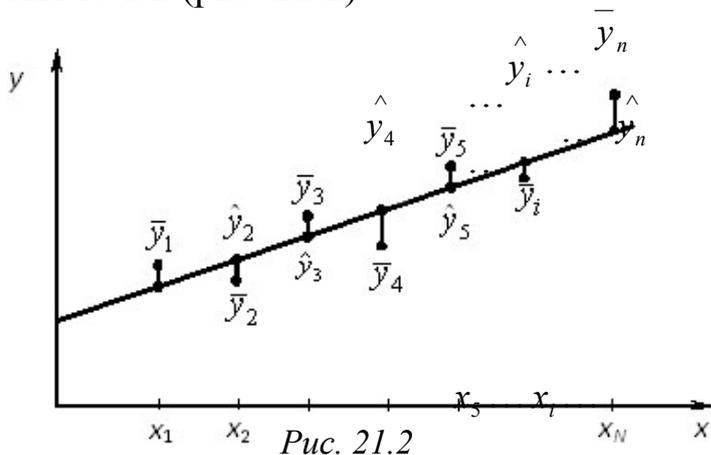
- нелинейных:

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_{12} + b_{11}x_1^2 + b_{22}x_2^2 + b_{32}x_1^2x_2 + \dots$$

Создание математической модели включает четыре этапа:

- проверку постулатов регрессионного анализа;
- выбор вида полинома и определение его коэффициентов;
- проверку значимости коэффициентов;
- проверку адекватности модели.

Определение математических моделей с использованием метода наименьших квадратов (МНК). Пусть точки \bar{y}_i соответствуют средним экспериментальным значениям в каждом сечении (при x_1, x_2 и т. д.), а точки \hat{y}_i – прямой, построенной на основе расположения этих точек (рис. 21.2).



Разность между точками на модели и средними значениями обозначим δ_i :

$$\delta_i = \hat{y}_i - \bar{y}_i.$$

При использовании МНК линия должна проводиться таким образом, чтобы сумма квадратов отклонений между средними значениями и точками модели была минимальной:

$$B = \sum_{i=1}^N \delta_i^2 \rightarrow \min.$$

Условием нахождения минимума является выполнение условий:

$$\frac{\partial B}{\partial b_0} = 0, \quad \frac{\partial B}{\partial b_1} = 0.$$

Запишем

$$B = \sum_{i=1}^N \left(\hat{y}_i - \bar{y}_i \right)^2 = \sum_{i=1}^N \left(b_0 + b_1x_i - \bar{y}_i \right)^2. \quad (21.1)$$

Производная

$$\frac{\partial B}{\partial b_0} = \sum_{i=1}^N 2 \left(b_0 + b_1x_i - \bar{y}_i \right) = 0.$$

Отсюда получим

$$Nb_0 + b_1 \sum x_i = \sum \bar{y}_i.$$

Производная

$$\frac{\partial B}{\partial b_1} = \sum_{i=1}^N 2 (b_0 + b_1 x_i - \bar{y}_i) x_i = 0.$$

Следовательно,

$$b_0 \sum x_i + b_1 \sum x_i^2 = \sum x_i \bar{y}_i.$$

Получаем систему уравнений

$$\left. \begin{aligned} Nb_0 + b_1 \sum x_i &= \sum \bar{y}_i, \\ b_0 \sum x_i + b_1 \sum x_i^2 &= \sum x_i \bar{y}_i, \end{aligned} \right\}$$

которая называется *системой нормальных уравнений*.

Решая эту систему, получим

$$b_0 = \frac{\begin{vmatrix} \sum y_i & \sum x_i \\ \sum x_i y_i & \sum x_i^2 \end{vmatrix}}{\begin{vmatrix} N & \sum x_i \\ \sum x_i & \sum x_i^2 \end{vmatrix}}, \quad b_1 = \frac{\begin{vmatrix} N & \sum \bar{y}_i \\ \sum x_i & \sum x_i \bar{y}_i \end{vmatrix}}{\begin{vmatrix} N & \sum x_i \\ \sum x_i & \sum x_i^2 \end{vmatrix}},$$

$$b_0 = \frac{\sum \bar{y}_i \sum x_i^2 - \sum x_i \bar{y}_i \sum x_i}{N \sum x_i^2 - (\sum x_i)^2},$$

$$b_1 = \frac{N \sum x_i \bar{y}_i - \sum x_i \sum \bar{y}_i}{N \sum x_i^2 - (\sum x_i)^2}.$$

Коэффициенты уравнения регрессии часто называют коэффициентами регрессии.

Проверка адекватности модели. На рис. 21.3 приведены примеры адекватной (1) и неадекватной (2) моделей. Вертикальные отрезки показывают интервал, в котором находятся экспериментальные значения.

Проверка адекватности проводится при помощи критерия Фишера

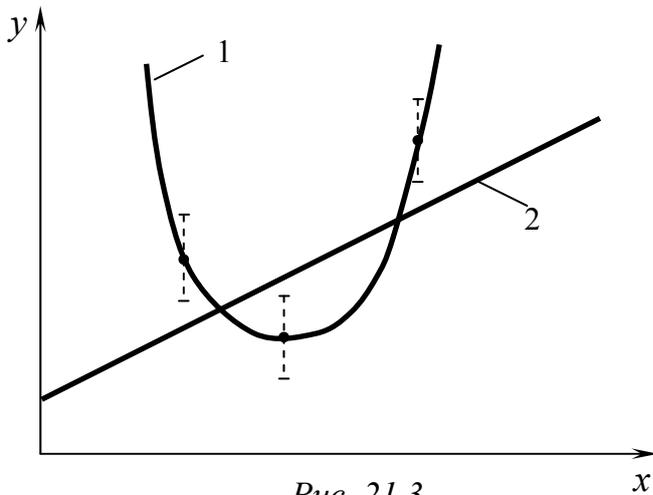


Рис. 21.3

$$F_{\Theta} = \frac{\sigma_{ад}^{*2}}{\sigma_y^{*2}},$$

где $\sigma_{ад}^{*2}$ – дисперсия адекватности; σ_y^{*2} – дисперсия воспроизводимости.

Если $F_{\Theta} < F_T$, где F_T – табличное значение [11], найденное при заданном уровне значимости p , то считается, что модель адекватна, в противном случае – неадекватна.

Степени свободы при нахождении F_T равны

$$f_1 = N - (k + 1),$$

$$f_2 = N,$$

где $k + 1$ определяет число коэффициентов регрессии, k равно числу факторов.

Дисперсия адекватности имеет вид

$$\sigma_{ад}^{*2} = \frac{\sum_{i=1}^N \delta_i^2}{N},$$

где $\delta_i = \hat{y}_i - \bar{y}_i$.

Дисперсия воспроизводимости – это дисперсия среднего значения выходного параметра. Последовательно ее можно рассчитать следующим образом.

Дисперсия y_i в каждом i -м сечении

$$\sigma_i^{*2} = \frac{\sum_{j=1}^n (y_{ij} - \bar{y}_i)^2}{n - 1},$$

где n – число параллельных опытов.

Дисперсия выходного параметра y

$$\sigma_y^{*2} = \frac{\sum_{i=1}^N \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2}{N(n - 1)}.$$

Дисперсия среднего значения выходного параметра (дисперсия воспроизводимости)

$$\sigma_{\bar{y}}^{*2} = \frac{\sigma_y^{*2}}{n}$$

или

$$\sigma_{\bar{y}}^{*2} = \frac{\sum_{i=1}^N \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2}{Nn(n-1)}.$$

Проверка значимости коэффициентов регрессии. Значимость коэффициентов регрессии заключается в сравнении их абсолютных значений с ошибками эксперимента. Считается, что коэффициент значим, если его абсолютная величина $|b_k| > \Delta b_k$, где Δb_k – ошибка при его определении, устанавливаемая по выражению

$$\Delta b_k = t_\gamma \sigma_{b_k}^*$$

где $\sigma_{b_k}^*$ – среднеквадратическое отклонение коэффициента b_k ; или по формуле

$$\Delta b_k = \frac{t_\gamma \sigma_{\bar{y}}^*}{\sqrt{N}}.$$

Лекция 22

ТЕОРИЯ АКТИВНОГО МНОГОФАКТОРНОГО ЭКСПЕРИМЕНТА

Эксперимент *активный*, если факторы изменяются по ранее оговоренному правилу (алгоритму), и *многофакторный* при условии варьирования несколькими или всеми факторами.

Пример. Взвешивание трех объектов: A , B и C (табл. 22.1).

Таблица 22.1

Традиционная схема взвешивания

Номер опыта	A	B	C	Результат взвешивания
1	–	–	–	y_1
2	+	–	–	y_2
3	–	+	–	y_3
4	–	–	+	y_4

Если объект находится на весах, то в табл. 22.1 стоит знак «+», если нет – то знак «–».

Веса $P(A) = y_2 - y_1$; $P(B) = y_3 - y_1$; $P(C) = y_4 - y_1$.

Дисперсии $D[P(A)] = D[y_2 - y_1] = D[y_2] + D[y_1] = 2D[y]$, так как измерения проводятся на одном приборе, то $D[y_1] = D[y_2] = D[y]$.

Очевидно, что дисперсии весов объектов B и C также равны $2D[y]$.

Теперь проведём взвешивание по плану табл. 22.2.

Таблица 22.2

Предлагаемая схема взвешивания

Номер опыта	A	B	C	Y
1	+	–	–	y_1
2	–	+	–	y_2
3	–	–	+	y_3
4	+	+	+	y_4

Например, вес $P(A) = \frac{y_1 - y_2 - y_3 + y_4}{2}$, а дисперсия имеет вид

$$D[A] = \frac{D[y_1 - y_2 - y_3 + y_4]}{4} = \frac{4D[y]}{4} = D(y).$$

Следовательно, точность увеличивается в два раза, т.е. многофакторность позволяет сократить объем опыта.

Полный факторный эксперимент типа 2^k

Пусть $y = f(x_1, x_2, \dots, x_k)$. Обозначение 2^k означает, что каждый из k факторов варьируется на двух уровнях.

Полным факторным экспериментом (ПФЭ) называется эксперимент, в котором реализуются все возможные сочетания уровней факторов.

При ПФЭ сначала выбирается начальный уровень факторов $x_{10}, x_{20}, x_{30}, \dots, x_{k0}$, потом для каждого фактора определяется уровень варьирования $\Delta x_1, \Delta x_2, \Delta x_3, \dots, \Delta x_k$ и нижний и верхний значения фактора (рис. 22.1):

$$\tilde{x}_{1Н} = \frac{x_{1Н} - x_{10}}{\Delta x_1} = \frac{-\Delta x_1}{\Delta x_1} = -1;$$

$$\tilde{x}_{1В} = \frac{x_{1В} - x_{10}}{\Delta x_1} = \frac{\Delta x_1}{\Delta x_1} = +1.$$

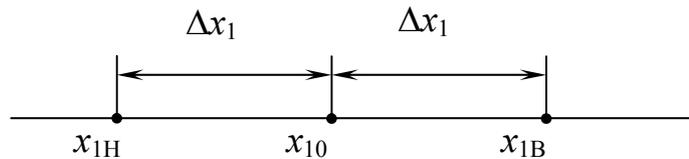


Рис. 22.1

Перед проведением ПФЭ составляется матрица планирования эксперимента (МПЭ). Пусть $y = f(x_1, x_2)$ и требуется найти коэффициенты модели $y = b_0 + b_1x_1 + b_2x_2$, тогда матрица будет иметь вид, показанный в табл. 22.3.

Таблица 22.3

Матрица ПФЭ 2^2

Номер опыта	x_1	x_2	Y
1	+	+	y_1
2	+	-	y_2
3	-	+	y_3
4	-	-	y_4

Знак «+» в матрице означает, что фактор равен «+1», то есть принимает значение на верхнем уровне, а знак «-» – на нижнем. Расположение точек показано на рис. 22.2.

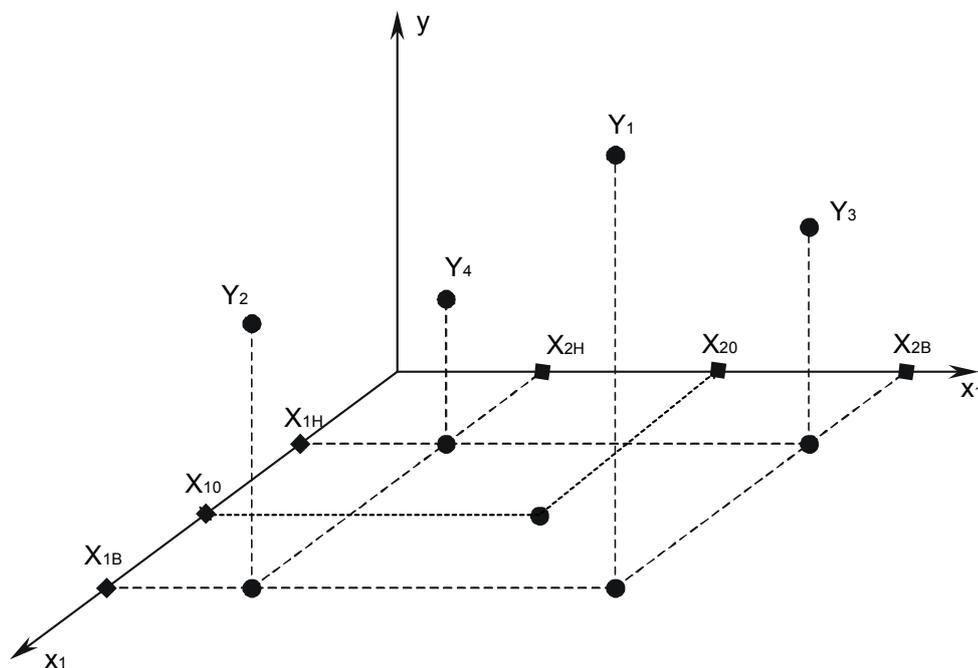


Рис. 22.2

Свойства МПЭ:

1) *симметричность*, означающая, что сумма элементов каждого столбца равна 0:

$$\sum_{i=1}^N x_{ji} = 0;$$

2) *нормированность* – сумма квадратов элементов каждого столбца равна числу опытов N (числу строк); учитываем только независимые опыты:

$$\sum_{i=1}^N x_{ji}^2 = N;$$

3) *ортогональность* – сумма почленных произведений любых двух столбцов равняется нулю:

$$\sum_{i=1}^N x_{ji} x_{ki} = 0.$$

Эти свойства легко проверить на матрице ПФЭ 2^2 , а также ПФЭ 2^3 (табл. 22.4).

Таблица 22.4

Матрица ПФЭ 2^3

Номер опыта	x_1	x_2	x_3	Y
1	–	–	–	y_1
2	+	–	–	y_2
3	–	+	–	y_3
4	–	–	+	y_4
5	+	+	–	y_5
6	+	–	+	y_6
7	–	+	+	y_7
8	+	+	+	y_8

Свойства матрицы ПФЭ позволяют значительно упростить формулы для определения коэффициентов регрессии. В соответствии с МНК, если $y = b_0 + b_1x_1 + b_2x_2$, запишем

$$B = \sum_{i=1}^N E_i^2 \rightarrow \min, \text{ где } E_i = b_0 + b_1x_{1i} + b_2x_{2i} - \bar{y}_i.$$

Производные

$$\frac{\partial B}{\partial b_0} = 0, \quad \frac{\partial B}{\partial b_1} = 0, \quad \frac{\partial B}{\partial b_2} = 0.$$

Следовательно,

$$\frac{\partial B}{\partial b_0} = \sum_{i=1}^N 2(b_0 + b_1x_{1i} + b_2x_{2i} - \bar{y}_i) = 0$$

или

$$\left. \begin{aligned} Nb_0 + b_1 \sum x_{1i} + b_2 \sum x_{2i} &= \sum \bar{y}_i, \\ b_0 \sum x_{1i} + b_1 \sum x_{1i}^2 + b_2 \sum x_{2i}x_{1i} &= \sum x_{1i}\bar{y}_i, \\ b_0 \sum x_{2i} + b_1 \sum x_{1i}x_{2i} + b_2 \sum x_{2i}^2 &= \sum x_{2i}\bar{y}_i. \end{aligned} \right\}$$

Расчетные формулы, учитывая свойства матрицы ПФЭ, принимают вид

$$b_0 = \frac{\sum \bar{y}_i}{N}; \quad b_1 = \frac{\sum x_{1i}\bar{y}_i}{N}; \quad b_2 = \frac{\sum x_{2i}\bar{y}_i}{N};$$

т. е.

$$b_1 = \frac{y_1 + y_2 - y_3 - y_4}{4}, \quad b_2 = \frac{y_1 - y_2 + y_3 - y_4}{4}, \quad b_0 = \frac{y_1 + y_2 + y_3 + y_4}{4}.$$

Для того чтобы b_0 определить по единым правилам, в МПЭ вводят дополнительный столбец x_0 (табл. 22.5).

Таблица 22.5

Матрица ПФЭ 2^2 с дополнительным столбцом

Номер опыта	x_0	x_1	x_2	Y
1	+	+	+	y_1
2	+	+	-	y_2
3	+	-	+	y_3
4	+	-	-	y_4

Возникает вопрос о возможности использования такого алгоритма в случае нелинейности модели:

$$y = b_0 + b_1x_1 + b_{12}x_1x_2,$$

где $b_{12}x_1x_2$ определяет нелинейность, называемую *эффектом взаимодействия*. Система нормальных уравнений принимает вид

$$\left. \begin{aligned} Nb_0 + b_1 \sum x_1 + b_2 \sum x_2 + b_{12} \sum x_1x_2 &= \sum y_i, \\ b_0 \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_1x_2 + b_{12} \sum x_1^2 &= \sum x_{1i}y_i, \\ b_0 \sum x_2 + b_1 \sum x_1x_2 + b_2 \sum x_2^2 + b_{12} \sum x_1x_2^2 &= \sum x_{2i}y_i, \\ b_0 \sum x_1x_2 + b_1 \sum x_1^2x_2 + b_2 \sum x_1x_2^2 + b_{12} \sum x_1^2x_2^2 &= \sum x_{1x_2i}y_i, \end{aligned} \right\}$$

учитывая свойства матрицы ПФЭ, получим

$$b_0 = \frac{\sum y}{N}, \quad b_1 = \frac{\sum x_1 y}{N}, \quad b_2 = \frac{\sum x_2 y}{N}, \quad b_{12} = \frac{\sum (x_1 x_2) y}{N}.$$

Таким образом, видно, что b_{12} может быть найден по тому же алгоритму.

Для определения коэффициентов модели

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 = 0$$

матрица ПФЭ имеет вид, показанный в табл. 22.6.

Таблица 22.6

Матрица ПФЭ 2^3 для нелинейной модели

Номер опыта	x_0	x_1	x_2	x_3	$x_1 x_2$	$x_1 x_3$	$x_2 x_3$	Y
1	+	-	-	-	+	+	+	y_1
2	+	+	-	-	-	-	+	y_2
3	+	-	+	-	-	+	-	y_3
4	+	-	-	+	+	-	-	y_4
5	+	+	+	-	+	-	-	y_5
6	+	+	-	+	-	+	-	y_6
7	+	-	+	+	-	-	+	y_7
8	+	+	+	+	+	+	+	y_8

Она обладает свойствами симметричности, нормированности и ортогональности.

Рассмотрим теперь модель $y = b_0 + b_1 x_1 + b_2 x_2^2$.

Система нормальных уравнений будет иметь вид

$$\left. \begin{aligned} Nb_0 + b_1 \sum x_1 + b_2 \sum x_2^2 &= \sum y, \\ b_0 \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_1 x_2^2 &= \sum x_1 y, \\ b_0 \sum x_2^2 + b_1 \sum x_1 x_2^2 + b_2 \sum x_2^4 &= \sum x_2^2 y, \end{aligned} \right\}$$

тогда

$$\left. \begin{aligned} Nb_0 + Nb_2 &= \sum y, \\ Nb_1 &= \sum x_1 y, \\ Nb_0 + Nb_2 &= \sum y. \end{aligned} \right\}$$

Первое и третье уравнения идентичны, а из двух уравнений невозможно определить три неизвестные.

Если привести матрицу, то видно, что столбцы при x_0 и при x_2^2 неразличимы,

$$b_0 \rightarrow \beta_0 + \beta_{22},$$

то есть наблюдается смешение эффектов (табл. 22.7).

Таблица 22.7

Матрица ПФЭ 2^2 для модели с квадратичными элементами

Номер опыта	x_0	x_1	x_2^2	Y
1	+	-	+	y_1
2	+	+	+	y_2
3	+	-	+	y_3
4	+	+	+	y_4

Следовательно, нелинейные члены, в которых факторы во второй и более высокой степени, при помощи ПФЭ определить нельзя (используются другие планы).

Дробный факторный эксперимент

ПФЭ требует проведения большого объема опытов ($k = 3, 2^3=8$, но в каждом не менее двух, т.е. всего 16 опытов).

При двух факторах на двух уровнях МПЭ будет иметь вид, показанный в табл. 22.8.

Таблица 22.8

Матрица ПФЭ 2^3

Номер опыта	x_0	x_1	x_2	$x_1 x_2$	Y
1	+	-	-	+	y_1
2	+	+	-	-	y_2
3	+	-	+	-	y_3
4	+	+	+	+	y_4

Можно определить коэффициенты модели $y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2$. Предположим, нам априорно известно, что процесс линейный, и тогда определения коэффициента b_{12} не требуется ($b_{12} = 0$). В этом случае результаты ПФЭ обладают избыточностью. Возникает вопрос, нельзя ли за счет избыточности получить дополнительную информацию или сократить число опытов?

Составим матрицу (табл. 22.9).

Таблица 22.9

Усеченная матрица ПФЭ 2^3

Номер опыта	x_0	x_1	x_2	x_3	$x_1 x_2$	$x_1 x_3$	$x_2 x_3$	Y
1	+	-	-	+	+	-	-	y_1
2	+	+	-	-	-	-	+	y_2
3	+	-	+	-	-	+	-	y_3
4	+	+	+	+	+	+	+	y_4

Эта матрица обладает свойствами симметричности, нормированности, ортогональности. Используя ее, можно рассчитать коэффициенты модели $y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3$, но

$$b_0 \rightarrow \beta_0 + \beta_{12};$$

$$b_1 \rightarrow \beta_1 + \beta_{23};$$

$$b_2 \rightarrow \beta_2 + \beta_{13}.$$

Так как заранее известно, что процесс линейный, то $\beta_{12} = \beta_{13} = \beta_{23} = 0$. Получим модель, для которой провели четыре эксперимента, сократив опыт в два раза, но с применением априорной информации. Такой дробный факторный эксперимент (ДФЭ) называется *полуреplikой*.

При ПФЭ 2^{15} проводится 32768 экспериментов, а если применяется 1/16-реплики, то достаточно поставить всего 16 экспериментов.

Лекция 23

ПРОВЕДЕНИЕ И ОБРАБОТКА РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА

Подготовка к проведению эксперимента

К проведению эксперимента нужно тщательно подготовиться: проверить и откалибровать приборы, подготовить материалы и журнал эксперимента (чтобы была ясна последовательность его проведения), определить вид эксперимента – ПФЭ или ДФЭ, провести *рандомизацию* опытов (определение случайного порядка проведения опытов). Рандомизация применяется для исключения систематической ошибки. Для этого берется таблица случайных, равномерно распределенных чисел, если опытов 16, то выбирают цифры от 1 до 16. Например, 8, 13, 3, 12, 11, 15, 6, 14, 2, 4, 5, 7, 1, 9, 10, 16.

Составляется матрица ПФЭ 2^3 и проводится эксперимент (табл. 23.1).

Таблица 23.1

ПФЭ 2^3

Номер опыта	Порядок проведения опытов	x_0	x_1	x_2	x_3	$y^{(1)}$	$y^{(2)}$	\bar{y}
1	8, 13	+	–	–	–	2,40	1,46	1,93
2	3, 12	+	+	–	+	3,58	2,78	3,18
3	11, 15	+	–	–	+	6,10	4,50	5,30
4	6, 14	+	–	+	–	4,76	3,12	3,94
5	2, 4	+	+	+	–	5,35	4,15	4,75
6	5, 7	+	+	–	–	4,10	2,70	3,40
7	1, 9	+	–	+	+	5,42	4,00	4,71
8	10, 16	+	+	+	+	15,41	12,99	14,20

Обработка результатов эксперимента

1. Нахождение средних значений в каждой строке (см. табл. 23.1).
2. Расчет дисперсии и дисперсии среднего в каждой строке матрицы

$$\sigma_j^{*2} = \frac{\sum_{i=1}^n (y_{ji} - \bar{y}_i)^2}{n-1},$$

где i – параллельные опыты, j – номер строки.

Например, $\sigma_1^{*2} = \frac{(2,40 - 1,93)^2 + (1,46 - 1,93)^2}{2 - 1} = 0,44$ и т. д.

Дисперсии среднего в каждой строке матрицы равны

$$\begin{aligned} \overline{\sigma_1^{*2}} = \frac{\sigma_1^{*2}}{n} = \frac{0,44}{2} = 0,22; \quad \overline{\sigma_2^{*2}} = 0,16; \quad \overline{\sigma_3^{*2}} = 0,64; \quad \overline{\sigma_4^{*2}} = 0,67; \quad \overline{\sigma_5^{*2}} = 0,36; \\ \overline{\sigma_6^{*2}} = 0,49; \quad \overline{\sigma_7^{*2}} = 0,504; \quad \overline{\sigma_8^{*2}} = 1,46. \end{aligned}$$

3. Проверка однородности дисперсий. Проводится с помощью критерия Кохрена

$$G_{\text{Э}} = \frac{\overline{\sigma_{\max}^{*2}}}{\sum_{j=1}^N \overline{\sigma_j^{*2}}}.$$

Экспериментальное значение $G_{\text{Э}}$ сравнивается с табличным значением $G_{\text{Т}}$ [14]. Если $G_{\text{Э}} < G_{\text{Т}}$, то дисперсии однородны.

$$G_{\text{Э}} = \frac{1,46}{0,16 + 0,64 + 0,67 + 0,36 + 0,49 + 0,504 + 1,46} = 0,34.$$

Табличное значение определяется при заданном уровне значимости p и числах степеней свободы $f_1 = n - 1, f = N(n - 1)$.

Если $p = 0,05$ и $f_1 = 2 - 1 = 1, f_2 = 8$, то $G_{\text{Т}} = 0,679$, то есть выполняется условие $G_{\text{Э}} < G_{\text{Т}}$, значит дисперсии однородны.

4. Проверка значимости различий средних. Сравнение средних проводится путем сравнения критерия Стьюдента [14], найденного экспериментальным путем $t_{\text{Э}}$, с табличным значением $t_{\text{Т}}$. Если экспериментальное значение меньше табличного, то различия незначимы.

Для двух близких значений \bar{y} (2-я и 6-я строки)

$$t_{\text{Э}} = \frac{\bar{y}_6 - \bar{y}_2}{\sigma^* \sqrt{\frac{1}{n_2} + \frac{1}{n_6}}},$$

где $n_2 = n_6 = 2$.

Средневзвешенная дисперсия

$$\sigma^{*2} = \frac{\sigma_1^{*2} f_1 + \sigma_2^{*2} f_2}{f_1 + f_2} = \frac{0,32 + 0,98}{2} = 0,65,$$

так как $f_1 = f_2 = n - 1 = 1$. Стандарт $\sigma^* \approx 0,8$.

Получим

$$t_{\Theta} = \frac{3,40 - 3,18}{0,8 \cdot 1} = \frac{0,22}{0,8} \approx 0,27.$$

Величину t_T находим для $p = 0,05$ и $f = n_1 + n_2 - 2 = 2$. Из таблицы [14] найдем $t_T = 4,3$.

Так как $t_{\Theta} < t_T$, то различие незначимо, оно могло появиться за счет ошибок эксперимента. В этом случае нужно уменьшить ошибки эксперимента или увеличить интервалы варьирования.

Определение коэффициентов регрессии (коэффициентов уравнения)

Будем искать коэффициенты полинома

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3.$$

Получим

$$b_0 = \frac{1,93 + 3,18 + 5,30 + 3,94 + 4,75 + 3,40 + 4,71 + 14,20}{8} = 5,18,$$

$$b_1 = \frac{-1,93 + 3,18 - 5,30 - 3,94 + 4,75 + 3,40 - 4,71 + 14,20}{8} = 1,21 \text{ и т.д.}$$

С учетом рассчитанных коэффициентов запишем

$$y = 5,8 + 1,21x_1 + 1,72x_2 + 1,67x_3 + 1,37x_1x_2 + 0,64x_1x_3 + 0,88x_2x_3.$$

Проверка значимости коэффициентов

Абсолютное значение коэффициента должно быть больше ошибки при его определении, т.е.

$$|b_j| > \Delta b_j, \Delta b_j = \pm t \sigma_{bj}^*.$$

Так как $f = N(n-1) = 8(2-1) = 8$, то для уровня значимости $p = 0,05$ найдем $t = 2,306$ и

$$\sigma_{bj}^{*2} = \frac{\sigma_{\bar{y}}^2}{N} = \frac{0,564}{8} = 0,0705.$$

Получим

$$\sigma_{bj}^* = 0,266$$

и

$$\Delta b_j = \pm 2,306 \cdot 0,266 = 0,613.$$

Сравнение Δb_j с коэффициентами регрессии показывает, что все коэффициенты значимы. Если какие-то из них оказываются незначимыми, то соответствующие члены необходимо исключить из модели.

Проверка адекватности модели

Определяется с использованием критерия Фишера

$$F_{\text{Ф}} = \frac{\sigma_{\text{ад}}^{*2}}{\sigma_{\bar{y}}^{*2}}.$$

Дисперсия адекватности имеет вид

$$\sigma_{\text{ад}}^{*2} = \frac{\sum_{i=1}^N \Delta y_i^2}{f},$$

где $\Delta y_i = \hat{y}_i - \bar{y}_i$; $f = N - (k + 1)$; \hat{y}_i – значение y , найденное по математической модели; \bar{y}_i – среднее значение; k – число факторов (в данном случае $k = 3$).

Дисперсия воспроизводимости имеет вид

$$\sigma_{\bar{y}}^{*2} = \frac{\sum_{i=1}^N \sigma_i^{*2}}{f} = \frac{0,22 + 0,16 + \dots + 1,46}{8(2-1)} = 0,564,$$

так как $f = N(n-1)$.

Для определения $\sigma_{\text{ад}}^{*2}$ (табл. 23.2) рассчитаем сначала все \hat{y}_i .

В качестве примера определим \hat{y}_1 . Получим $\hat{y}_1 = 5,18 = 1,21(-1) + 1,72(-1) + 1,67(-1) + 1,37(+1) + 0,64(+1) + 0,88(+1) = 3,46$.

Таблица 23.2

Пример оформления расчета

Номер опыта	\hat{y}	\bar{y}	Δy	Δy^2
1	3,46	1,93	1,53	2,34
2	4,80	3,18	1,62	2,61
3	...	5,30
4	...	3,94
5	...	4,75
6	...	3,40
7	...	4,71
8	...	14,20

Получим

$$\sum_{i=1}^N \Delta y_i^2 = \frac{2,34 + 2,61 + \dots}{8} = 24,46.$$

Дисперсия адекватности

$$\sigma_{\text{ад}}^{*2} = \frac{24,46}{4} = 6,12$$

и экспериментальное значение критерия Фишера

$$F_{\text{э}} = \frac{6,12}{0,564} = 11.$$

Табличное значение $F_{\text{Т}}$ находится для

$$f_1 = N - (k + 1) = 4, \quad f_2 = N(n - 1) = 8, \quad p = 0,05,$$

тогда $F_{\text{Т}} = 4,1$.

Так как условие $F_{\text{Т}} > F_{\text{э}}$ не выполняется, то гипотеза об адекватности несправедлива. Для получения адекватности можно уменьшать интервалы варьирования, изменять центр плана эксперимента или перейти к нелинейным моделям.

Лекция 24

КРУТОЕ ВОСХОЖДЕНИЕ ПО ПОВЕРХНОСТИ ОТКЛИКА

Метод Бокса – Уилсона

Это метод оптимизации на основе крутого восхождения по поверхности отклика (рис. 24.1).

В методе Бокса – Уилсона (крутого восхождения) используются свойства градиента функции и применяются методы ПФЭ и ДФЭ.

Градиент функции $y = f(x_1, x_2)$ имеет вид

$$\nabla y = \frac{\partial y}{\partial x_1} \vec{i} + \frac{\partial y}{\partial x_2} \vec{j}, \quad (24.1)$$

где \vec{i} и \vec{j} – единичные векторы (орты), ∇y указывает направление наиболее быстрого изменения функции. Линейный полином будет иметь вид

$$y = b_0 + b_1 x_1 + b_2 x_2, \quad (24.2)$$

где b_1, b_2 – частные производные от функции по параметрам.

На рис. 24.2 для примера приведён график функции $y = b_0 + b_1 x$.

Очевидно, что $b_1 = \operatorname{tg} \alpha = \frac{\partial y}{\partial x}$.

Очевидно также, что $b_2 = \frac{\partial y}{\partial x_2}$.

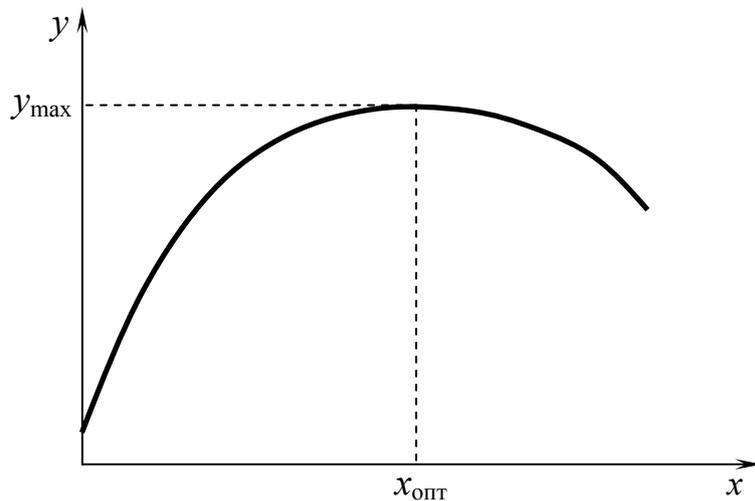


Рис. 24.1

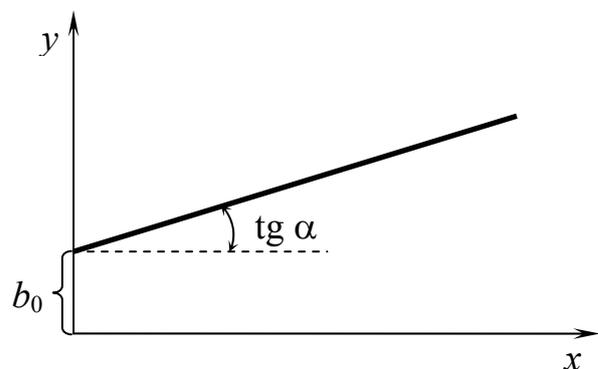


Рис. 24.2

Сравнивая выражения (24.1) и (24.2), видим, что уравнение регрессии (24.2) указывает направление градиента. Коэффициенты b_1 и b_2 можно определить с использованием ПФЭ или ДФЭ. Если изменять b_1x_1 и b_2x_2 пропорционально коэффициентам регрессии, то можно осуществить движение по градиенту.

Сначала выбираются начальные значения x_1 и x_2 , затем интервалы варьирования Δx_1 и Δx_2 , составляется матрица ПФЭ 2^2 (табл. 24.1).

Таблица 24.1

Матрица ПФЭ 2^2

Номер опыта	x_1	x_2	y
1	–	–	y_1
2	–	+	y_2
3	+	–	y_3
4	+	+	y_4

Определим коэффициенты

$$b_1 = \frac{-y_1 - y_2 + y_3 + y_4}{4};$$

$$b_2 = \frac{-y_1 + y_2 - y_3 + y_4}{4}.$$

Используя математическую модель (24.2) и изменяя слагаемые пропорционально b_1 и b_2 , двигаемся с выбранным шагом по градиенту в новую точку отсчета, где проводится мысленный (численный) эксперимент. Будем двигаться до тех пор, пока не начнется уменьшение y . В области точки максимума проводится новый ПФЭ. Если во всех четырех точках при проведении ПФЭ значение функции окажется меньше, чем в точке максимума, то экстремум обнаружен.

Для того чтобы более точно определить экстремум, надо провести ПФЭ при меньших интервалах варьирования или перейти к планам второго порядка.

Пример матрицы планирования и расчета крутого восхождения. Пусть $y = f(x_1, x_2)$. Начальные значения параметров, интервалы варьирования и уровни факторов приведены в табл. 24.2.

Таблица 24.2

Значения факторов и интервалов варьирования

Фактор	Уровень			Интервал варьирования (J)
	основной	верхний	нижний	
x_1	1,5	2,0	1,0	0,5
x_2	7,0	8,0	6,0	1,0

Результаты ПФЭ и движения по градиенту с использованием мысленных опытов (расчетов по математической модели) приведены в табл. 24.3.

Таблица 24.3

Определение данных для движения по градиенту

Номер опыта	Кодирование значений факторов		
	x_1	x_2	y
1	–	–	15,0
2	+1	–1	10,0
3	–1	+1	85,0
4	+1	+1	82,0
$b_i J_i$	$-2 \cdot 0,5 = -1$	$-4,5$	–
Шаг при изменении x_2 на 0,5	$-0,11$	$-0,5$	–
Округление	$-0,1$	$-0,5$	–

Коэффициенты регрессии

$$b_0 = \frac{\sum_{i=1}^n y_i}{4} = \frac{15+10+85+82}{4} = 48, \quad b_1 = \frac{-15+10-85+82}{4} = -2, \quad b_2 = -4,5.$$

Следовательно, линейная математическая модель

$$y = 48 - 2x_1 - 4,5x_2.$$

Для мысленного опыта 5 (табл. 24.4), например, получим

$$x_1 = \frac{1,4 - 1,5}{0,5} = -0,2, \quad x_2 = \frac{6,5 - 7}{1} = -0,5,$$

$$y_5 = 48,0 - 2(-0,2) - 4,5(-0,5) = 48,0 + 0,4 + 2,25 = 50,65 \text{ и т.д.}$$

Таблица 24.4

Результаты расчета движения по градиенту

Номер опыта	Кодирование значений факторов		
	x_1	x_2	y
5	1,4	6,5	50,65
6	1,3	6,0	...
7	1,2	5,5	95,15
8	1,1	5,0	...
9	1,0	4,5	...

$$y_7 = 48,0 - 2x_1 - 4,5x_2 = 95,15.$$

Расчеты (мысленные эксперименты) используются только для сокращения объема эксперимента. Параллельно с мысленными опытами проводится эксперимент, но не в каждой точке, и сравниваются $y_{\text{расч}}$ и $y_{\text{эсп}}$. Когда расхождение становится значительным, снова проводят ПФЭ для определения нового градиента.

Лекция 25

ПЛАНЫ ВТОРОГО ПОРЯДКА

Если зависимость (1) (рис. 25.1) нелинейная, а мы будем аппроксимировать её линейным уравнением (2), то это приведёт к большой ошибке ε . Следовательно, необходимо провести эксперимент ещё в одной точке, например в начальной x_0 .

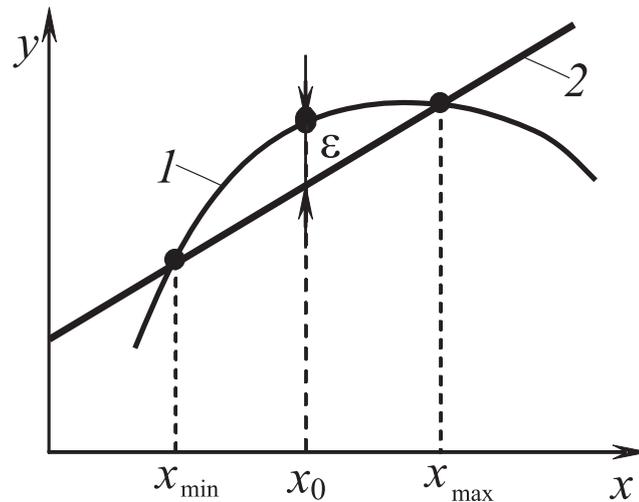


Рис. 25.1

То есть, если модель нелинейна, то нужно строить планы 3^k (табл. 25.1).

Таблица 25.1

План 3^2

Номер опыта	x_1	x_2	y
1	–	–	y_1
2	–	+	y_2
3	+	–	y_3
4	+	+	y_4
5	0	+	y_5
6	+	0	y_6
7	0	–	y_7
8	–	0	y_8
9	0	0	y_9

Эксперимент 3^k проводят редко из-за большого количества опытов, которое можно сократить, если воспользоваться композиционными планами Бокса – Уилсона.

Они делятся:

- 1) на композиционные ортогональные планы 2-го порядка;
- 2) композиционные рототабельные планы 2-го порядка.

Композиционные ортогональные планы 2-го порядка. Предположим, необходимо определить коэффициенты математической модели в виде полинома

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{11}x_1^2 + b_{22}x_2^2.$$

Для определения коэффициентов такого полинома к точкам, в которых проводится ПФЭ или ДФЭ, добавляют так называемые «звездные точки» на расстоянии α от центра плана (рис. 25.2).

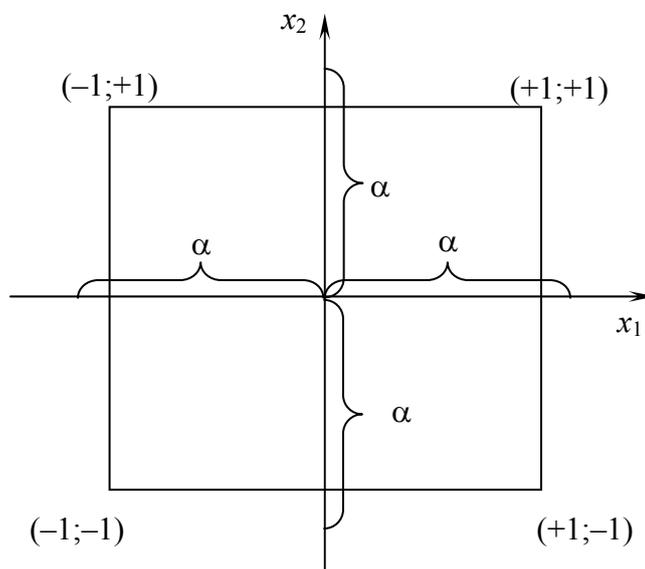


Рис. 25.2

Проводят эксперимент в этих точках и записывают МПЭ (табл. 25.2).

Таблица 25.2

Матрица планирования эксперимента

Номер опыта	x_1	x_2	$x_1 x_2$	x_1^2	x_2^2	y
1	+	+	+	+	+	y_1
2	+	-	-	+	+	y_2
3	-	+	-	+	+	y_3
4	-	-	+	+	+	y_4
5	0	$+\alpha$	0	0	α^2	y_5
6	$+\alpha$	0	0	α^2	0	y_6
7	0	$-\alpha$	0	0	α^2	y_7
8	$-\alpha$	0	0	α^2	0	y_8
9	0	0	0	0	0	y_9

Для нелинейных коэффициентов выбирают «звездные точки», которые удалены от центра эксперимента на α (α – плечо).

Этот план становится ортогональным, если подобрать соответствующим образом плечо α . Характеристики ортогональных планов приведены в табл. 25.3.

Таблица 25.3

Характеристики ортогональных планов

Характеристика плана	Число факторов		
	2	3	4
ПФЭ	2^2	2^3	2^4
Число «звездных точек»	4	6	8
Число точек в центре	1	1	1
Общее число опытов	9	15	25
Плечо α	1,0	1,215	1,414

Расчет коэффициентов регрессии проводится по формулам, приведенным в литературе [11].

Композиционные рототабельные планы 2-го порядка (табл. 25.4, 25.5). Критерием эффективности в ортогональных планах является ортогональность, а критерием рототабельности – одинаковая точность предсказания величины параметра оптимизации в точках, расположенных на одинаковом расстоянии от центра плана.

Условие рототабельности плана – однородность дисперсии в указанных точках, которая проверяется по критерию Фишера.

Формулы для расчета коэффициентов приведены в [11].

Таблица 25.4

Композиционный рототабельный план 2-го порядка

Номер опыта	x_1	x_2	x_1x_2	x_1^2	x_2^2	y
1	–	–	+	+	+	\bar{y}_1
2	+	–	–	+	+	\bar{y}_2
3	–	+	–	+	+	\bar{y}_3
4	+	+	+	+	+	\bar{y}_4
5	$+\alpha$	0	0	$+\alpha^2$	0	\bar{y}_5
6	$-\alpha$	0	0	$+\alpha^2$	0	\bar{y}_6
7	0	$+\alpha$	0	0	$+\alpha^2$	\bar{y}_7
8	0	$-\alpha$	0	0	$+\alpha^2$	\bar{y}_8
9	0	0	0	0	0	\bar{y}_9
10	0	0	0	0	0	\bar{y}_{10}
11	0	0	0	0	0	\bar{y}_{11}
12	0	0	0	0	0	\bar{y}_{12}
13	0	0	0	0	0	\bar{y}_{13}

Таблица 25.5

Характеристики композиционного рототабельного плана 2-го порядка

Характеристика плана	Число факторов		
	2	3	4
Ядро плана	2^2	2^3	2^4
Число «звездных точек»	4	6	8
Число точек в центре	5	6	7
Общее число опытов	13	20	31
Плечо α	1,414	1,682	2,00

Лекция 26

ВВЕДЕНИЕ В ТЕОРИЮ МАССОВОГО ОБСЛУЖИВАНИЯ

Рассмотрим характерный пример и убедимся, как часто встречаются случаи, при анализе которых целесообразно применять теорию массового обслуживания.

Например, сборочные работы в цехах радиотехнического производства без использования конвейера. Сборку осуществляет бригада, располагающая несколькими рабочими местами. На сборку поступают детали в случайные моменты времени. Полный комплект, необходимый для сборки, получается в случайные моменты времени, поэтому рабочие могут простаивать в ожидании деталей или в очереди стоят несколько полных комплектов деталей. Поскольку детали имеют отклонения и у рабочих разная квалификация, то для выполнения операции сборки затрачивается разное время и выходной поток продукции бригады также случайный.

Итак, поток поступающих деталей, состояние загрузки рабочих, время выполнения операции, выходной поток продукции случайны. При наличии случайности, которая является нормальным состоянием технологического процесса, закономерности этого процесса являются вероятностными.

Можно привести и другие примеры, которые могут рассматриваться как система массового обслуживания (СМО).

Система массового обслуживания и ее элементы. Введем для модели системы массового обслуживания терминологию, позволяющую абстрагироваться от конкретных процессов. В рассмотренном примере приводился технологический процесс сборки, при котором детали приходили на место сборки в случайные моменты времени и образовывали в совокупности случайный поток. Абстрагируясь от конкретных обстоятельств, будем считать, что работа (обслуживание) проводится над *входящим потоком заявок* (требований).

Этим термином будем называть совокупность заявок, поступающих на обслуживание в случайные (или неслучайные) моменты времени. В технологических процессах мы имеем дело с конкретным рабочим местом сборки и рабочим (бригадой). Рассматривая абстрактную модель, будем считать, что обслуживание заявки производится в одном «канале обслуживания», то есть СМО *одноканальная*. В многоканальных системах работа происходит в двух или более каналах. Подходя к каналам обслуживания в случайный момент времени, оче-

редная заявка может заставить все каналы уже занятыми. Тогда дальнейший путь заявки зависит от внутренней организации системы массового обслуживания. В *системах с отказами* (потерями) вновь поступившая на обслуживание заявка, застав все каналы уже занятыми, покидает систему. В простейшей системе с ожиданием заявки заявка находится в очереди и ожидает обслуживания неограниченное время в соответствии с правилом «первый пришел – первый обслужен». В системах с неограниченным временем ожидания заявок в *очереди с ожиданием* может быть различная *дисциплина очереди* (дисциплина отбора заявок), под которой понимают порядок поступления заявок из очереди в канал обслуживания, когда он освобождается.

Дисциплина очереди включает в себя: правила отбора требований на обслуживание, наличие или отсутствие приоритета, шкалу приоритета, правило применения приоритета. Существует *система с ограниченным ожиданием*, в которых участие заявки в очереди может ограничиться ее длиной или временем ожидания, которые могут быть детерминированными или случайными величинами. В технологии радиоаппаратостроения чаще всего встречаются очереди без приоритета с дисциплиной «первый пришел – первый обслужен».

Заявки, прошедшие обслуживание, образуют *выходящий поток*, который может являться входящим для последующего канала обслуживания. Выходящий поток, как правило, тоже случайный.

Таким образом, входящий поток заявок, канал обслуживания, очередь и выходящий поток образуют теорию массового обслуживания.

Но схема системы массового обслуживания может быть более сложной. Если каналы расположены последовательно, то процесс обслуживания в одном канале является фазой всего процесса обслуживания. Такая система называется многофазной. Еще более сложной являются система, имеющая последовательно и параллельно (или наоборот) расположенные каналы, и система с сетевой структурой. Основная задача изучения системы массового обслуживания – ее количественный анализ. Для этого устанавливаются зависимости между параметрами входящего потока, канала обслуживания, очереди и выходящего потока. Анализ позволяет определить в системе «слабые» звенья и их влияние на эффективность обслуживания, чтобы путем расширения «узких мест» добиться улучшения эффективности. Для решения такой задачи могут быть привлечены методы оптимизации.

Перспективной представляется и обратная задача – по заданным параметрам выходящего потока и канала обслуживания определить параметры входящего потока, который будет оптимально обслужен.

Вероятностное описание функционирования одноканальной системы массового обслуживания. Рассмотрев описание важнейших элементов СМО – входящего потока, времени обслуживания и дисциплины очереди, перейдем к вероятностному описанию закономерностей функционирования системы.

Наиболее удобно для описания состояния системы рассмотреть вероятность того, что в системе находится k заявок $P(k)$. Пользуясь ею, можно найти другие характеристики системы: вероятность появления очереди, среднее время ожидания в очереди, среднее время обслуживания, функцию распределения времени ожидания и др.

Условимся вести рассмотрение при следующих ограничениях: входящий поток пуассоновский, время обслуживания экспоненциальное, система состоит из одного канала с ожиданием.

Рассмотрим вероятность того, что в момент времени $t + \Delta t$ в системе будет находиться k заявок $P(k, t + \Delta t)$. Но такая ситуация, когда через промежуток времени Δt в системе имеется k заявок, возможна в четырех случаях.

1. В момент времени t в системе было k заявок. За время Δt ни одна заявка не пришла, ни одна не ушла (обслуживание за время Δt не было закончено). Если ни одна заявка не пришла, то это означает, что случайное время между заявками больше, чем интервал Δt . Вероятность такого случая для пуассоновского входящего потока имеет вид

$$P(T_3 > \Delta t) = e^{-\lambda \Delta t} \cong 1 - \lambda \Delta t.$$

Здесь $e^{-\lambda \Delta t}$ мы разложили в степенной ряд и ограничились рассмотрением членов, содержащих Δt в первой степени.

Чтобы ни одна заявка не покинула канал обслуживания, нужно, чтобы $t_{об} > \Delta t$. В случае экспоненциального распределения времени обслуживания вероятность этого события равна

$$P(t_{об} > \Delta t) = e^{-\mu \Delta t} \cong 1 - \mu \Delta t.$$

Вероятность такого события, когда в момент времени t в системе было k заявок, а за промежуток времени Δt ни одна заявка не пришла и не покинула систему, получается перемножением всех трех вероятностей по теореме умножения вероятностей

$$P(k, t)(1 - \lambda \Delta t)(1 - \mu \Delta t).$$

2. В момент времени $t + \Delta t$ в системе будет k заявок при следующих условиях. В момент времени t в системе была $(k - 1)$ заявка. За время Δt пришла одна заявка и не ушла ни одна. Вероятность такого события будет равна

$$P(k - 1, t)\lambda\Delta t(1 - \mu\Delta t).$$

Вероятность прихода хотя бы одной заявки получена как вероятность противоположного события, при котором не приходит ни одной заявки:

$$P(1, \Delta t) = 1 - P(0, \Delta t) = \lambda\Delta t.$$

3. В момент времени t в системе была $(k + 1)$ заявка, и ни одна не пришла, одна ушла.

Вероятность этого события равна

$$P[(k + 1), t](1 - \lambda\Delta t)\mu\Delta t.$$

4. В момент времени t было k заявок, одна пришла, и одна ушла. Вероятность такого события равна

$$P(k, t)\lambda\Delta t\mu\Delta t.$$

Другие ситуации при данных условиях считаем невозможными, так как вероятности других переходов системы за промежуток Δt являются величинами высшего порядка малости. Рассмотренные выше ситуации друг от друга независимы, поэтому вероятность того, что через промежуток времени Δt в системе будет k заявок, равна сумме вероятностей каждого отдельного события:

$$P(k, t + \Delta t) = P(k, t)[1 - (\lambda + \mu)\Delta t] + P(k - 1, t)\lambda\Delta t + P(k + 1, t)\mu\Delta t.$$

Здесь члены, содержащие Δt^2 , отброшены как величины высшего порядка малости.

Перенесем слагаемые $P(k, t)$ влево, разделим обе части уравнения на Δt и перейдем к пределу при $\Delta t \rightarrow 0$; получим

$$\begin{aligned} \frac{dP(k, t)}{dt} &= \lim_{\Delta t \rightarrow 0} \frac{\Delta P(k, t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{P(k, t + \Delta t) - P(k, t)}{\Delta t}; \\ \frac{dP(k, t)}{dt} &= \lambda P(k - 1, t) - (\lambda + \mu)P(k, t) + \mu P(k + 1, t). \end{aligned}$$

Аналогично мы можем найти вероятность того, что в момент времени $t + \Delta t$ в системе не будет ни одной заявки, т.е. $P(0, t + \Delta t)$.

Возможны две ситуации: первая – в момент времени t в системе не было заявок и за промежуток времени Δt в систему не пришла ни одна заявка, вторая – в момент времени t в системе была заявка, за время Δt она ушла.

Первая вероятность есть

$$P(0, t)(1 - \lambda\Delta t),$$

вторая –

$$P(1, t)\mu\Delta t$$

и

$$P(0, t + \Delta t) = P(0, t)(1 - \lambda\Delta t) + P(1, t)\mu\Delta t.$$

Получаем

$$\frac{dP(0, t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{P(0, t + \Delta t) - P(0, t)}{\Delta t} = -\lambda P(0, t) + \mu P(1, t).$$

Таким образом, мы пришли к системе линейных однородных уравнений, называемых уравнениями Эрланга:

$$\left. \begin{aligned} \frac{dP(0, t)}{dt} &= -\lambda P(0, t) + \mu P(1, t), \\ \frac{dP(k, t)}{dt} &= \lambda P(k-1, t) - (\lambda + \mu)P(k, t) + \mu P(k+1, t), \end{aligned} \right\}$$

с неизвестными функциями $P(0, t), P(1, t), \dots, k = 0, 1, 2, \dots$.

Параметрами в решении системы уравнений будут выступать плотность входящего потока заявок λ и величина μ , обратная среднему времени обслуживания. Произвольную постоянную, входящую в решение, можно определить из условия

$$\sum_{k=0}^{\infty} P(k, t) = 1.$$

Система уравнений Эрланга может быть решена, однако практическое использование этого решения связано с вычислительными трудностями. Если ограничиваться рассмотрением систем массового обслуживания только в установившемся режиме работы (опуская переходные режимы), то решение оказывается более простым. Доказано, что в рассматриваемых системах может существовать предельный режим при $t \rightarrow \infty$ (установившийся режим), при котором все вероятности состояния системы $P(0, t), P(1, t), \dots$ стремятся к постоянным пределам p_0, p_1, p_2, \dots , перестают зависеть от t и их производные, естественно, обращаются в нуль.

Существует

$$\lim_{t \rightarrow \infty} P(k, t) = p_k, \quad k = 0, 1, 2, \dots$$

Система дифференциальных уравнений Эрланга переходит при этом в систему алгебраических уравнений с неизвестными p_0, p_1, p_2, \dots , которые определить гораздо проще. Итак, для установившегося режима

$$\left. \begin{aligned} \mu p_1 - \lambda p_0 &= 0, \\ \mu p_{k+1} + \lambda p_{k-1} - (\lambda + \mu) p_k &= 0. \end{aligned} \right\}$$

Из второго уравнения системы нетрудно образовать формулу рекуррентности, которую можно использовать для решения системы уравнений:

$$p_{k+1} = \frac{1}{\mu} [(\lambda + \mu) p_k - \lambda p_{k-1}].$$

Все известные вероятности в стационарном режиме, как и в нестационарном, связаны условием

$$\sum_{k=0}^{\infty} p_k = 1.$$

Вернемся к решению системы. Из первого уравнения следует

$$p_1 = \frac{\lambda}{\mu} p_0.$$

Из формулы рекуррентности получаем

$$p_2 = \frac{1}{\mu} [(\lambda + \mu) p_1 - \lambda p_0] = \frac{\lambda^2}{\mu^2} p_0; \quad p_3 = \frac{\lambda^3}{\mu^3} p_0.$$

Нетрудно усмотреть закономерность. Для любого k

$$p_k = \left(\frac{\lambda}{\mu} \right)^k p_0.$$

Так как сумма всех вероятностей равна единице

$$\sum_{k=0}^{\infty} p_k = p_0 \sum_{k=0}^{\infty} \left(\frac{\lambda}{\mu} \right)^k = 1,$$

то получаем

$$p_0 = \frac{1}{\sum_{k=0}^{\infty} \left(\frac{\lambda}{\mu} \right)^k} = 1 - \frac{\lambda}{\mu}.$$

Предполагается, что $\frac{\lambda}{\mu} < 1$. Кроме того,

$$\sum_{k=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^k = \left(\frac{\lambda}{\mu}\right)^0 + \left(\frac{\lambda}{\mu}\right)^1 + \left(\frac{\lambda}{\mu}\right)^2 + \dots = \frac{1}{1 - \frac{\lambda}{\mu}}.$$

Окончательно получаем

$$p_k = \left(\frac{\lambda}{\mu}\right)^k \left(1 - \frac{\lambda}{\mu}\right).$$

Последнее выражение есть закон распределения количества в одноканальной системе с ожиданием. Оно называется формулой Эрланга. В формулах часто встречается отношение λ/μ , поэтому удобно его обозначить одной буквой $\alpha = \lambda/\mu$. Она выражает отношение среднего количества поступающих в систему заявок, приходящихся на одну обслуженную заявку (или к среднему времени обслуживания одной заявки), и называется приведенной плотностью потока заявок (также интенсивностью нагрузки системы). Пользуясь полученной формулой, рассмотрим основные закономерности функционирования одноканальной системы. Результаты расчета для разных α даны в таблице.

α	p_0	p_1	p_2	p_3	$p(k \geq 1)$	$p(k \geq 2)$	$m(v)$
0	1	0	0	0	0	0	0
0,1	0,9	0,09	0,009	0,0009	0,1	0,01	0,01
0,2	0,8	0,16	0,032	0,0064	0,2	0,04	0,05
0,3	0,7	0,21	0,063	0,0189	0,3	0,09	0,13
0,4	0,6	0,24	0,096	0,0384	0,4	0,16	0,27
0,6	0,4	0,24	0,144	0,0864	0,5	0,36	0,9
0,7	0,3	0,26	0,147	0,1029	0,7	0,49	1,63
0,8	0,2	0,16	0,128	0,1024	0,8	0,64	3,2
0,9	0,1	0,09	0,081	0,0729	0,9	0,81	8,1
1	0	0	0	0	1	1	∞

Очередь не образуется при поступлении заявки, когда в системе нет ни одной заявки (но тогда канал обслуживания простаивает) или находится одна заявка и вторая еще не поступила (один канал загружен). Вероятность образования очереди

$$P(k \geq 2) = 1 - (p_0 + p_1).$$

Можно построить кривую среднего количества заявок, находящихся в очереди v . Чтобы в очереди была одна заявка, необходимо в системе иметь две заявки (одна в очереди, одна в канале обслуживания). Вероятность такого состояния системы согласно формуле Эрланга

$$p_2 = \alpha^2 (1 - \alpha) = P(k = 2).$$

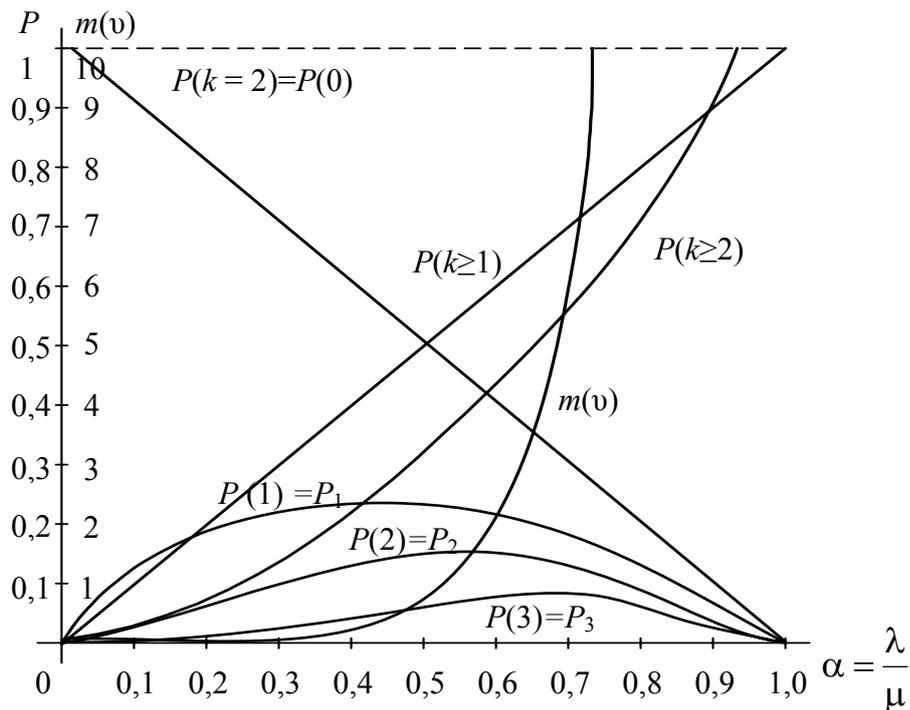
Вероятность того, чтобы в очереди находилось две заявки, равна

$$p_3 = \alpha^3(1 - \alpha) = P(k = 3).$$

По определению

$$\begin{aligned} m(v) &= \sum_{v=1}^{\infty} vP(k = 1 + v) = 1p_2 + 2p_3 + \dots = \alpha^2(1 - \alpha) + 2\alpha^3(1 - \alpha) + \dots = \\ &= \alpha^2(1 + \alpha + \alpha^2 + \dots) = \frac{1}{1 - \alpha}, \text{ так как } 1 + \alpha + \alpha^2 + \dots = \frac{1}{1 - \alpha}. \end{aligned}$$

Зависимость $m(v)$ показана на рисунке.



Если приведенная плотность потока заявок $\alpha = 0,5$, то вероятность загрузки системы $P(k \geq 1) = 0,5$ и вероятность простоя системы $p_0 = 0,5$, т.е. довольно большая. При этом вероятность образования очереди равна $0,25$, а среднее количество заявок в очереди — $0,5$ (см. рисунок). Изменяя λ/μ , можно добиться, чтобы очереди практически не было, но тогда система будет простаивать (см. кривую p_0 на рисунке). И наоборот, если довести α , например, до $0,8$, то p_0 уменьшится до $0,2$, т.е. вероятность простоя мала, но в очереди будет в среднем стоять $3,2$ заявки.

Из изложенного выше можно сделать очень важный вывод. Если нельзя исключить элемент случайности в потоке отказов (в поступлении деталей на сборку) и в процессе обслуживания (в процессе сборки), то неизбежны простои и очередь. Можно путем увеличения производственных мощностей практически исключить очереди, если обеспечить малое $\alpha = \lambda/\mu$. Но это не всегда экономически выгодно, так как связано с большими простоями. При $\alpha > 1$ длина очереди непрерывно увеличивается и становится бесконечно большой, следовательно, система так и не приходит к стационарному режиму. Об этом свидетельствует поведение средней длины очереди $m(v)$ в районе абсциссы $\alpha \cong 1$. Полученные выше зависимости дают представление о том, с какой вероятностью в системе может возникнуть то или иное состояние. Причем оказывается, что вероятность образования очереди при реальных α значительна. В этой связи большой интерес представляет определение того, сколько времени находится заявка в очереди, ожидая обслуживания. Для этого найдем закон распределения времени ожидания и среднее время ожидания заявки в системе.

Опуская выводы, можно записать закон распределения

$$P(T_{\text{ож}} > t_{\text{ож}}) = \frac{\lambda}{\mu} e^{-(\mu-\lambda)t_{\text{ож}}}.$$

По известной формуле теории вероятностей определяем среднее время ожидания в очереди

$$m(t_{\text{ож}}) = \frac{\lambda}{\mu} \frac{1}{\mu - \lambda}.$$

Среднее время нахождения в системе тех заявок, которые сразу не попадают на обслуживание, равно

$$m(t_{\text{ож}} + t_{\text{об}}) = m(t_{\text{ож}}) + \frac{1}{\mu} = \frac{1}{\mu - \lambda},$$

где $1/\mu$ – среднее время обслуживания заявки в канале.

Таким образом, имея в своем распоряжении два параметра λ и μ и пользуясь приведенными выше аналитическими выражениями, можно получить все вероятностные закономерности, описывающие систему в целом. Это очень много дает для практики. Действительно, для определения λ и μ с достаточной точностью нужно провести по 50 – 100 наблюдений; а для определения закона распределения методами математической статистики без использования аналитических результатов теории массового обслуживания необходимо выполнить 1000 – 10 000 наблюдений.

Библиографический список

1. Кофанов, Ю. Н. Теоретические основы конструирования, технологии и надежности радиоэлектронных средств : учеб. для вузов / Ю. Н. Кофанов. – М. : Радио и связь, 1991. – 380 с. – ISBN 5-256-00862-5.
2. Вентцель, Е. С. Теория вероятности / Е. С. Вентцель. – М. : Высш. шк., 1999. – 576 с. – ISBN 5-06-003650-2.
3. Он же. Исследование операций : задачи, принципы, методология / Е. С. Вентцель. – М. : Наука, 1988. – 208 с.
4. Адлер, Ю. П. Планирование эксперимента при поиске оптимальных условий / Ю. П. Адлер [и др.]. – М. : Наука, 1976. – 278 с.
5. Голинкевич, Т. А. Прикладная теория надежности : учеб. для вузов / Т. А. Голинкевич. – М. : Высш. шк., 1985. – 168 с.
6. Фомин, А. В. Допуски в РЭА / А. В. Фомин [и др.]. – М. : Сов. радио, 1973. – 128 с.
7. Маслов, А. Я. Оптимизация радиоэлектронной аппаратуры / А. Я. Маслов [и др.] ; под ред. А. Я. Маслова, А. А. Чернышова. – М. : Радио и связь, 1982. – 200 с.
8. Норенков, И. П. Системы автоматизированного проектирования. В 9 кн. Кн. 1. Принципы проектирования и структура / И. П. Норенков. – М. : Высш. шк., 1986. – 326 с.
9. Он же. Основы автоматизированного проектирования : учеб. для вузов / И. П. Норенков. – 2-е изд., перераб. и доп. – М. : Изд-во МГТУ им. Н.Э. Баумана, 2002. – 336 с. – ISBN 5-7038-2090-1.
10. Антонов, А. В. Системный анализ : учеб. для вузов / А. В. Антонов. – М. : Высш. шк., 2005. – 454 с. – ISBN 5-06-04862-4.
11. Ахназарова, С. Л. Оптимизация эксперимента в химии и химической технологии / С. Л. Ахназарова, В. В. Кафаров. – М. : Высш. шк., 1978. – 319 с.
12. Талицкий, Е. Н. Анализ точности электронных средств : метод. указания к лаб. работам / Е. Н. Талицкий ; Владим. гос. ун-т. – Владимир, 2002. – 72 с.
13. Он же. Вероятностно-статистические методы при проектировании электронных средств : метод. указания к лаб. работам / Е. Н. Талицкий ; Владим. гос. ун-т. – Владимир, 2003. – 48 с.
14. Бородюк, В. П. Статистические методы в инженерных исследованиях (лабораторный практикум) : учеб. пособие / В. П. Бородюк [и др.] ; под. ред. Г. К. Круга. – М. : Высш. шк., 1983. – 216 с.

Учебное издание

ТАЛИЦКИЙ Евгений Николаевич

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ПРОЕКТИРОВАНИЯ
ЭЛЕКТРОННЫХ СРЕДСТВ

Курс лекций

Подписано в печать 19.11.07.

Формат 60x84/16. Усл. печ. л. 9,30. Тираж 100 экз.

Заказ

Издательство

Владимирского государственного университета.
600000, Владимир, ул. Горького, 87.