ФЕДЕРАЛЬНОЕ АГЕНТСТВО ВОЗДУШНОГО ТРАНСПОРТА
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ
ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
ГРАЖДАНСКОЙ АВИАЦИИ»

Кафедра технической эксплуатации летательных аппаратов и авиадвигателей
П.К. Кабков

ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
И СИСТЕМНЫЙ АНАЛИЗ

Рекомендовано Учебно-методическим объединением вузов Российской Федерации по образованию в области эксплуатации авиационной и космической техники для межвузовского использования в качестве учебного пособия

Москва-2005

УДК 519(075.8)
ББК 22.17я73
К 12

Печатается по решению редакционно-издательского совета Московского государственного технического университета ГА
Рецензенты: канд. техн. наук, доц. Е.Д. Герасимова;
гл. технолог проектно-изыскательного и научно-исследовательского института «Аэропроект» В.А. Шиманский

Кабков П.К.
К 12 Исследование операций и системный анализ: Учебное пособие. – М.: МГТУ ГА, 2005. – 96 с., 12 табл., 41 рис.
ISBN 5-86311-461-4

В данном учебном пособии изложены основы системного анализа и методы исследования операций как прикладные методы системного анализа

Данное учебное пособие издается в соответствии с учебной программой для студентов специальности 130300 дневного обучения.

Рассмотрено и одобрено на заседаниях кафедры 28.12.04 г. и методического совета 14.12.04 г.

К $\frac{2705140400 - 017}{\text{Ц33(03)-05}}$ ББК 22.17я73
Св. план 2005 г.
поз.17

КАБКОВ Павел Кондратьевич

ИССЛЕДОВАНИЕ ОПЕРАЦИЙ И СИСТЕМНЫЙ АНАЛИЗ
Учебное пособие

Редактор Е.А. Колотушкина

Подписано в печать 11.05.2005 г.

Печать офсетная	Формат 60x84/16	5,14 уч.-изд. л.
5,58 усл.печ.л.	Заказ №1360/	Тираж 400 экз.

Московский государственный технический университет ГА
125993 Москва, Кронштадтский бульвар, д.20
Редакционно-издательский отдел
125493 Москва, ул. Пулковская, д.6а
ISBN 5-86311-461-4

ВВЕДЕНИЕ

Техническая эксплуатация самолётов как область научной, инженерно-технической и производственно-хозяйственной деятельности, является специфической системой, состоящей из совокупности объектов и средств технической эксплуатации, летного и инженерно-технического состава и системы управления его деятельностью.

Техническая эксплуатация призвана обеспечивать работоспособность и исправность авиационной техники, своевременную готовность её к использованию по назначению при наименьших трудовых и материальных затратах.

Научной основой рассмотрения технической эксплуатации как специфической системы является системный анализ, представляющий собой совокупность методов, ориентированных на исследование сложных систем.

Содержательная особенность этих методов зависит от характера рассматриваемых в системном анализе объектов.

Из всего многообразия методов, применяемых в системном анализе, в настоящем пособии рассматриваются методы, имеющие прикладное значение при выполнении задач, решаемых воздушным транспортом и его эксплуатационными предприятиями и связанных главным образом с технической эксплуатацией летательных аппаратов и авиационных двигателей.

Основные направления системного анализа, обеспечивающие специальные дисциплины, связаны с методами определения эффективности и показателей качества систем, вероятностно-математическими методами анализа и методами исследования операций. В методах исследования операций основное внимание уделено Марковским и полумарковским случайным процессам и системам массового обслуживания.

Как известно, исследование операций – наука, занимающаяся количественным обоснованием решений во всех областях целенаправленной человеческой деятельности [19]. При этом под операцией понимают любое мероприятие (или систему действий), объединенное единым замыслом и направленное к достижению определённой цели. В системном анализе, как это будет рассмотрено в настоящем пособии, аспект цели является одной из важнейших характеристик системы. Вследствие этого и методологические принципы, и математические методы системного анализа и исследования операций имеют много общего. Можно даже сказать, что прикладные математические методы исследования операций являются одним из важнейших инструментов системного анализа.

По дисциплине «Исследование операций и системный анализ» раньше было издано четыре части конспекта лекций [31,32,33,34]. Настоящее учебное пособие написано с учетом опыта автора по чтению лекций и проведению лабораторных работ и практических занятий по этой дисциплине, оно в ряде мест дополнено и методически усовершенствовано.

1. ОСНОВЫ СИСТЕМНОГО АНАЛИЗА

1.1.Методология системного анализа

Методология – наука о методах исследований, изучения принципов, лежащих в основе рассмотрения того или иного объекта, явления. Другими словами, методология – совокупность принципов, категорий, понятий, лежащих в основе данной области знаний.

Несмотря на то, что методология содержит общие принципиальные положения, она играет конструктивную роль, очерчивая область ведения данной науки и определенным образом организуя исследования в этой науке.

Можно выделить четыре основных уровня современного методологического знания, уровня методологии.

Первый уровень – уровень философской методологии, представляющий собой анализ общих принципов познания и категории науки в целом. Эта сфера методологии является разделом философского знания и разрабатывается специфическими для философии методами.

Второй уровень – уровень общенаучных методологических принципов и форм исследования. Общенаучный характер методологических концепций второго уровня означает их междисциплинарную природу. Применяется он на стыках традиционных наук, к новым синтетическим областям науки. Методологически концепции этого уровня не претендуют на решение мировоззренческих общефилософских задач, они разрабатываются в сфере нефилософского, главным образом в рамках современной логики науки. Принципиально их возможно переносить из одной области научного или технического знания в другие.

Третий уровень – уровень конкретно-научной методологии. На этом уровне организуются методы, принципы и процедуры исследования, применяемые в сложившихся научных дисциплинах (физика, химия, биология, математика и т.п.), и разрабатывается понятийный аппарат определенной научной дисциплины.

Четвертый уровень – уровень методики и техники исследования. На этом уровне производится разработка способов получения информации (данных) в процессе исследования (объекты, явления и т.п.) и методов обработки экспериментальных данных. Разработанные на этом уровне методики используются для определенной научной дисциплины или даже ее раздела.

Системный подход – исследование объектов как систем. Заметим, что термины *системный подход, системный анализ, теория систем, системология* в литературе [17, 51, 65] часто употребляются как синонимы. Вне зависимости от применяемого термина отметим, что системный подход является методологической основой системных исследований и относится к методологии второго уровня.

Отнесение его к методологии второго уровня обусловлено прежде всего тем, что системные исследования по своей сути носят междисциплинарный характер. Это же определило привлечение к системному анализу таких научных дисциплин, как математика, теория вероятностей, теория массового обслуживания, теория графов, которые в совокупности сформировали такое научное направление, как исследование операций. Системный анализ привлекает также результаты таких наук, как кибернетика, теория информации, логика, теория подобия и моделирования и пр.

Конкретизация методологии любой науки состоит в раскрытии содержания принципов этой науки, а также ее основных понятий и определений. Это положение относится также и к системологии, системному подходу.

Выделим два основных взаимодополняющих принципа системного подхода.

Первым принципом системного подхода является принцип всестороннего подхода к изучаемому объекту, рассмотрение его во всех взаимосвязях, взаимозависимостях и действиях. Когда говорят о необходимости системного подхода к решению какой-либо проблемы, то, в первую очередь, имеют в виду именно этот принцип.

Заметим, однако, что здесь необходимо различать внешние взаимосвязи объекта и внутренние взаимозависимости составляющих его частей.

С точки зрения внешних взаимосвязей систему можно представить в виде некоторого объекта со входами и выходами (рис. 1.1.). Система S является объектом с четко выраженными границами. Она имеет внешние воздействия (входы) - $\omega_i \in \Omega$, т.е. элементы, воспринимающие внешние входные воздействия, и выходы - $y_i \in Y$, т.е. некоторые величины, характеризующие результаты внешних воздействий. Сама система имеет некоторые параметры $x_i \in X$, характеризующие ее состояние.

Изображенная схема часто называется моделью «черного ящика», при которой интересуются только влиянием внешних воздействий на реакцию системы, не вдаваясь в процессы, происходящие в самом объекте.

С точки зрения внутренних взаимосвязей их учет может быть сделан с различной степенью подробности. На рис. 1.2а внутренний состав приведен на уровне крупных подсистем и их взаимосвязей. На рис. 1.2б – на уровне элементов. Кроме состава самой системы, необходимо также учитывать и влияние на результат ее функционирования параметров составляющих систему элементов.

Вторым основным принципом системного подхода является принцип эмерджентности. Его сущность состоит в том, что свойства системы как целого не содержатся в частях этой системы, не сводятся к простой совокупности свойств частей, составляющих систему. Система как целое приобретает совершенно новое свойство, характерное ей как новому целостному образованию. Другими словами, целое обладает такими качествами, которых нет у его частей.

В литературе [50] приводится такой пример проявления этого свойства. Пусть имеется цифровой автомат S, преобразующий любое целое число на его входе в число на единицу больше входящего (рис. 1.3а). Если соединить два таких автомата в кольца (рис. 1.3б), то в полученной схеме образуется новое свойство: она генерирует на выходе А последовательности только из четных чисел, а на выходе B – только из нечетных чисел. Если соединить два автомата S параллельно (рис. 1.3в), то с точки зрения генерирования чисел изменений в новой системе не будет, но она приобретает новое свойство с точки зрения повышения надежности.

Другим примером может явиться самолет как система, состоящая из двух основных частей – планера и авиационного двигателя. Планер (да и то специально сконструированный, легкий и с крылом соответствующего аэродинамического качества) может летать только благодаря восходящим потокам воздуха и не может летать по произвольно назначенному маршруту.

Авиационный двигатель, создающий тягу, самостоятельно не летает. Соединение же этих элементов приводит к новому свойству – способности аппарата тяжелее воздуха летать на значительные расстояния по заранее заданному маршруту.

Такое «внезапное» появление новых качеств у систем стали обозначать термином эмерджентность. Английское слово emergence означает возникновение из ничего, внезапное появление, неожиданную случайность. В специальной литературе на русском языке пока не найден эквивалентный русский термин. Иногда это свойство не совсем точно называют целостностью системы, подразумевая под этим не только прямое значение этого слова, но и возникновение новых свойств.

Свойство эмерджентности признается и официально. При государственной экспертизе изобретением признается и новое, ранее не известное соединение хорошо известных элементов, если при этом возникают новые полезные свойства.

1.2.Основные понятия и определения

Говоря о сущности системного подхода, мы уже оперировали некоторыми понятиями и определениями (система, подсистема, модель системы), не раскрывая их содержания. Продолжая конкретизацию методологии изучаемой науки, раскроем содержание основных понятий: собственно система, структура системы и модель системы.

1.2.1. Понятие «система» и классификация систем

На интуитивном и бытовом уровнях у каждого человека понятие «система» складывается в результате ассоциаций, связанных с употреблением термина *система* в сочетании со словами солнечная, нервная, отопительная или уравнений, показателей и т.п.

С научной точки зрения понятие «система» является одним из центральных в системном подходе, в системологии. Очень многие авторы поэтому анализировали это понятие, развивали определение системы до различной степени формализации. Например, в [64] собрано 35 различных определений системы. Такая множественность определений объясняется как тем, что сама наука системология находится в стадии становления, так и тем, что в зависимости от объекта и целей исследования в определениях подчеркивается та или иная специфичность объекта исследования.

Приведем некоторые наиболее общие определения систем. Людвиг фон Берталанфи – основатель современной общей теории систем: «Система может быть определена как совокупность элементов, находящихся в определенных отношениях друг с другом и со средой» [9]. Флейшман Б.С. – современный специалист по теории экосистем: «Под системой понимается множество элементов со связями между ними» [65].

Приведенные определения могут быть отнесены к любым типам систем (солнечная система, система уравнений, транспортная система, система управления воздушным движением и т.п.).

Вопросы системологии, относящиеся к системам различного типа, изложены в работах [17, 50, 51, 60, 66]. При рассмотрении конкретных типов систем даются обычно определения, отражающие особенности этого типа систем.

Рассмотрим классификацию систем и выделим те типы систем, которые нас будут интересовать. Всю возможную совокупность систем можно разделить на абстрактные и материальные (рис. 1.4).

Классификация систем

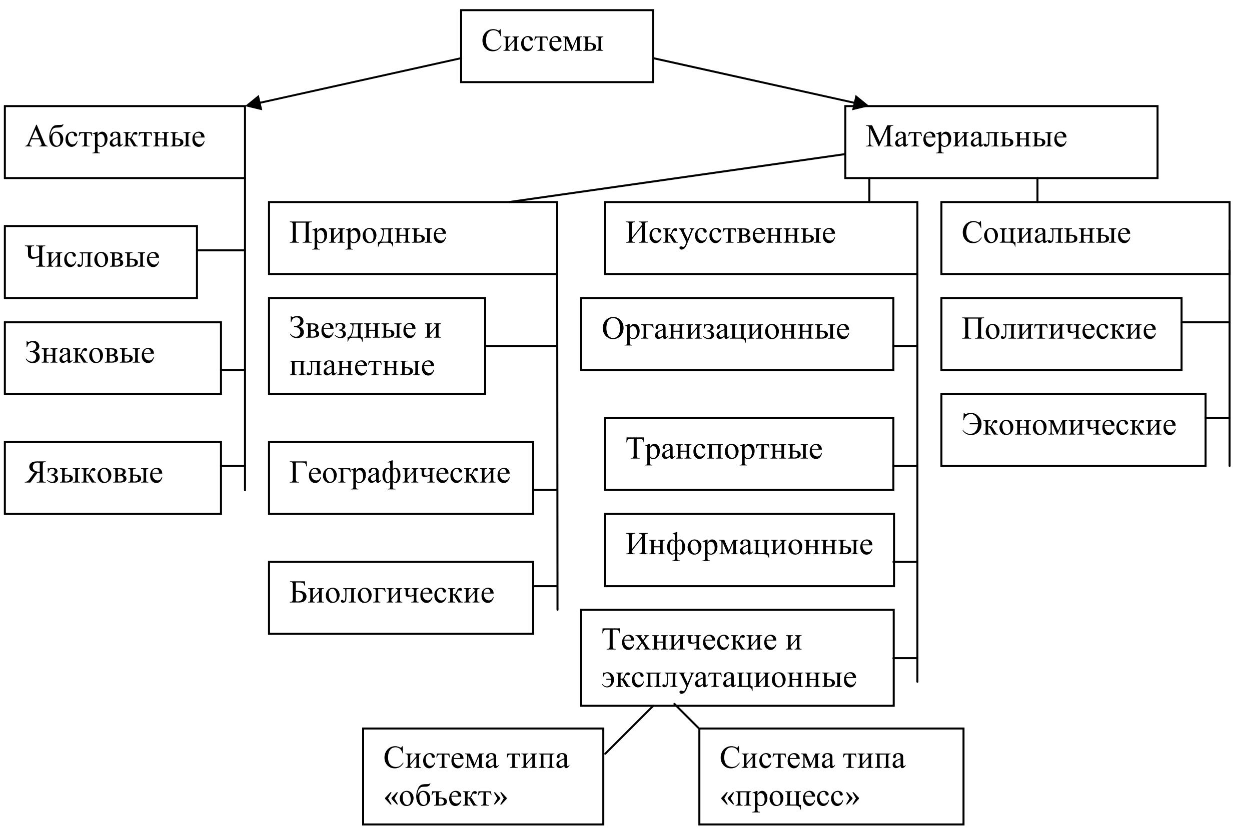


Рис. 1.4

Абстрактные системы являются продуктами человеческого мышления. Это математические, знаковые и логические системы, системы понятий, гипотез, языковые системы и т.д.

В качестве материальных систем можно выделить природные, искусственные и социальные.

Природные системы – это неорганические и органические совокупности объектов: звёзды и планетные системы, географические и геологические образования, биологические (растительные и животные) совокупности.

Искусственные (созданные человеком) - это объекты материальной культуры человечества. Здесь многообразие типов систем возрастает вместе с прогрессом науки, техники и производства. Некоторые из типов этих систем: организационные (управления), транспортные, информационные, технические системы, эксплуатационные, технологические и т.п.

Социальные системы занимают своеобразное положение. С одной стороны, любая общественная система вырастает на материальном базисе, а с другой стороны, её идеологической основой могут быть абстрактные и утопические идеи и положения. Хотя это и интересная сама по себе область знания, мы ее касаться не будем.

Вопросы, связанные с общим направлением исследований по системному анализу, освещены в работах [50, 51, 52, 64].

Наиболее развитой областью исследований и разработок является теория технических систем. Это направление исследований объединяется общим названием системотехника. Общие вопросы системотехники изложены в работах [15, 60, 66].

Предметом нашего рассмотрения будут являться искусственные системы, в первую очередь технические и эксплуатационные системы, из которых можно выделить системы типов «объект» и «процесс». Вопросы теории эксплуатационных систем изложены в работах [23, 57, 58, 67].

Кроме приведенной выше общей классификации систем, они могут быть классифицированы по другим признакам.

По положению системы в иерархии: системы различных уровней подчиненности (линейной и «штатной» подчиненности, сферы полномочий и пр.).

По связям с внешней средой: открытые (по крайней мере с одним входом и выходом), замкнутые (без связей с окружением).

По изменениям состояния: динамическое (состояния изменяются со временем), статическое (состояния во времени не меняются).

По характеру параметров системы и её поведения: детерминированные (параметры строго определены и поведение системы однозначно зависит от ее параметров и предсказуема), стохастические (параметры системы изменяются случайным образом, поведение системы неоднозначно и может быть определено только с определенной степенью вероятности).

По характеристикам состава: подсистема, система, надсистема.

По степени сложности системы: предельно сложные (народное хозяйство страны в целом), очень сложные (полностью автоматизированные предприятия; производственный комплекс), сложные (самолет, автомобиль), простые (болтовое соединение).

По размерам системы: большая, небольшая.

В приведенных примерах классификации появились новые понятия и термины. Продолжая рассмотрение основных понятий и определений, в первую очередь раскроем понятия большая и сложная система. Строгих определений этих понятий нет, поэтому дадим разъяснения сути этих терминов.

Под сложной системой понимают такую систему, в которой изменение значения какого-либо параметра влечет за собой изменение многих других параметров, причем зависимости между этими изменениями редко бывают линейными. Математическое описание функционирования такой системы является достаточно сложным.

Под большой системой понимают систему, которая является большой как с точки зрения разнообразия составляющих её элементов, так и с точки зрения количества одинаковых частей, количества выполняемых функций и стоимости.

Заметим, что часто понятия большая система и сложная система употребляются как синонимы.

И понятие системы вообще, большой и сложной системы в особенности являются относительными. При детальном рассмотрении состава любой системы некоторые части системы могут быть, в свою очередь, разбиты на составные части и т.д. Те части системы, которые при данном конкретном рассмотрении мы считаем неделимыми, называют элементами системы.

Части системы, состоящие более чем из одного элемента, называют подсистемами.

Например, в каком-нибудь механическом агрегате, рассматриваемом как система, болтовое соединение можно считать подсистемой, состоящей из болта и гайки. Но если тот же болт рассматривать с точки зрения структуры материала и его химического состава, то он, в свою очередь, будет системой, состоящей из различных веществ и элементов.

Поэтому при рассмотрении материальных объектов, создаваемых человеком, целесообразно границы системы определить не по качествам составляющих её элементов, а целями или задачами, которые должна решать система.

Поэтому аспект цели является одним из важнейших при характеристике системы. Этот же принцип применяется при подразделении системы на подсистемы, т.е. подсистемой следует называть такую часть системы, которая выполняет определенную задачу в интересах общей цели.

Заметим, что большие системы, как правило, создаются не для выполнения одной цели, а нескольких, т.е. являются многоцелевыми (например, многоцелевые военные самолеты).

1.2.2 Структура системы

Само слово структура (лат. Structure – строение, связь) означает относительно устойчивую связь (отношение) и взаимодействие элементов, сторон, частей предмета, явления, процесса как целого [38]. В общей теории систем [52] под структурой понимается установленное отношение между элементами множества или операциями над ними.

Имеет место следующая детализация любой системы: система как некоторое целостное образование («черный ящик»), состав системы (элементы и подсистемы) и, наконец, структура системы, описывающая определенные связи (отношения) между элементами и подсистемами системы.

Следует однако иметь в виду, что должны рассматриваться не любые возможные связи, а только те, которые необходимы и достаточны для достижения цели системы. Поэтому для технических систем, в которых аспект цели является определяющим, можно дать следующие определения.

Структурой системы называют совокупность необходимых и достаточных для достижения цели отношений между элементами системы.

В ряде случаев вопрос о структуре становится главным, проблемным. Более того, вообще система не может быть без структуры, определяющей отношения между её элементами.

Характерно также, что при конкретизации видов связей между элементами системы возникает необходимость выявления различных типов структур в одной и той же системе.

Рассмотрим, какие могут быть типы структур технических систем (рис. 1.5).

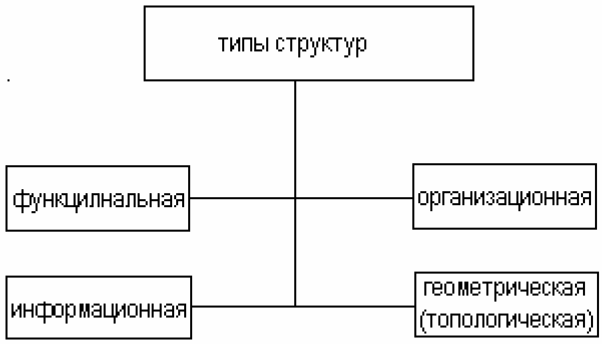


Рис 1.5

Функциональная структура – структура, отражающая функциональные связи и взаимодействия между элементами.

Организационная структура – структура, определяющая административное деление и подчиненность в системе.

Информационная структура – структура, раскрывающая пути потоков информации, взаимосвязь элементов системы, собирающих, обрабатывающих, передающих и использующих информацию.

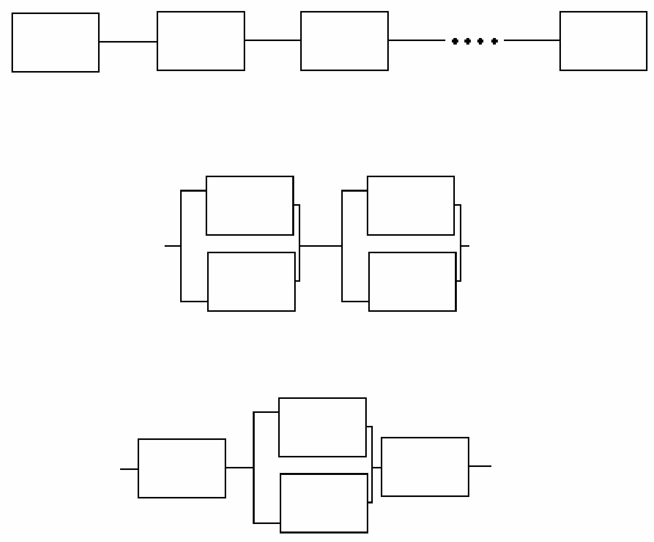
Геометрическая (топологическая) структура – структура, отражающая расположение элементов системы в пространстве, необходимое для выполнения её функций (например, радиотехнические системы местоопределения самолета).

При подробном рассмотрении какой-либо конкретной системы могут быть выделены и другие типы структур.

Для наглядности структуру системы часто изображают графически в виде элементов, взаимосвязи между которыми изображают соединительными линиями. При необходимости указывается стрелками направление взаимодействия (ориентированный граф).

Графическая конструкция структуры в ряде случаев используется как название той или иной структуры. На рис 1.6 приведены линейные структуры последовательно и параллельно соединенных элементов и линейная структура смешанного соединения элементов. Приведенные виды соединений элементов используются для расчета надежности функциональных систем методом структурных схем [30].

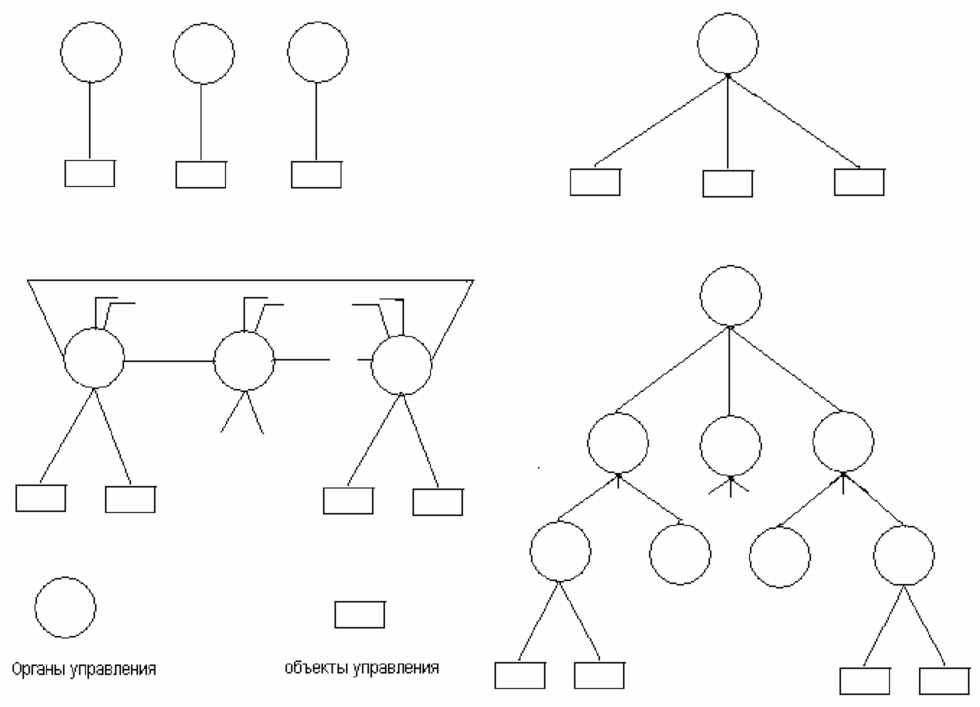
Для организационных структур в зависимости от характера управления могут быть виды структур, представленные на рис. 1.7.



а) последовательного соединения элементов
б) параллельного соединения элементов
в) смешенного соединения элементов

Рис. 1.6

Виды структур управления



а) децентрализованная
б) централизованная
в) централизованная рассредоточенная
г) иерархическая

Рис. 1.7

1.2.3 Модель системы

Моделирование есть замена изучения интересующего нас объекта (явления, устройства, системы, процесса) в натуре изучением его на модели. Основной смысл моделирования заключается в том, чтобы по результатам исследований с моделями можно было бы давать необходимые ответы о характере эффектов и о различных параметрах, связанных с объектом, в натурных условиях.

На практике применяются следующие виды моделирования: физическое, моделирование методами аналогии и математическое моделирование.

Физическое моделирование – это такое моделирование, при котором сохраняется физическая природа исследуемых явлений. При этом используются критерии подобия – безразмерные величины, связывающие физические параметры исследуемых явлений. Примером физического моделирования является изучение аэродинамических характеристик самолёта в аэродинамической трубе.

В случае моделирования аналогиями модель производит иное физическое явление, отличное по своей природе от исследуемого, но описываемое теми же уравнениями. Метод моделирования аналогиями основан на формальной аналогии дифференциальных уравнений, описывающих различные по физической сущности процессы. Примером аналогии является электродинамическая и электротепловая аналогии.

Математическое моделирование основано на построении математических зависимостей или алгоритмов, отражающих отношения элементов системы или процессов в ней.

Математические модели могут быть детерминированными и стохастическими. Первые математически однозначно описывают происходящие в объекте процессы. Стохастические модели учитывают вероятностные характеристики процессов, происходящих в системе.

Методы исследования операций – это по сути дела различного рода математические модели.

1.3. Система типа «процесс»

Рассматривая ранее классификацию систем, из группы технических систем мы выделили системы типа «объект», элементами которых являются предметы (двигатель, машина, строение, агрегат и т. п.), и системы типа «процесс», элементами которых являются операции (изготовление, транспортировка, обслуживание и т п)

Процесс (лат processus -ход, прохождение, продвижение) - закономерная последовательность следующих друг за другом моментов развития чего-либо.

Примеры: производственный процесс как последовательная смена трудовых операций, процесс роста, процесс мышления, болезнетворный воспалительный процесс, процесс обучения и т. п.

Системы типа «процесс», как и системы типа «объект» можно классифицировать по их особенностям, например:

абстрактные процессы (процесс мышления);

природные процессы (геологические процессы, процессы в растительном и животном мире);

технологические процессы (процессы производства, процессы технического обслуживания и ремонта ЛА);

общественно-политические процессы.

Мы будем рассматривать искусственные технические и технологические процессы, которые организует человек с целью осуществления желательных для него изменений. В такого типа процессах, как и для систем типа «объект», существенным является аспект цели.

Объект действия в процессе называют операндом. Процесс, при котором сам операнд (Od) или его свойства претерпевают изменения при участии людей и технических средств с целью достижения желаемого состояния операнда, называют преобразованиями. Преобразование есть следствие определенных воздействий, основанных на механических , химических, электрических и других воздействиях и описываемых некоторой инструкцией – рецептом , алгоритмом, технологией . Преобразование вызывается либо неудовлетворительным состоянием Od^1 либо потребностью в Od^2.

Целенаправленное воздействие на операнд выполняется операторами. Это воздействие осуществляется в виде потоков материи –S, энергии-En и информации-I. Осуществляют же эти воздействия люди - Me, технические средства систем-Ts и окружение –Umg.
Схема воздействия преобразования приведена на рис. 1.9.

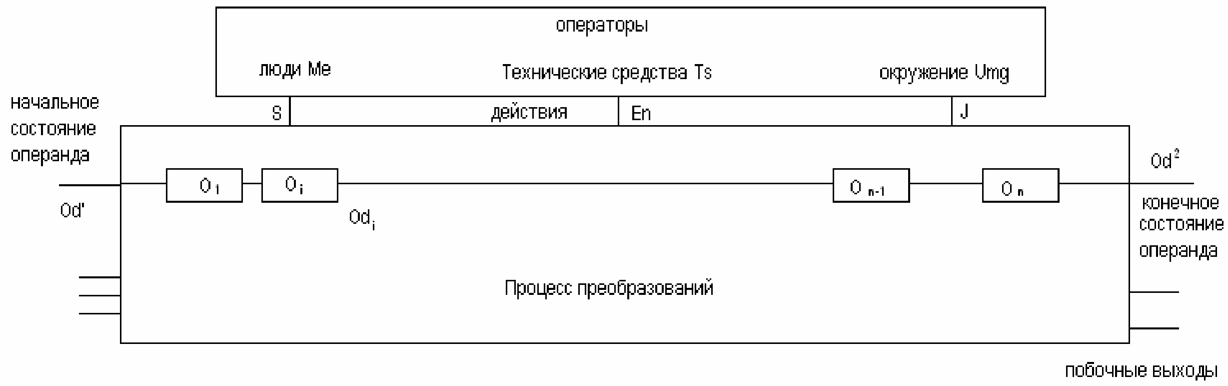


Рис. 1.9

Операндами преобразований могут быть материальные, энергетические и информационные объекты, а также живые существа, в частности, люди (например, в процессе обучения операндами являются студенты).

Одним из основных операндов в техническом процессе является материя - S, применение которого может осуществляться в виде таких преобразований, как переработка, обработка, транспортировка и хранение.

Операнд энергия – En может быть подвергнут следующим преобразованиям: превращение, трансформирование, транспортировка, накопление.

Операнд информация - I может быть преобразован, изменена его форма ,он может быть передан и накоплен.

Человек-Me в качестве операнда может выступать, например, в медицинском учреждении, где ему могут изменить существо отдельного органа (вместо потерянной ноги –протез), изменить его форму (был болен - выздоровел), транспортировать (доставка в лечебное учреждение).

Операндами в высшем учебном заведении являются студенты, которые из абитуриентов в процессе обучения превращаются в специалистов.

Технический процесс как сложная система может быть разбит на подпроцессы ,а последние на операции (рис. 1.10).

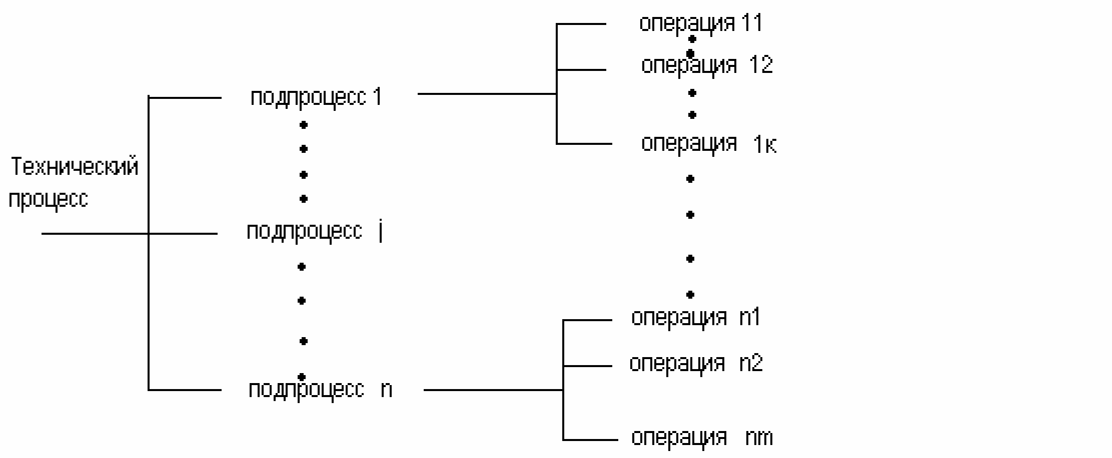


Рис. 1.10

Операцией называется элементарный процесс, соответствующий одному рабочему действию (обточка, нагрев, транспортировка, заправка топливом и пр.).

Согласованная совокупность операций представляет собой технологический процесс.

Для реализации технологического процесса одних рабочих операций, как правило , недостаточно. Возникает необходимость в ряде вспомогательных операций. Могут быть следующие виды вспомогательных операций:

-подготовительные операции (набор инструмента, подведение суппорта и пр.);

-операции обслуживания (заточка инструмента, смазка, удаление стружки и пр.);

-операции управления и регулирования (изменение рабочего режима).

В техническом процессе неизбежны также побочные входы и выходы. Примеры побочных входов: смазочные материалы, катализаторы; примеры побочных выходов: дым, стружка, помеха и пр. Учитывать побочные выходы необходимо, т. к. они могут быть вредными как для обслуживающего персонала, так и для окружающей среды.

1.4. Анализ и синтез в системных исследованиях

Анализ (analisys- греч. расчленение) - термин, имеющий два основных значения. Во-первых, это мысленное или реальное разделение целого на части и метод познания, основанный на этом. Во-вторых, это синоним научного исследования вообще, именно, когда говорят «подвергнуть анализу».

Синтез (synthesis-греч. соединение) – термин, означающий мысленное или реальное соединение частей в единое целое и метод познания, основанный на этом принципе.

Наиболее четко и последовательно понятия анализ и синтез разделяются в химии. Здесь анализ химически сложного соединения означает выявление составляющих элементов этого соединения. Разработкой методов анализа в этом случае занимается специальное направление этой науки - аналитическая химия. Синтез в химии означает получение известного сложного вещества из неизвестных его составляющих или получение нового ранее неизвестного вещества с заданными свойствами.

Такое четкое и категорическое разделение анализа и синтеза сделать нельзя , если заниматься исследованием систем, особенно сложных.

Если расчленить систему на составляющие ее элементы, то утрачивается одно из основных свойств системы- ее эмерджентность или целостность, обладающая свойствами, не сводящимися к свойствам отдельных составляющих систем.

Кроме того, исчезают при расчленении системы на части и существенные свойства ее частей, выделенных из целого. Так, отделенная от летательного аппарата система управления своей функции как отдельная часть выполнять не может. Способность управлять самолетом эта система может реализовать только в единстве с летательным аппаратом.

С другой стороны, синтез - объединение элементов в систему, понимание целого невозможно без вскрытия роли (функций) элементов, составляющих систему.

Таким образом, анализ и синтез дополняют, но не заменяют друг друга. Системный подход совмещает оба рассмотренных метода.

Это совмещение наглядно проявляется в разработке структур системы. Напомним, что структура означает устойчивую связь и взаимодействие элементов объекта как целого. При этом рассматриваются не любые возможности связи, а только те, которые необходимы для осуществления цели системы.

Таким образом, с одной стороны, структура отражает разделение объекта на элементы, с другой стороны, она отражает взаимосвязь, целостность объекта, т. е. в структуре системы совмещены и анализ, и синтез системы.

Представляется, что само название науки «системный анализ» является неточным, поскольку он противоречит понятию целостности системы. Однако, это название в настоящее время является устоявшимся и означает по сути дела исследования систем вообще (и анализ, и синтез).

2. ЭФФЕКТИВНОСТЬ СИСТЕМ И ПОКАЗАТЕЛИ ИХ КАЧЕСТВА

2.1 Определение понятий «эффективность и качество»

Для оценки систем вводится следующая иерархия показателей, которые характеризуют их на различных уровнях подробности: эффективность, показатели качества, технические параметры. Под эффективностью системы будем понимать степень её приспособленности к достижению поставленных задач.

Это определение соответствует приведенному в п. 1.2.1 заключению, что аспект цели является одним из важнейших при характеристике системы.

Сказанное относительно эффективности системы полностью относится и к понятию эффективности операции. Под операцией понимают любое мероприятие, объединенное единым замыслом и направленное на достижение определенной цели [19].

Для количественной оценки эффективности используют показатели эффективности. Конкретный вид показателя эффективности и его содержательная сущность зависит от вида системы и целей, которые перед ней стоят.

Выбор показателей эффективности и качества относится к наиболее сложному этапу системного анализа.

2.2 Подходы к выбору математических выражений для показателей эффективности и качества

В общем виде для критерия эффективности можно выделить следующие категории фактов, от которых зависит этот показатель:

- известные параметры системы и условия её функционирования (α_1, α_2, …..), которые являются постоянными и изменены быть не могут;
- неизвестные условия или факторы (Y_1, Y_2, …..);

- элементы выбора в ходе решения задачи ($x_1, x_2, \ldots$).

Таким образом, в общем виде можно записать

$$W=W(\alpha_1,\alpha_2\ldots Y_1, Y_2,\ldots, x_1,x_2,\ldots) \qquad (2.1)$$

Прежде всего отметим, что показатели могут быть единичными и комплексными. Единичный показатель количественно характеризует только одно свойство системы.

Комплексный показатель отражает не менее двух свойств (задачей, целей) системы. В качестве примера единичных показателей назовём его показатели безотказности, к которым в теории надежности относятся следующие показатели: интенсивность отказов $\lambda(t)$, вероятность работы $P(t)$ на интервале времени от 0 до t , средняя наработка до отказа T_1.

В теории надежности и других специальных дисциплинах будут подробно рассмотрены эти и другие единичные показатели.

Сложнее дело обстоит с формированием комплексных (обобщенных) показателей. Для их формирования применяются различные способы. Если система многоцелевая и нет оснований отдавать предпочтение какой-либо одной главной цели, то формируется совокупность показателей эффективности.

Например, для авиационной транспортной системы применяется следующая совокупность показателей:

- безопасность полётов;
- регулярность вылетов;
- интенсивность использования и экономичности.

В этом случае формируется не комплексный показатель, а комплекс показателей.

Другим подходом является построение обобщенного показателя на основе аддитивных и мультипликативных преобразований частных критериев.

Общий вид мультипликативного обобщенного критерия

$$U=\prod_{i=1}^{n} W_i^{\lambda_i} \qquad (2.2)$$

где λ_I – некоторые вещественные числа (в частном случае $\lambda_I=1$).

Нередко в качестве обобщенного мультипликативного показателя берут показатель в виде дроби, в числителе которой ставят те показатели, которые желательно увеличить (положительные показатели), а в знаменателе – те, которые желательно уменьшить (отрицательные показатели).

$$U=\frac{W_1\cdot W_2\ldots W_m}{W_{m+1}W_{m+2}\ldots W_k}$$

Общим недостатком подобного рода обобщенных критериев является то, что недостаток эффективности по одному показателю всегда можно скомпенсировать за счет другого.

Аддитивные комплексные показатели формируются в виде суммы частных показателей

$$U=\Sigma\, a_i\, w_i\ , \qquad (2.4)$$

где a_i – положительные или отрицательные коэффициенты.

Положительные коэффициенты ставятся при тех частных показателях, которые желательно максимизировать, отрицательные – при тех, которые желательно минимизировать.

Обобщенный показатель, построенный на основе аддитивных преобразований, так же как и мультипликативный, не исключает возможность взаимных компенсаций значениями частных показателей. Однако, если в качестве коэффициентов a_i использовать меру ценности (полезности) соответствующего частного показателя, то обобщенный критерий будет более объективно соответствовать степени влияния на него частных показателей. Правда, в этом случае проблема перемещается в другую плоскость: как определить весовые коэффициенты a_i, чтобы они объективно отражали степень влияния каждого частного показателя на обобщенный. Один из распространенных способов решения этой задачи – экспертное оценивание коэффициентов a_i .

Экспертное оценивание может быть сделано, например, следующим образом.

Назначается группа из m экспериментов. Им предполагается оценить значение каждого из n частных показателей. Например, по стобалльной шкале. Каждый j-й эксперт назначает числа c_{ij}, соответствующие каждому i-му частному показателю.

Далее производится нормировка

$$a_{ij} = \frac{c_{ij}}{\sum_{i=1}^{n} c_{ij}} \qquad \sum_{i=1}^{n} a_{ij} = 1 \qquad (2.5)$$

Получено m наборов нормированных коэффициентов a_{ij} (по числу экспертов).

Усредняя, получаем значения каждого весового коэффициента

$$a_{ij} = \frac{1}{m} \sum_{j=1}^{m} a_{ij}$$

В некоторых случаях из множества частных показателей может быть выделен один, главный показатель w_i . Его и следует максимизировать. На остальные частные показатели накладываются ограничения (устанавливаются допустимые значения W_g). Эти ограничения могут быть в виде неравенств

$$W_2>W_{g2},\ \ldots,\ W_m>W_{mg},\ W_{m+1}< W_{gm+1},\ \ldots\ W_k < W_{gk} \qquad (2.7)$$

Приведенная запись означает, что первые m-1 частных показателей не могут быть меньше, а остальные k-m показателей не должны превышать установленных допустимых значений.

Ограничения могут быть, например, на стоимость объекта, на его вес или габариты.

2.3 Задачи выбора

2.3.1. Постановка задач выбора

Выбор является действием, придающим всей деятельности целенаправленность, подчиняя деятельность определенной цели или совокупности целей.

Задачи выбора чрезвычайно разнообразны, различны методы их решения и способы описания задач. В настоящее время сложились и используются три основных языка выбора: критериальный, бинарных отношений и функции выбора [51]. В настоящем пособии рассматривается только критериальный язык.

Критериальный язык является наиболее развитым и наиболее часто употребляемым в технических приложениях. Основан критериальный язык на предположении, что качество функционирования каждой системы или операции можно оценить специальным показателем (критерием) эффективности или качества, которые рассмотрены в п.п. 2.1 и 2.2. Эти показатели называют также целевой функцией, функцией предпочтения, функцией полезности и т.п. Общий вид критерия эффективности был приведен ранее (2.1). Из тех факторов, определяющих этот показатель, (известные параметры α_1, $\alpha_{2....}$, неизвестные факторы Y_1, Y_2..., и элементы выбора x_1, x_2,…), выбору подлежат значения величин x_1, x_2, …. В зависимости от конкретной задачи содержательная сущность этих параметров может быть самой разнообразной.

Сам выбор осуществляется из некоторого множества альтернатив. Это множество может состоять из ряда одноименных систем (действий) с различными значениями величин x_1, x_2, …. или разнообразных систем (действий), но предназначенных для решения одной и той же задачи.

Кроме наличия множества альтернатив должна быть определена цель, ради которой предстоит осуществить выбор. Эта цель может быть выражена непосредственно целевой функцией (критерием эффективности) или быть сформулирована только качественно.

Сформулируем в общем виде постановку задачи выбора для случая, когда имеется один элемент выбора $x_{1,}$ и выбор необходимо сделать между двумя альтернативами x_1^1 x_1^2. Пусть соответствующие целевые функции (критерии эффективности) суть W_1^1 $(\alpha_1,\alpha_2\ldots Y_{1,} Y_{2,\ldots,} X_1^1)$ и $W_1^2(\alpha_1,\alpha_2\ldots Y_{1,} Y_{2,\ldots,} X_1^2)$.

Альтернатива X_1^1 предпочтительнее X_1^2 , если W_1^1 $(\alpha_1,\alpha_2\ldots Y_{1,} Y_{2,\ldots,} X_1^1) >$ $W_1^2(\alpha_1,\alpha_2\ldots Y_{1,} Y_{2,\ldots,} X_1^2)$, если система (действие) тем лучше, чем больше показатель W_1 и W_1^1 $(\alpha_1,\alpha_2\ldots Y_{1,} Y_{2,\ldots,} X_1^1)$, если система (действие) тем лучше, чем меньше показатель $W_{1.}$

Методы решения задачи выбора даже в такой простейшей постановке весьма разнообразны и для их разработки применяются такие научные дисциплины как теория оптимизации, математическая статистика, математическое программирование, теория случайных процессов, теория массового обслуживания и метод сетевого планирования.

2.3.2 Задачи оптимизации

Сущность оптимизации состоит в выборе таких значений x_i , чтобы показатель эффективности W обратился в максимум (минимум).

Если все Y_j=0, то имеет место детерминированный случай. Если хотя бы одно Y_j=0, то имеет место оптимизация выбора в условии неопределенности.

Для детерминированного случая имеем следующую общую зависимость

$$W=W(\alpha_1,\alpha_2\ldots,x_{1,}x_2\ldots) \qquad (2.8)$$

В первую очередь следует попытаться построить эту математическую функцию в явном виде. Если есть возможность получить математическую функцию, где аргумент – искомая величина, то поиск оптимального значения сводится к обычной математической процедуре поиска экстремума, дифференцированию по этому аргументу и приравниванию нулю производной.

Задача оптимизации в условиях неопределенности формулируется следующим образом: при заданных условиях $\alpha_1,\alpha_2\ldots$ с учетом неизвестных факторов Y_1,Y_2 …. Найти такие элементы решения $x_{1,}$ x_2 … , которые по возможности обращали бы в максимум показатель эффективности W. Здесь оговорка «по возможности» отражает неопределенность, которая возникает вследствие неизвестности значения факторов $Y_1,Y_2\ldots$

Если неизвестные факторы Y_1,Y_2 …. являются случайными величинами, для которых известны их функции распределения (или они могут быть получены из статистических данных), то для решения используют следующие приёмы: искусственное сведение к детерминированной схеме и «оптимизация в среднем».

Первый приём заключается в том, что вместо случайных величин используют их математические ожидания. Заметим, что чем меньше дисперсия величины Y_j, тем точнее результат. Если же распределение случайной величины «размытое» (дисперсия большая), то следует с осторожностью пользоваться способом замены случайной величины её математическим ожиданием, т.к. в том случае неизбежны большие ошибки.

Суть «оптимизации в среднем» состоит в нахождении математического ожидания самого показателя эффективности при известных распределениях случайных величин Y_1, Y_2 …. Если плотность распределения этих величин есть $f\,(Y_1, Y_2\ \ldots)$, то

$$W=MW=\iint\ldots\int W(\alpha_1,\alpha_2\ldots,x_1,x_2\ldots)\,f\,(Y_1,Y_2\ \ldots)dY_1\cdot dY_2\ldots \qquad (2.9)$$

Кроме W желательно также определить и дисперсию этой величины D[W].

3. ВЫБОР ПРИ МНОГОКРИТЕРИАЛЬНЫХ ЗАДАЧАХ

3.1. Общие сведения о многокритериальных задачах оптимизации

Почти всякая сложная практическая задача принятия решения является многокритериальной. В многокритериальной задаче оптимизации сравнение решений по предпочтительности осуществляется не непосредственно, а при помощи заданных на множестве X всех альтернатив числовых функций f_1, f_2,….., $f_м$ (m>2), называемых частными (локальными) критериями.

Критерий f_i образует векторный критерий f=(f_1, f_2,….., $f_м$) . В п 2.2 настоящего пособия написано, что многокритериальную задачу можно свести к однокритериальной путём формирования общего критерия, построенного на основе аддитивных и мультипликативных преобразований частных критериев.

Недостатками такого вида обобщенных критериев является то, что для мультипликативного критерия недостаток по одному критерию может быть скомпенсирован за счет другого, а в аддитивных критериях возникает сложность с определением весовых коэффициентов частных показателей.

В том случае, когда среди множества критериев есть один доминирующий, то оптимизация проводится по этому критерию с положительными ограничениями на другие (2.12). Иногда используется комбинация двух критериев: суммарной стоимости продукции и затрат на её производство. В данном деле часто используют такие два соотношения: эффективность - стоимость.

Если f(x) – функция, характеризующая производство, и F(x) – функция, характеризующая затраты, то задача сводится к увеличению f(x) при одновременном уменьшении F(x) , т.е. необходимо добиваться максимума производства при минимуме затрат. Поскольку минимум затрат равен нулю, то с нулевыми затратами произвести какую-либо полезную работу нельзя. Естественно, такой крайний случай является бессмысленным, но сама тенденция произвести побольше с возможно меньшими затратами является разумной.

3.2. Множество Парето

Рассмотренные выше способы многокритериальности выбора предполагают выделение единственной «наилучшей» альтернативы. Способ, при котором производится выбор не одной альтернативы, а некоторого множества альтернатив, удовлетворяющих определённым условиям, представляет собой обобщенное понятие максимума числовой функции в случае их множества.

Подход к анализу многокритериальных задач в этом случае состоит в следующем. Для сокращения множества исходных вариантов исключаются из анализа те варианты решения, которые заведомо хуже других.

Предположим, что мы сделали некоторый выбор X. Далее предположим, что существует некоторый другой выбор X', такой, что для всех критериев $f_i(x)$ имеет место неравенство

$$f_i(x) > f_i(x) \quad i=1,2,\ldots n, \qquad (3.1)$$

причем хотя бы одно из неравенств строгое. Очевидно, в этом случае выбор x' предпочтительнее выбора x. Поэтому все векторы x, удовлетворяющие приведенным неравенствам следует исключить из рассмотрения.

Если имеется множество векторов X*, для которых не существует такого X', которое удовлетворяло бы неравенству 3.1, то такое множество значений X* называется множеством Парето.

Состояния X* называются оптимальными по Парето, если не существует других допустимых состояний, которые были бы не хуже и хотя бы для одного x_i лучше X*.

Многомерный случай трудно представить наглядно, поэтому рассмотрим графический двумерный случай. Пусть имеется две критериальные функции, определяющие цели системы $f_1(x)$ и $f_2(x)$, при этом для достижения наилучшего результата обе функции необходимо увеличить, т.е. $f_1(x) \rightarrow \max$ $f_2(x) \rightarrow \max$.

Возьмем плоскость в координатах (f_1, f_2). Тогда каждому допустимому значению переменной x отвечает одна точка на плоскости (f_1, f_2). Равенства $f_1= f_1(x)$ и $f_2=f_2(x)$ определят параметрическое задание некоторой кривой abcd на этой плоскости (рис. 3.1).

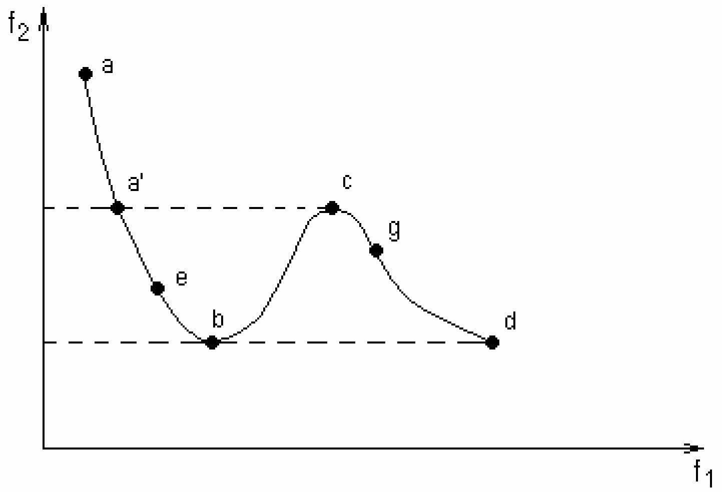


Рис. 3.1

Рассмотрим, к какому участку обозначенной кривой можно применить принцип Парето. Так, участок bc не может быть отнесён к множеству Парето, т.к. с ростом f_1 растёт и f_2. На этом участке изменению переменной *x* соответствует одновременное увеличение обеих целевых функций, следовательно, эти варианты следует исключить из рассмотрения.

Из тех же соображений исключаем и участок a'b. На этом участке любой точке *e* на участке cd найдётся некоторая точка g, для которой и $f_{1g} > f_{1e}$ и $f_{2g} > f_{2e}$.

На изображенном графике претендовать на принадлежность к множеству Парето могут участки cd и участок aa' (без точки a').

Принцип Парето заключается в том, чтобы в качестве решения принимались те значения x (тот вектор x), которые принадлежат множеству Парето, т.е. удовлетворяют условию (3.1).

Как видим, принцип Парето не выделяет одного единственного решения, а указывает на некоторое множество решений, сужая, таким образом, множество возможных альтернатив. В пределах этого выделенного множества окончательный выбор может быть сделан, исходя из каких-либо дополнительных, но существенных соображений, а предварительное выделение множеств предпочтительных решений облегчает этот выбор.

3.3. Численные методы построения множества Парето.

Построение множества Парето представляет собой довольно трудную задачу численного анализа, поскольку необходимо иметь дело с многомерной областью альтернатив.

Для иллюстрации способа построения множества Парето рассмотрим случай двух критериев $f_1(x)$ и $f_2(x)$, при этом улучшение качества системы происходит при увеличении обоих критериев, т.е.

$$f_1(x) \to \max \qquad (3.2)$$
$$f_2(x) \to \max.$$
$$X \in G_x$$

Область G_x – область существования рассматриваемых критериев.

Решение задачи начинается с построения области G_x в координатах x_1- x_2 (рис 3.2а)

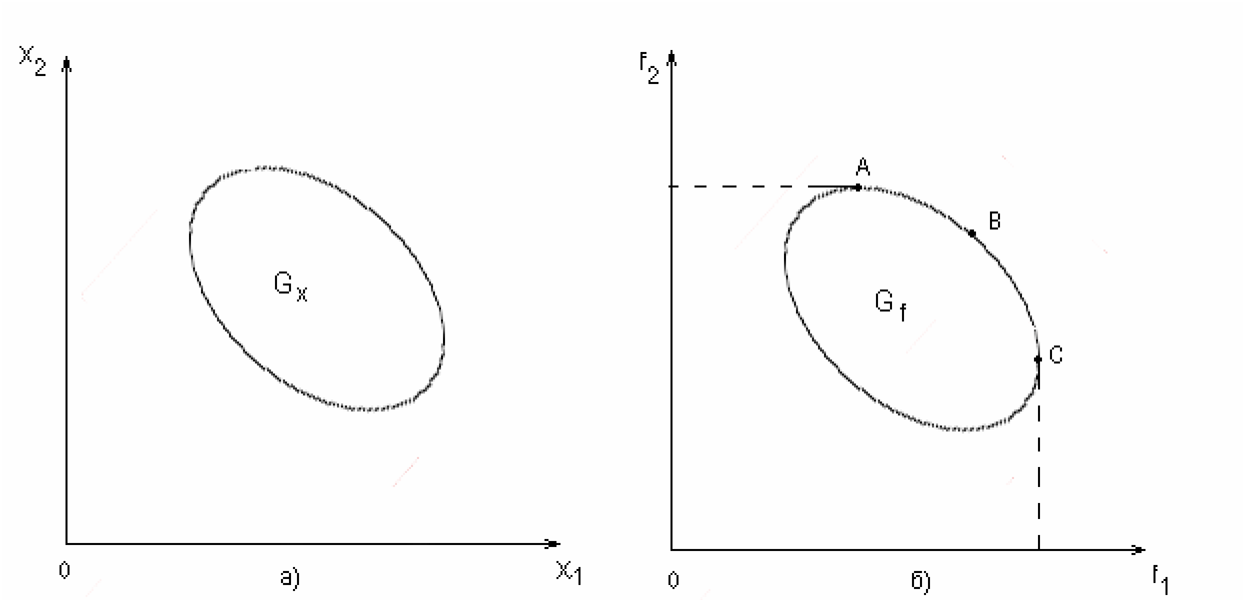


Рис. 3.2

Следующей процедурой является отображение множества G_x на множество G_f в плоскости критериев f_1- f_2 (рис. 3.2б), при этом f_1 и f_2 рассчитываются для каждой точки множества G_x т.е.

$$f_1 = f_1(x)$$
$$f_2 = f_2(x). \qquad (3.3)$$
$$f \in G_f$$

Полученное множество G_f в плоскости критериев есть множество достижимости. Из всего этого множества G_f к множеству Парето может быть отнесена только часть граничной области этого множества, именно дуга АВС, ограниченная сверху точкой А, соответствующей $f_2(a)=\max f_2$, и справа точкой В, соответствующей $f_1(b)=\max f_1$.

4. ВЕРОЯТНОСТНЫЕ ЗАКОНЫ РАСПРЕДЕЛЕНИЯ ПРИ АНАЛИЗЕ СИСТЕМ И ИХ ЭЛЕМЕНТОВ

4.1. Законы распределения дискретных случайных величин

4.1.1. Равномерный закон

Равномерный закон встречается тогда, когда все значения дискретной случайной величины равновероятны. При передаче цифровых данных, как правило, вероятности появления той или иной цифры одинаковы и имеют, таким образом, равномерное распределение.

Если имеется **n** значений дискретной случайной величины, то вероятность любого из этих значений P_i равна

$$P_i = \begin{cases} 0 \quad (-\infty < X_i < 1), \\ \dfrac{1}{n} \quad (1 \le X_i \le n) \\ 0 \quad (n < X_i < \infty). \end{cases} \qquad (4.1)$$

Функция распределения F(X) имеет ступенчатый вид с одинаковыми приращениями при каждом значении X

$$F(X) = \begin{cases} 0 \quad (-\infty < X_i < 1), \\ \sum\limits_{x=1}^{x} P_i = \dfrac{x_i}{n} \quad (1 \le X_i \le n) \\ 1 \quad (n < X_i < \infty) \end{cases} \qquad (4.2)$$

Графики величин P_i и F(X) приведены на рис 4.1

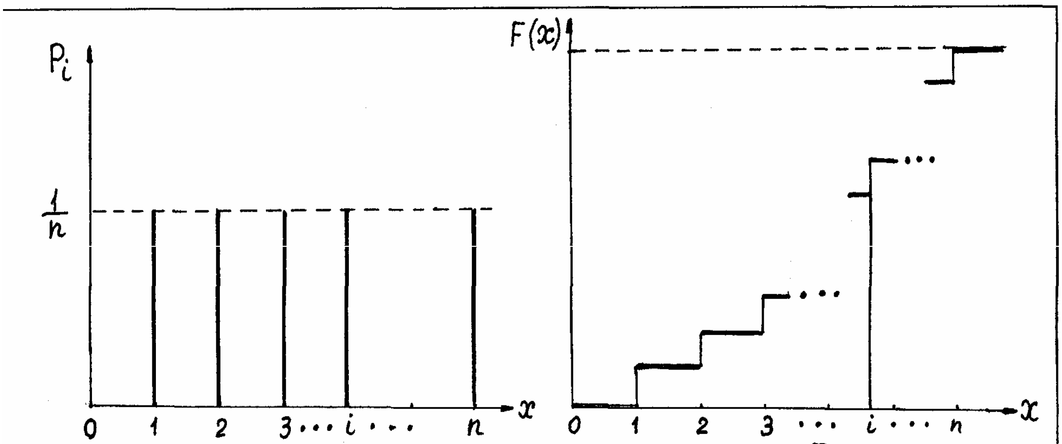


Рис.4.1

4.1.2. Биноминальный закон распределения (закон Бернулли)

Биноминальный закон даёт вероятность того, что в последовательности из **n** независимых испытаний событие наступает ровно **k** раз. В каждом из испытаний интересующее нас событие может произойти с одной и той же вероятностью **p**, соответственно не появление этого события **q=1-p.**

Вероятность того, что все первые **k** испытаний приведут к появлению искомого события, а при остальных **n-k** испытаниях его не будет, равна $\mathbf{p^k q^{n-k}}$ **.**
Вероятность любой другой комбинации **k** положительных и **n-k** отрицательных исходов также равна $\mathbf{p^k q^{n-k}}$.Число таких комбинаций равно числу сочетаний из **n** элементов по **k** т.е. $\underset{n}{\overset{k}{C}}$ **,** следовательно , искомая вероятность появления **k** событий в последовательности из **k** испытаний равна :

$$\mathbf{p_i = p(n_1,k) = c_n^k p^k\ q^{(n-k)}} \qquad 4.3$$

Для заданного значения **n** и **k=0,1,2,… n** эти вероятности являются последовательными членами разложения по формуле бинома Ньютона выражения $\mathbf{(p+q)^n}$ т.е.

$$\mathbf{(p+q)^n = cn^0 p^0 q^n c_n 1 p^1 q^{n-1} + c_n^2 p^2 q^{n-2} + \ldots + c_n^k p^k q^{(n-k)} + \ldots + c_n^n p^n q^0 = 1} \quad (4.4)$$

Отсюда и название - биноминальный закон. Схема биноминального распределения впервые рассмотрена Я. Бернулли, поэтому этот закон называют также законом Бернулли.
Функция распределения для рассматриваемого закона имеет вид:

$$\mathrm{F(x)} = \sum_{k=0}^{i} P(n.k) = \sum_{k=0}^{i} \underset{n}{\overset{k}{C}}\ \mathrm{p^k\ q^{n-k}} \qquad (4.5)$$

На рисунке 4.2 приведены графики величин **p(n k)** и **F(x)** для **p=q=0.5** и **n=6**
Функция частот $\mathbf{p_i}$ биноминального закона, как это следует из формулы (4.3), зависит от **p** и **n**
Биноминальное распределение применяют при определении качества принимаемой продукции выборочным методом контроля, находит применение при оценке вероятности безотказной работы технических систем, работающих в циклическом режиме.

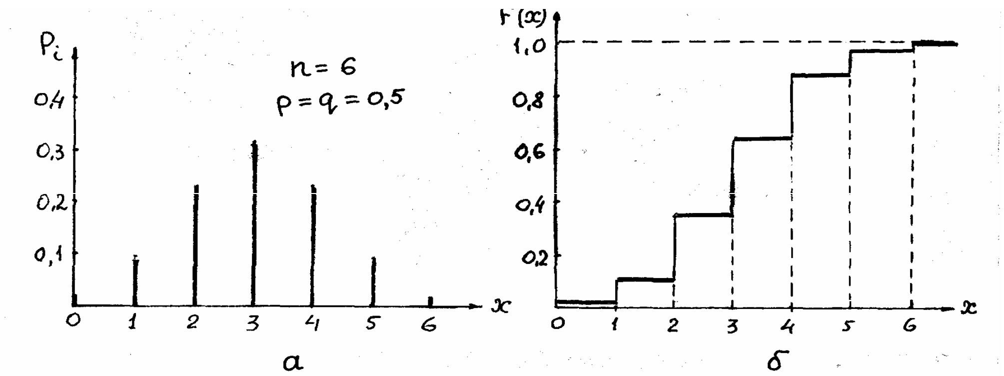


Рис.4.2

4.1.3. Закон Пуассона

Закон Пуассона является предельным случаем биноминального распределения, когда вероятность **p**. осуществления интересующего нас события в единичном эксперименте очень мала, но число экспериментов **n**, производимых в одной серии, достаточно велико. При этом произведение **n p** стремится к некоторой постоянной положительной величине λ.

$$\lambda = \lim_{n\to\infty\ p\to 0} np \qquad (4.6)$$

Поэтому закон Пуассона часто называют также законом редких событий.

Функция частот дискретного распределения Пуассона для целочисленных значений **k** имеет вид:

$$p_\kappa = \frac{\lambda^\kappa}{\kappa!} e^{-\lambda} \qquad (4.7)$$

P_k есть вероятность того ,что рассматриваемое событие в достаточно длинной серии испытаний появится ровно **k** раз.
Функция распределения описывается уравнением:

$$\mathbf{F(x)} = \sum_{k=0}^{\infty} P_{\mathbf{k}} = \sum_{k=0}^{\infty} \frac{\lambda}{k!} e^{-\lambda} \qquad (4.8)$$

Параметр λ. имеет различный физический смысл в зависимости от существа рассматриваемой задачи. Так , например , если рассматривается работа телефонной станции, то параметр λ есть среднее число вызовов за некоторое время **t.**

Закону Пуассона подчиняется также число распавшихся ядер радиоактивного вещества и т.п.

С помощью этого закона описываются такие процессы, как появление внезапных отказов в сложных системах, число отказов однотипного оборудования за данный интервал времени и т.п.

Распределение Пуассона широко применяется при контроле производства и качества продукции. Точное решение возникающих при этом вопросов даётся формулой биноминального распределения, однако если доля брака **p** не больше **0,1** , а проверяемая партия большая , то решение по формулам биноминального распределения требует продолжительных вычислений , тогда как с достаточной для практики точностью задача решается непосредственно при помощи распределения Пуассона.

4.2. Закон распределения непрерывных случайных величин

4.2.1. Нормальный закон распределения

Нормальный закон распределения непрерывных случайных величин (часто называемый законом Гаусса) занимает среди других законов распределения особое положение .Его главная особенность состоит в том ,что он является предельным законом , к которому приближаются другие законы распределения при весьма часто встречающихся типичных условиях , например , распределение суммы независимых случайных величин при достаточно большом числе слагаемых стремится к нормальному .

Нормальный закон распределения характеризуется плотностью вероятности вида:

$$\mathbf{F(x)} = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-m)^2}{2\sigma^2}} \qquad (4.9)$$

Как видим этот, закон является двухпараметрическим: величина **m** есть математическое ожидание, а величина σ **-**среднее квадратическое отклонение величины **X**
Функция нормального распределения имеет вид:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-(t-m)^2}{2\sigma^2}} dt \qquad (4.10)$$

Кривая плотности распределения нормального закона имеет симметричный холмообразный вид (рис 4.3). Максимальная ордината кривой равна $\frac{1}{\sigma\sqrt{2\pi}}$, соответствует точке **x=m.** По мере удаления от точки **m** величина функции **f(x)** уменьшается и асимптотически приближается к оси абсцисс правее и левее математического ожидания.

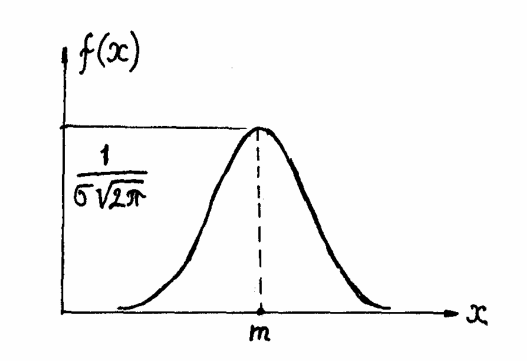


Рис. 4.3

В случае **m=0** и σ =1 распределение является центрированным и нормированным, для которого плотность вероятностей выражается формулой

$$\varphi_0(X) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \qquad (4.11)$$

Соответственно функция нормированного и центрированного распределения имеет вид:

$$F_0(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt \qquad (4.12)$$

Эту функцию часто называют стандартной функцией нормального распределения.

Функции **φ0(x)** и **F0(X)** табулированы, при этом в таблицах даются только положительные значения аргумента. Вход в таблицу для значения аргумента, математического ожидания **m** и среднего квадратичного отклонения **σ** производится по значению величины

$$S = \frac{x-m}{\sigma} \qquad (4.13)$$

Если аргумент отрицателен, то используется четность нормированной плотности распределения

$$\varphi_0(S) = \varphi_0(-S) \qquad (4.14)$$

Значение $F_0(S)$ в этом случае определяется по формуле:

$$F_0(S) = 1 - F_0(S) \qquad (4.15)$$

Хотя нормальное распределение допускает возможность принятия случайной величиной отрицательного значения, это распределение часто является хорошей моделью для описания различных технических параметров, в том числе показателей надежности.

В теории надежности нормальным распределением описывают наработки на отказ элементов технических устройств, вследствие их износа и старения. В этом случае обычно наработка на отказ измеряется в часах, циклах работы и т.п., т.е. является существенно положительной величиной.

Следует также иметь в виду, что в технических приложениях обычно имеют дело с такими значениями параметров **m** и **σ** (m >> o), что возникающие ошибки являются незначительными.

В приложении 1 приведена таблица 1.1 функции нормального распределения для функции вида (4.12).С помощью таблиц вероятностей нормального закона решаются задачи определения значений вероятностей попадания случайной величины **X** в некоторый интервал **(x_i, x_{i+1})** вероятностей выполнения неравенства **x ≤ a** или неравенства **x > b.**

На практике встречаются случаи, когда необходимо определить, какое значение случайной величины соответствует заданной вероятности. Эта задача решается с помощью специальной числовой характеристики распределения вероятностей, называемой ***квантилью.***

Для действительной случайной величины **X** с функцией распределения **F(x)** квантилью порядка **p (0 < p < 1)** называется число $К_p$ такое, что $F(К_P) < p$.

Если **F(x)** - непрерывная строго монотонная функция, то $К_p$ - единственное решение уравнения **F(x) = P.**, т.е. $К_p$ - функция **p**., обратная функции **F(x)** .,

$$К_p = G(p) = F^{-1}(p.) \qquad (4.16)$$

Если **F(x)** непрерывна и $p_2 > p_1$, то вероятность неравенства $Кp_1 < x < Кp_2$ равна p_2-p_1.
Квантиль $К_{1/2}$ есть медиана случайной величины **X**. Квантили $К_{1/4}$ и $К_{3/4}$ называются квартилями, а $К_{0.1}$ $К_{0.2}$ …$К_{0.9}$ - децилями.

В приложении 1 приведена таблица 1.2 квантилей нормального распределения. Понятие ***квантиль*** и соответствующие таблицы применяются также и для других видов распределения.

4.2.2 Экспоненциальный закон распределения

Распределение случайной положительной величины называется экспоненциальным, если его плотность распределения вероятностей имеет вид

$$f(x) = \lambda e^{-\lambda X} \qquad (4.17)$$

Экспоненциальный закон является однопараметрическим, λ - параметр распределения, является строго положительной константой, **e = 2,71828...** - основание натурального логарифма.

Функция экспоненциального распределения может быть определена путем интегрирования функции **f(x)**

$$F(X) = \int_0^x \lambda e^{-\lambda t} dt = 1 - l^{-\lambda t} \qquad (4.18)$$

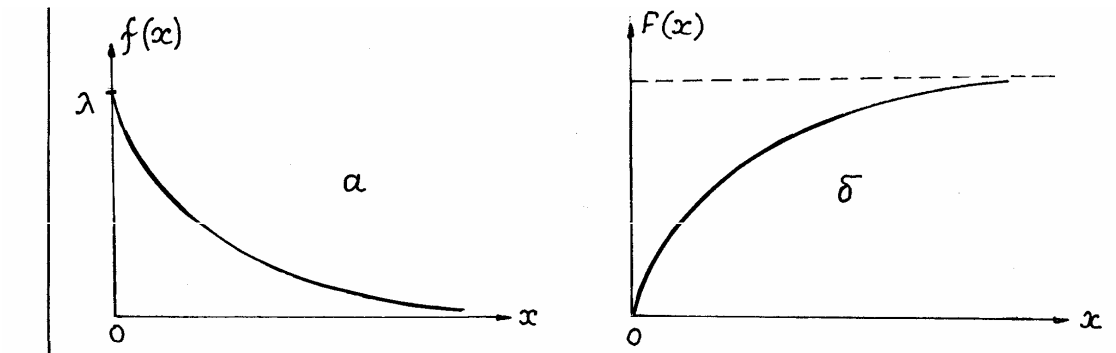
Математическое ожидание **M(x)** и среднее квадратическое отклонение σ_x экспоненциального распределения совпадают и равны обратному значению параметра λ,

$$M(X) = \sigma_x = \frac{1}{\lambda} \qquad (4.19)$$

Графики плотности вероятности экспоненциального распределения и функции **F(x)** приведены на рис. 4.4.

Закон экспоненциального распределения нашел широкое применение в технических приложениях. Экспоненциальное распределение часто называют основным законом надежности. Он описывает надежность работы изделия в период его нормальной эксплуатации, когда постепенные отказы еще не проявляются и надежность характеризуется внезапными отказами. Параметр λ, в этом случае характеризует интенсивность отказов.

Показательный закон играет большую роль в теории Марковских случайных процессов. Экспоненциальному распределению подчиняется величина промежутка между двумя смежными событиями простейшего (Пуассоновского) потока событий. Это же распределение широко используется в теории массового обслуживания.



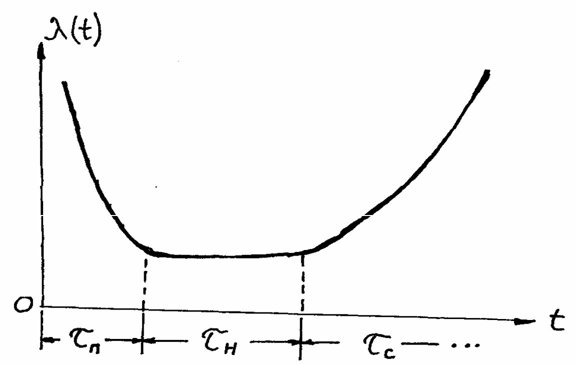
а) график плотности **f(x)**
б) функция **F(x)** экспоненциального распределения

Рис. 4.4

4 . 2 . 3 . Распределение Вейбулла

При рассмотрении экспоненциального закона распределения было подчеркнуто, что параметр этого закона -λ, является строго положительной константой. При рассмотрении надежности технических систем этот параметр характеризует интенсивность отказов.

Однако многочисленные экспериментальные данные в области анализа надежности технических элементов и систем показывают, что в значительном числе случаев параметр **λ**, изменяется так, как это изображено на кривой, представленной на рис. 4.5.



$\tau_п$ – **период "приработки" ("обкатки")**
$\tau_н$ – **период нормальной эксплуатации**
$\tau_с$ – **старение и износ**

Рис. 4.5

Из этого графика видно, что весь интервал времени работы технической системы можно разбить на три периода. В начале первого периода величина **λ(t)** имеет высокое значение и довольно быстро убывает. Это объясняется наличием в технической системе элементов с явными и скрытыми дефектами, которые приводят к относительно быстрому выходу из строя этих элементов. Этот период принято называть периодом «приработки» («обкатки»).

Второй период - период нормальной эксплуатации, когда интенсивность отказов приблизительно постоянна и относительно невысока.

Последний период - период старения и износа, во время которого отказы элементов системы происходят за счет необратимых физико-химических явлений.

Рассмотренный ход изменения функции **λ(t)** может быть описан классом степенных зависимостей, имеющих следующий вид

$$\lambda(t) = \lambda_0 \alpha (\lambda_0 t)^{\alpha - 1} \qquad (4.20)$$

где λ_0 и α - некоторые положительные числовые параметры, причем значения, $\alpha < 1$, $\alpha = 1$ **и** $\alpha > 1$ отвечают поведению функции отказов соответственно в период приработки, нормальной эксплуатации и старения.

Количество отказавших за время **t** элементов для нестационарного потока отказов, **λ(t),** очевидно, равно

$$n_{отк}(t) = \int_0^t \lambda(t) dt \qquad (4.21)$$

Функция распределения, которая в данном случае будет означать вероятность отказа к моменту времени **t**, может быть записана в следующем виде

$$F(t) = 1 - e^{-\int_0^t \lambda(t) dt} \qquad (4.22)$$

Если теперь в эту формулу подставить выражение для **λ(t)** из (4.20), и проинтегрировать показатель степени, то получим:

$$F(t) = 1 - e^{-(\lambda_0 t)^{\alpha}}, \quad t \geq 0 \quad (4.23)$$

Соответственно плотность вероятности

$$f(t) = \lambda_0 \alpha t^{\alpha - 1} \times e^{-(\lambda_0 t)^{\alpha}}, \quad t \geq 0 \quad (4.24)$$

Выражения (4.23) и (4.24) и есть распределение Вейбулла (в литературе [7,8] встречается также наименование "Распределение Вейбулла - Гнеденко"). Как видим, распределение Вейбулла является двухпараметрическим.

Рассмотренное ранее экспоненциальное распределение является частным случаем распределения Вейбулла, а именно соответствует значениям $\alpha = 1$, т.е. описывает процесс отказов в период нормальной эксплуатации.

В прикладных задачах теории надежности [30] распространена следующая форма записи закона Вейбулла

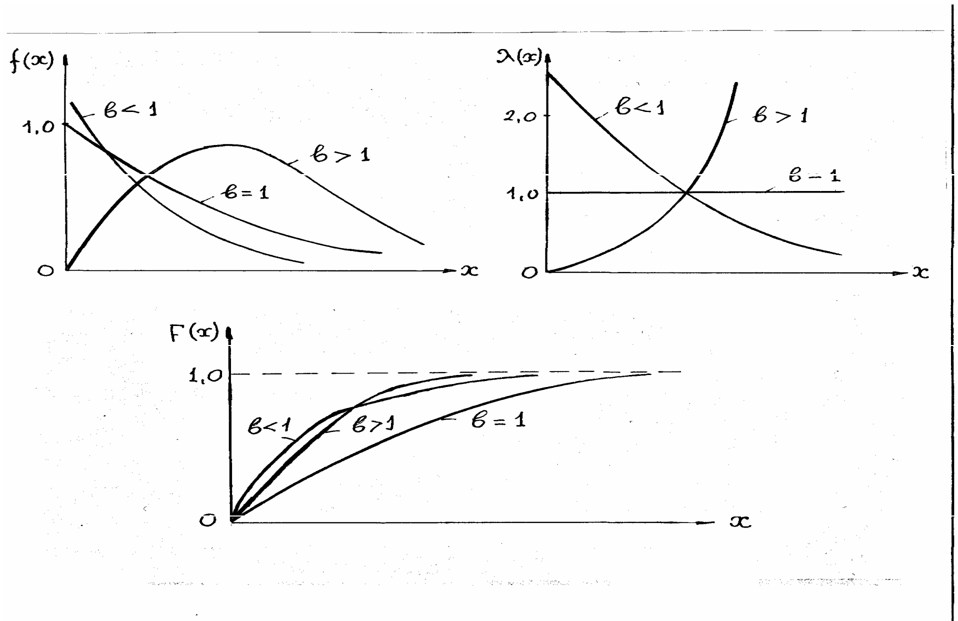
$$f(x) = \frac{b}{a} \left(\frac{x}{a}\right)^{b-1} e^{-\left(\frac{x}{a}\right)^{d}} \qquad (4.25)$$

$$F(x) = 1 - e^{-\left(\frac{x}{a}\right)^{b}} \qquad (4.26)$$

Сравнение с исходной формулой (4.24) показывает, что $b = \alpha$ и $a = \frac{1}{\lambda_0}$.

Величина **b** является параметром формы распределения, а величина **a** - параметром масштаба. В. приложении 1 приводится таблица 1.3 для определения коэффициентов **a** и **b** .Вход в таблицу производится по величине $v = \frac{\sigma}{t_{cp}}$.

На рис. 4.6 приведены кривые распределения Вейбулла и зависимостей **λ(t)** при различных значениях параметров **a** и **b** для закона Вейбулла в записях (4.25) и (4.26).



Графики зависимостей f(x), F(x) и λ(x) распределения Вейбулла для различных значений параметра b.

Рис. 4.6

5 СТАТИСТИЧЕСКАЯ ОБРАБОТКА ДАННЫХ

5.1 Основные задачи, решаемые статистическими методами

В естественных и технических науках понятие "*статистика*" означает анализ массовых явлений, основанный на применении методов теории вероятностей. Методы статистической обработки исходных данных, разработка и использование которых апеллируют к вероятностной природе этих данных, принято называть методами *математической статистики*.

Кратко рассмотрим основные задачи, решаемые методами математической статистики.

Первой задачей является задача собственно статистической обработки полученных экспериментальных данных. Результатами такой обработки могут быть таблицы распределения, ряды распределения или гистограммы.

Вторая задача состоит в определении закона распределения случайной величины, статистические данные которой обработаны. Для этого по внешнему виду, например, гистограммы распределения, выдвигается гипотеза о возможном законе распределения исследуемой случайной величины, и затем проверяется правдоподобие выдвинутой гипотезы с помощью того или иного критерия согласия.

Третья задача заключается в определении неизвестных параметров закона распределения, установленного с помощью критериев согласия, с использованием тех же статистических данных.

Наконец, четвертая основная задача состоит в выборе, в широком смысле этого слова, наиболее предпочтительного поведения в условиях существенной неопределенности, обусловленной случайным характером величин, которые определяют поведение принимающего решение человека. На практике важное значение имеет статистический приемочный контроль, позволяющий дать обоснованное заключение о годности партии изделий или обоснованного выбора между двумя или несколькими изделиями (системами) или линиями поведения.

Кроме этих основных задач математической статистики, она применяется для анализа погрешности измерений и позволяет дать рекомендации о необходимой точности измерительных приборов (систем).

Статистические методы используются для установления связи между случайными величинами (методы корреляции и регрессии), исследование рядов динамики и прогнозирования изменения тех или иных параметров.

В физике и механике сложились целостные теоретические направления, использующие вероятно - статистические методы (статистическая физика и статистическая механика).В последние годы развилась новая отрасль знаний теория распознавания образов, позволяющая идентифицировать те или иные процессы, явления или изображения, относя их к определенному классу явлений.

5.2 Вариационный ряд, гистограмма и эмпирическая функция распределения

Исходным пунктом любого статистического исследования случайной величины **X** является совокупность из **n** наблюдений, в результате которых величина **X** принимает значения $\mathbf{X_1}$, X_2,.-.$\mathbf{X_n}$ Заметим, что среди этих значений некоторые могут быть и одинаковыми. Если эти значения охватывают все **N** возможных однотипных объектов, подлежащих исследованию, т.е. **n = N,** то это множество объектов называют *генеральной совокупностью.*

В том случае, если имеется такая совокупность **n** наблюдений, что **n < N,** то она называется *выборкой,* а величина **n** - *объемом выборки.*

На практике чаще всего имеют дело именно с выборкой, поскольку обследование всей генеральной совокупности бывает либо слишком трудоемко (слишком большое **N**), либо принципиально невозможно (в случае бесконечных генеральных совокупностей).

Выборка должна быть *представительной* (репрезентативной), т.е. она по своим статистическим свойствам должна отражать, представлять свойства всей генеральной совокупности.

Для наглядности полученный вариационный ряд преобразуют следующим образом. Ось абсцисс делят на интервалы $(\mathbf{x_i}, \mathbf{x_{i+1}})$ длины $\Delta \mathbf{X}=\mathbf{X_{i+1}}-\mathbf{X_i}$

Длину интервала следует выбирать такую, чтобы количество интервалов не было большим, но и не настолько малым, чтобы не искажать особенности распределения статистических данных.

Приближенно длина интервала может быть определена по формуле

$$\Delta\mathrm{X} = \frac{x_{\max} - x_{\min}}{1 + 3.2\lg n} \qquad (5.1)$$

После разбиения вариационного ряда на интервалы определяется значение статистической плотности распределения для каждого **i-го** интервала

$$f^*_i(x) = \frac{\Delta n_i}{n\Delta x}\ , \qquad (5.2)$$

где Δn_i -число значений членов вариационного ряда, попавших в i-й интервал. Чертёж значений $f^*_i(x)$, отложенных для каждого интервала, называется гистограммой частот.

Статистическая плотность распределения $f^{*}(x)$ является аналогом плотности распределения непрерывной случайной величины и имеет размерность, обратную размерности случайной величины **X:**

$$[f^{*}(x)] = [x]^{-1} \qquad (5.3)$$

Частость, отображающая вероятность нахождения величины **X** в i-ом интервале, равна

$$p_i^{*} = \frac{\Delta n_i}{n} \qquad (5.4)$$

Аналогом теоретической функции распределения **F(x)** является эмпирическая функция распределения $F^{*}(x)$ случайной величины **X**.

В том случае, если экспериментальные значения случайной величины $\mathbf{x_i}$ сгруппированы, функция **F*(x)** определяется соотношением

$$F_k^{*}(x) = \frac{n(x)}{n} \qquad (5.5)$$

где **n(x)** число чисел вариационного ряда, удовлетворяющее неравенству $\mathbf{x_к < x}$ Для случая группированных данных

$$F_l^{*}(x) = \frac{\sum\limits_{i=1}^{l} \Delta n_i}{n} \qquad (5.6)$$

или

$$F_l^{*}(x) = \sum_{i=1}^{l} P_i \qquad (5.7)$$

Δn_i - число значений членов вариационного ряда, попавших в i-й интервал;

Здесь P_i^{*} - частость, соответствующая **i**-му интервалу разбиения (5.4). Из определения эмпирической функции распределения непосредственно следует часто используемое другое название - *накопленная частость***.**

5.3. Полные и усеченные (цензурированные) выборки

На практике при организации наблюдений или испытаний систем и их элементов, например, при регистрации отказов техники, могут встречаться различные ситуации, отличные от простой регистрации результатов испытаний при фиксированном числе объектов.

Характерны в этом плане наблюдения за отказами элементов систем в процессе эксплуатации объектов, например летательных аппаратов. В этом случае, как правило, нет возможности так организовать эксплуатационные наблюдения, чтобы получить данные по надежности необходимого вида и в достаточном количестве.

В этом случае по тем или иным причинам производится прекращение эксплуатационных наблюдений.

Прекращение эксплуатационных наблюдений объекта до наступлений отказа называется цензурированием.

Выборка, которая формируется из обоих видов реализаций, т.е. содержит значения наработок до отказа и неполные реализации, называется цензурированной выборкой.

Могут быть однократно и многократно цензурированные выборки. В однократно цензурированной выборке значение времени цензурирования Θ для всех, не достигших времен отказа изделий одинаково, причем $\Theta \geq \mathbf{t_{откmax}}$, где $\mathbf{t_{откmax}}$ - наибольшая наработка до отказа изделий полной выборки. Многократно цензурированная выборка характеризуется значениями наработок до цензурирования, не равными между собой.

На рис. 5.1 приведена неранжированная временная диаграмма моментов отказов изделий для случая полной выборки. Для этого случая среднее значение времени отказа равно

$$T^*_{ср} = \frac{\sum_{i=1}^{N} t_i}{N} \qquad (5.8)$$

На рис.5.2 изображена неранжированная диаграмма однократно цензурированной выборки. Из общего числа **N** изделий у **n** изделий зарегистрированы отказы, а у **N – n** изделий прекращены наблюдения (произведено цензурирование) при **t =** Θ **.**

Для этого случая цензурированной выборки среднее значение времени отказа равно

$$T^*_{ср} = \frac{\sum_{i=1}^{n} t_i}{n} + \frac{(N-n)}{n}\Theta \qquad (5.9)$$

Вывод этой формулы с использованием метода максимального правдоподобия будет дан в п.6.2. Сейчас отметим, что второе слагаемое в формуле (5.9) учитывает то обстоятельство, что о **N - n** изделиях информации об отказах нет, но известно, что эти изделия до **t** ≤ Θ не отказали.

Более подробно вопросы однократного и многократного цензурирования будут изучаться в специальных учебных дисциплинах.

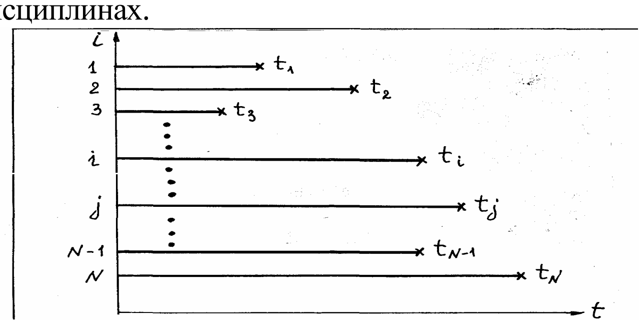


Рис. 5.1

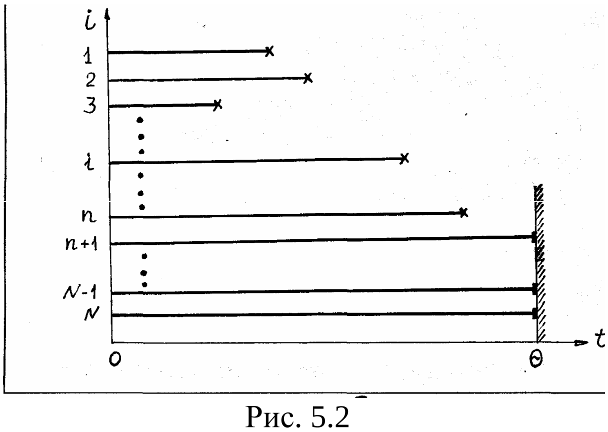


Рис. 5.2

6. ОПРЕДЕЛЕНИЕ ПАРАМЕТРОВ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ ПО ДАННЫМ ВЫБОРОК

6.1. Метод моментов

Для определения параметров распределения исследуемой случайной величины используется *метод моментов*. Понятие момента широко используется в механике для описания распределения масс (статистические моменты, моменты инерции и т.д.), действия силы на рычаг и т.п.

Частости, отражающие вероятность нахождения случайной величины **X** в **i**-ом интервале, можно рассматривать как силу, приложенную к середине **i**-го интервала.

Чаще всего на практике применяются моменты двух видов: начальные и центральные.

Начальный момент порядка **S** выражается формулой

$$\alpha_S = \sum_{i=1}^{k} x_i^S P_i^* \quad (6.1)$$

где P_i^* – частость, соответствующая **i**-му интервалу,

Xi - расстояние до средины интервала от начала координат (отсюда и название - начальный момент). При **S = 0** $\alpha_s = 1$, т.е. начальный момент нулевого порядка равен единице. При **S = 1** получаем

$$\alpha_S = \sum_{i=1}^{k} x_i^S P_i^* = M[x] = \bar{x}^* \quad (6.2)$$

т.е. начальный момент первого порядка есть математическое ожидание исследуемой случайной величины.

Центральные моменты - это моменты относительно математического ожидания $\bar{x}^*$.

Центральный момент порядка **S** выражается следующей формулой

$$\mu_S = \sum_{i=1}^{k} (x_i - \bar{x}^*)^S P_i^*$$

при **S = 0** $\mu_{s=s}$ **= 1,** т.е. нулевой центральный момент равен единице.

при **S=1**

$$\mu_1 = \sum^{k} (x_i - \bar{x}^*) P_i^* = \sum^{k} x_i p_i^* - \bar{x}^* \sum^{k} p_i^* = \bar{x}^* - \bar{x}^* = 0 \quad (6.4)$$

т.е. центральный момент равен нулю

при **S=2**

$$\mu_S = \sum_{i=1}^{k} (x_i - \bar{x}^*)^2 = D[x] \quad (6.5)$$

Второй центральный момент равен дисперсии, которая характеризует разброс случайной величины около математического ожидания.

Как известно $\sigma_x = \sqrt{D[x]}$

Математическое ожидание и дисперсия (или среднее квадратическое отклонение) - наиболее часто применяемые характеристики случайной величины. Для более подробного описания распределения применяются моменты высших порядков. Третий центральный момент служит для характеристики асимметрии, а четвертый центральный момент служит для характеристики "крутости" плотности распределения. При необходимости с этими характеристиками можно ознакомиться в литературе по теории вероятностей.

6.2. Метод максимального правдоподобия

Метод максимального правдоподобия - это один из методов нахождения *статистических оценок* неизвестных параметров распределения. Предполагается, что результаты наблюдений $\mathbf{x_1, x_2,...x_n}$ являются взаимно независимыми случайными величинами, принадлежащими одному закону распределения вероятностей.

Рассмотрим случай, когда распределение вероятностей зависит от одного параметра **ν**. В качестве оценок выбирают то значение параметра **ν**, при котором полученные результаты *"наиболее вероятны"*.

Смысл принципа *"наибольшей вероятности"* рассматриваемого метода заключается в следующем.

Вводится функция

$$\mathbf{L(x_b\ x_{2v}. .x_{n\,v,}) = f(x_1,v)\text{-}f(x_2,v)..\ .f(x_n,v),} \qquad (6.6)$$

где $\mathbf{f(x_i,v)}$ в случае *непрерывного распределения* есть *плотность вероятности* рассматриваемой случайной величины **X**, а в *дискретном* случае есть *вероятность* того, что случайная величина **X** примет значение $\mathbf{x_i}$.

Функцию $\mathbf{L(x_1, x_2,...x_n,v)}$ называют *функцией правдоподобия.* Оценкой максимального правдоподобия параметра **ν** называют такое его значение $\hat{v} = \hat{v}(x_1, x_2,...x_n),$ при котором функция $\mathbf{L(x_1, x_2,...x,v)}$ достигает наибольшего значения.

Для удобства определения максимума функции $\mathbf{L(x_1, x_2,...x_n)}$, рассматривают логарифм этой функции **InL**, так как максимум **InL** и максимум самой функции **L** совпадают. Уравнение

$$\frac{d}{dv}\ln L(x_1, x_2 x_n ... v) = 0 \qquad (6.7)$$

называют уравнением правдоподобия.
Учитывая формулу (6.7), получаем:

$$\sum_{i=1}^{n} \frac{\partial}{\partial v} \ln f(x_i, v) = 0 \qquad (6.8)$$

Проиллюстрируем применение метода максимального правдоподобия на нескольких примерах.

Пример 1. Наблюдалось **N** изделий, из них **n** изделий отказало. Искомый параметр $\boldsymbol{v} = \mathbf{P}$ - вероятность безотказной работы.

Функция правдоподобия

$$\mathbf{L(n,P) = P^{N\text{-}n}\,(1\text{-}P)^n} \qquad (6.9)$$

Уравнение правдоподобия

$$\frac{\partial L}{\partial P}=\frac{\partial}{\partial P}[P^{N-n}(1-P)^{n}]=0 \qquad (6.10)$$

Дифференцируем $\frac{\partial L}{\partial P}=(N-P)\cdot P\cdot(1-P)^{n}+P^{N-n}\cdot n(1-P)^{n-1}$

Разделив каждый член на $\mathbf{P^{N-n-1}(1-P)_{n-1}}$,получим

(N-n)(1-P)-P n=0, откуда **N-n=NP** и оценка

$$\hat{P}=\frac{N-n}{N}=1-\frac{n}{N} \qquad (6.11)$$

Пример 2. Имеем $\mathbf{x_1,x_2,\ldots x_n}$ случайных величин (наблюдений), распределение которых предполагается экспоненциальным, т.е. $F(x)=1-e^{-\lambda x}$ и $F(x)=\lambda e^{-\lambda x}$. Определить параметр λ этого закона по значениям экспериментальных данных. В соответствии с формулой (6.8) имеем

$$\sum_{i=1}^{n}\frac{\partial}{\partial\lambda}\ln\lambda e^{-\lambda x}=\sum_{i=1}^{n}\frac{\partial}{\partial\lambda}[\ln\lambda-\lambda x_i]=\sum_{i=1}^{n}[\frac{1}{\lambda}-x_i]=\frac{n}{\lambda}-\sum_{i=1}^{n}x_i=0$$

Отсюда непосредственно получаем:

$$\lambda=\frac{n}{\sum_{i=1}^{n}x_i} \qquad (6.12)$$

Так как для экспоненциального распределения $X_{cp}=\frac{1}{\lambda}$, то имеем

$$X_{cp}=\frac{\sum_{i=1}^{n}x_i}{n} \qquad (6.13)$$

Пример 3. Наблюдается **N** изделий. За период наблюдений Θ зафиксировано ***n*** отказавших изделий при значениях наработок

$\mathbf{t_1,t_2,\text{.-.}t_n}$, при этом $\mathbf{(t_i)max < 0}$. Таким образом, имеет место цензурированная справа выборка. Распределение, как и в предыдущем случае - экспоненциальное.

Для **n** значений $\mathbf{t_i} < \Theta$ каждому $\mathbf{t_i}$ можно в соответствии с **(6.6)** поставить в соответствие значение плотности вероятности $\mathbf{f(t_i.\lambda)}$ Для каждого же из неотказавших изделий мы можем утверждать, что вероятность его отказа после t > Θ равна 1 - F(Θ), а для всех **N-n** неотказавших изделий равна $\mathbf{[l\text{-}F(\Theta)]^{N-n}}$.

Таким образом, для случая цензурированной выборки функция правдоподобия запишется в следующем виде

$$L(\lambda)=[\prod_{i=1}^{n}f(t_i,\lambda)]\cdot[1-F(\Theta)]^{N-n} \qquad (6.14)$$

Составляем уравнение правдоподобия

$$\frac{\partial}{\partial\lambda}\ln L(\lambda)=\frac{\partial}{\partial\lambda}\sum_{i=1}^{n}\ln(f(t_i,\lambda)+(N-n)\ln[1-G(\Theta)]=0 \qquad (6.15)$$

Для экспоненциального распределения

$$f(t_i,\lambda)=\lambda e^{-\lambda t}....и....F(\Theta)=1-e^{-\lambda t}$$

Отдельные члены этого уравнения определяются следующим образом:

$$\frac{\partial}{\partial\lambda}\ln f(t_i,\lambda)=\frac{1}{\lambda}-t_i$$

$$\ln[1-F(\Theta)]=\ln[1-(1-e^{-\lambda\Theta})]=\ln[e^{-\lambda\Theta}]=\lambda\Theta$$

$$\frac{\partial}{\partial\lambda}[1-F(\Theta)]=\frac{\partial}{\partial\lambda}(-\lambda\Theta)=-\Theta$$

Таким образом, получаем

$$\frac{\partial}{\partial\lambda}\ln L(\lambda)=\frac{n}{\lambda}-\sum_{i=1}^{n}t_i-(N-n)\Theta=0$$

Откуда $$\frac{1}{\lambda}=T_{cp}=\frac{1}{n}\sum_{i=1}^{n}t_i+\frac{N-n}{n}\Theta \qquad (6.16)$$

Приведенные примеры показывают, что метод максимального правдоподобия является хорошим инструментом для обработки различного рода экспериментальных данных: для полных и цензурированных выборок, для дискретных и непрерывных случайных величин.

7. ПРОВЕРКА ГИПОТЕЗ О ХАРАКТЕРЕ ЗАКОНА РАСПРЕДЕЛЕНИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

7.1. Постановка задачи статистической проверки гипотез

Статистическая проверка гипотез - один из основных разделов математической статистики, в котором рассматриваются методы статистической проверки между статистическими данными и гипотезами об их вероятностной природе. При этом можно выделить два основных направления: проверка, принадлежит ли данная статистическая совокупность данных какому - то вероятностному закону распределения, и при наличии двух различных статистических данных определить, принадлежат ли они одной генеральной совокупности (одному распределению).

Постановка задачи первого направления состоит в следующем.

Предположим, что в результате эксперимента получена выборка $\mathbf{x_1,x_2,...x_n}$ из генеральной совокупности с известной функцией распределения **F(x).** По экспериментальным данным строится эмпирическая функция распределения **F*(x),** которая естественно является ступенчатой (рис. 7.1).

Статистической гипотезой называется любое предположение относительно вида теоретической функции распределения **F(x),** сделанное на основе выборки.

На рис.7.1 приведена предполагаемая функция **F(x).** Заметим, что практически о виде теоретической функции распределения удобнее судить по гистограмме частостей, построенной по экспериментальным данным.

Проверка статистической гипотезы заключается в выборе решения: принять гипотезу или отвергнуть ее.

Обычно предполагают, что имеются две непересекающиеся гипотезы. $\mathbf{H_0}$ и $\mathbf{H_1}$. Гипотезу $\mathbf{H_0}$ будем называть *основной*, а гипотезу $\mathbf{H_1}$, - *конкурирующей* или *альтернативной.* Заметим, что выбор, какую гипотезу принять за основную, а какую за альтернативную, условен. Однако, как правило, удобно основной гипотезой $\mathbf{H_0}$ называть конкретное предположение о виде теоретической функции распределения или предположение, важное с точки зрения решаемых практических задач.

Задача проверки статистических гипотез состоит в том, чтобы на основе выборки $\mathbf{x_1, x_2 \ldots x_n}$ принять (т.е. считать справедливой) либо основную гипотезу $\mathbf{H_0}$, либо альтернативную гипотезу $\mathbf{H_1}$

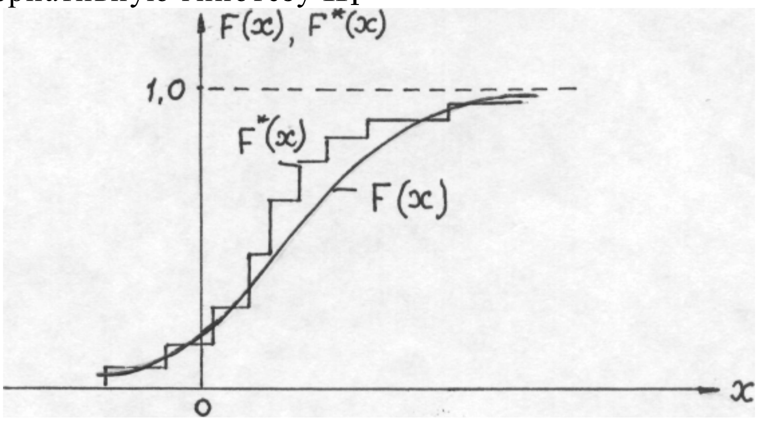


Рис.7.1

В настоящее время имеются специальные программные средства, позволяющие проверку статистических гипотез производить автоматически [4].

Второе направление проверки статистических гипотез состоит в выяснении принадлежности двух выборок одному распределению. В этом случае, имеем, например, две статистические выборки **(X)=$\mathbf{x_1\ x_2 \ldots x_к}$** **(Y)==$\mathbf{y_1\ y_2 \ldots y_l}$**
Возникает вопрос: принадлежат ли эти две выборки одному распределению (основная гипотеза **Ho**) или это разные распределения (альтернативная гипотеза $\mathbf{H_1}$). Такая ситуация может, например, возникнуть по данным об отказах по одному агрегату, но на разных авиатранспортных предприятиях, или сделанных на разных заводах; или проведенных в различных климатических условиях и т. п.

7.2. Критерии согласия

Статистическим критерием называется *правило,* позволяющее, основываясь только на выборке $x_1, x_2 \ldots x_n$, принять либо основную гипотезу **Ho,** либо альтернативную $\mathbf{H_1}$

Поскольку принятие основной **Ho** или альтернативной $\mathbf{H_1}$ гипотез основывается на выборке, состоящей из случайных чисел, $x_1, x_2 \ldots x_n$ то при любом критерии *ошибки* будут неизбежны.

При двухальтернативном выборе (либо **Ho**, либо $\mathbf{H_1}$) возможны четыре исхода:

принята гипотеза **Ho** и эта гипотеза верна;

принята гипотеза **Ho**, хотя она неверна;

принята гипотеза $\mathbf{H_1}$ и эта гипотеза верна;

принята гипотеза $\mathbf{H_1}$, хотя верна гипотеза **Ho**.

Из четырех выводов два являются ошибочными.

Ошибкой 1-го рода называют ошибку отклонения основной проверяемой гипотезы **Ho**, когда она верна. Вероятность этой ошибки обозначают через α

Ошибкой 2-го рода называют ошибку принятия гипотезы $\mathbf{H_1}$, когда верна основная гипотеза **Ho**. Вероятность ошибки 2-го рода обозначают через β.

Вопрос о том, какую из гипотез принять за основную, а какую за альтернативную, лежит вне области статистики и, в определенном смысле, этот выбор является произвольным. На практике обычно за основную гипотезу $\mathbf{H_0}$ принимают ту, когда важнее избежать ошибки 1-го рода.

При фиксированном объеме выборки обычно задаются величиной α - вероятностью отказа от основной гипотезы. Эту вероятность называют *уровнем значимости*

$$\alpha = P\{(x_1, x_2 ... x_n)\} \in W_k \qquad (7.1)$$

Выбор величины **α** зависит от сопоставления потерь, которые могут быть понесены в случае ошибочных заключений в ту или другую сторону. Чем весомее потери от ошибочного отвержения основной гипотезы **Ho** *f* тем меньше выбирается величина **α**. Чаще всего это затруднительно, поэтому пользуются стандартными уровнями значимости **α: 0,1; 0,05**; **0,025; 0,01; 0,005; 0,001**

Наиболее распространено значение **α = 0,05.**

Чем серьезнее последствия ошибки первого рода, тем меньшим должен быть уровень значимости, однако за понижения уровня значимости расплачиваются увеличением вероятности ошибки второго рода.

В связи с этим сначала назначают уровень значимости а, а затем выбирают такую процедуру проверки, которая обеспечивает минимальное значение **β**. Единственный способ одновременного уменьшения ошибок **α** и **β** - увеличение объема выборки **n**.

7.3. Применяемые критерии согласия

7.3.1. Критерий согласия Пирсона (критерий χ^2).

Критерий согласия Пирсона является одним из наиболее часто применяемых критериев для проверки согласованности теоретического распределения и статистических данных.

Пусть произведено **N** независимых опытов, в каждом из которых случайная величина **X** приняла некоторое значение **xi (i = 1,2,... n).** Результаты опытов сведены в **К** интервалов (групп) и построена гистограмма частот или гистограмма частостей.

По внешнему виду гистограммы можно высказать предположение о виде теоретического закона распределения, которому подчинены экспериментальные данные. Это предположение является основной гипотезой **Ho**.

Если Δn_i - . число значений величины в каждом интервале, то частота попадания в каждый интервал

$$P_i^* = \frac{\Delta n_i}{N} \qquad (7.2)$$

Зная (или предполагая) теоретический закон распределения, можно определить (по таблицам или расчетом) теоретические значения вероятностей попадания изучаемой случайной величины в каждый из интервалов. Если есть данные о теоретическом законе распределения, то теоретическая вероятность попадания случайной величины в **i-й** интервал равна

$$P_i = F(x_{i+1}) - F(x_i) \qquad (7.3)$$

Здесь **F(XI+I)** - значение функции распределения у правой границы интервала и **F(XI)** - соответственно у левой границы (в начале интервала), определенные по таблицам или расчетом..

Согласованность теоретического и статистического распределений определяется мерой расхождения в виде суммы квадратов отклонений значений P_i^* **и** P_i

$$U = \sum_{i=1}^{k} C_i (P_i^* - P_i)^2 \qquad (7.4)$$

В этой формуле **C_i** - некоторые весы интервалов. Пирсон предложил в качестве весов использовать отношение

$$C_i = \frac{N}{P_i} \qquad (7.5)$$

Пирсон показал, что при таком выборе величин **C_i** при больших значениях объема выборки **N** величина **U** практически не зависит от вида закона распределения **F(x)** и от величины **N**, а зависит только от числа интервалов **К**

Полученная таким образом мера распределения обычно обозначается χ^2

$$\chi^2 = N \sum \frac{(P_i^* - P_i)^2}{P_i} \qquad (7.6)$$

Учитывая формулу (7.2), получаем:

$$\chi^2 = \sum \frac{(\Delta n_i - NP_i)^2}{NP_i} \qquad (7.7)$$

Значение χ^2 зависит от числа степени свободы

r=k-s (7.8)

Здесь **k** - число интервалов, **s** - число независимых условий (связей), наложенных на распределение P_i^* (соответственно P_I). Эти условия следующие:

-обязательное условие $\sum_{i=1}^{k} P_i^* = 1$ или $(\sum_{i=1}^{k} P_i = 1)$

-условие совпадения средних значений

$\sum_{i=1}^{k} x_i^* \cdot P_i^* = m_x$ **,**т.е**.** теоретическое распределение подбирается таким образом,

чтобы совпали теоретическое и статистическое среднее значение;
- условие совпадения дисперсий (среднеквадратических отклонений)

$$\sum (x_i^* - m_x^*)^2 P_i^* = D_x$$

Так как первое условие всегда должно иметь место то**:**

r =k-1-*l* (7.9)

где ***l*** - число параметров, определяющих теоретическое распределение.

Например, для экспоненциального закона $\boldsymbol{l} = 1$, т.к. экспоненциальный закон определяется одним параметром распределения – **λ.**

Для нормального закона **l = 2**, т.к. параметрами нормального закона являются математическое ожиданий **m** и среднее квадратическое отклонение σ (или дисперсия **D**).

Для распределения Вейбулла **l= 2** , это **a** - параметр масштаба и **b** - параметр формы.

Таблицы значений χ^2 составляются с двумя входами: число степеней свободы **r** и доверительная вероятность (уровень значимости ошибки 1-го рода) **α**, или величина $\gamma = 1 - \alpha$.

Таблица значений χ^2 приведена в приложении 2.

7.3.2 Критерий Смирнова

Критерий Смирнова - непараметрический статистический критерий, применяемый для проверки гипотезы об однородности двух выборок.

Пусть $\mathbf{x_1, x_2, \ldots x_n}$ и $y_1, y_2, \ldots y_m$- взаимно независимые случайные величины. Проверяется гипотеза **Ho,** согласно которой обе выборки принадлежат одному и тому же распределению.

Для применения критерия Смирнова строятся статистические функции распределения $F_1^*(x)$ и $F_2^*(y)$ и находятся наибольшие значения из возможных отклонений этих функций

$$D_{n.m} = Sup[F_1^*(x) - F_2^*(y)] \qquad (7.10)$$

Гипотеза отвергается , когда $\mathbf{D_{m.n}}$

$$D_{n.m} \rangle \lambda_\alpha \sqrt{\frac{n+m}{m \cdot n}} \qquad (7.11)$$

Значения λ_α при заданных уровнях значимости приведены в таблице

α	**0.1**	**0.05**	**0.01**	**0.001**
λ_α	**1.22**	**1.36**	**1.63**	**1.95**

8. СТАТИСТИЧЕСКИЕ МЕТОДЫ ПРИЁМОЧНОГО КОНТРОЛЯ

8.1. Основные принципы статистических методов приёмочного контроля

Статистический приёмочный контроль – это совокупность методов статистического контроля качества массовой промышленной продукции с целью выявления её соответствия заданным требованиям.

При контроле качества промышленной продукции можно выделить два основных подхода: сплошной контроль (полный контроль всей партии) и выборочный контроль, когда по качеству выборки из **n** изделий, являющейся частью всего объема партии **N**, судят о всей партии.

При большом объеме партии сплошной контроль может быть затруднен как по времени, так и по стоимости самого процесса контроля. Но при сплошном контроле все равно возникает вопрос о том, сколько можно допускать "брака".

При выборочном контроле из партии извлекается *выборка* или *проба*. Выборка - это часть партии **(n < N)**. Она может представлять одно изделие или совокупность изделий, отобранных для контроля. Выборка из контролируемой партии должна быть случайной и представительной. Случайная выборка составляется из изделий, вероятность отбора каждого из которых одинакова. Выборка является представительной, если ее свойства отражают свойства всей контролируемой совокупности.

При контроле нештучной продукции (вода, бензин, газ и пр.) пользуются термином *проба*. Проба - это некоторое количество продукции, отобранное для контроля, она характеризуется объемом, взятым для пробы.

В стандартах на готовую продукцию, в технических условиях, технической документации и других нормативно - технических документах указываются планы контроля [11].

Планами контроля называется совокупность данных о виде контроля, объемах контролируемой партии, выборке или проб, о контрольных нормативных и решающих правилах.

Контрольный норматив – это значение показателей качества продукции, определенное нормативно – технической документацией и представляющее собой критерий для принятия решения по результатам контроля.

Приемочное число – это контрольный норматив, являющийся критерием для приемки партии продукции.

Браковочное число – контрольный норматив, являющийся критерием для забракования партии продукции.

Различают два варианта статистического контроля качества (надежности) изделий [63]:

- контроль по качественным признакам,
- контроль по количественным признакам.

Статистический приемочный контроль по качественному признаку представляет собой контроль качества продукции, в ходе которого все изделия разбиваются на две группы:

годные (кондиционные) и негодные (дефектные). Оценка всей партии проводится по величине доли дефектных изделий в выборке.

Статистический приемочный контроль по количественным признакам - это контроль качества продукции, в ходе которого определяются численные значения одного или нескольких ее параметров. Оценка всей партии проводится по статистическим характеристикам распределения определяемых параметров.

Статистический приемочный контроль может быть одноступенчатым, двухступенчатым, многоступенчатым и последовательным [11].

Одноступенчатый приемочный контроль - это такой контроль, когда решение относительно партии продукции принимается по результатам контроля только одной выборки или пробы.

При одноступенчатом контроле из контролируемой партии продукции объемом **N** случайным образом отбирают **n** единиц продукции, проверяют эту выборку и в ней подсчитывают число дефектных изделий **m**. Если число **m** меньше или равно приёмочному числу **C**, то партия изделий принимается. В противном случае она бракуется.

В том случае, если установлено приёмочное число C_2, то партия принимается при $m \leq C_1$ и бракуется при $m \geq C_2$.

Если же по результатам одноступенчатого контроля окажется, что $C_1 < m < C_2$, то производится двухступенчатый контроль.

При двухступенчатом контроле устанавливаются объем второй выборки n_2 и новое приемочное число C_3 и браковочное число C_4. Сравнение числа дефектных изделий с этими числами производится по совокупности двух выборок.

Последовательный контроль (последовательный анализ) - метод статистического исследования при проверке гипотез, при контроле после каждого наблюдения производится анализ всех предыдущих наблюдений. Максимальное число наблюдений (проверок нескольких выборок) заранее не устанавливается. Применяется как для приемки партии изделий, так и для сравнения двух систем.

8.2. Оперативная характеристика плана контроля

Оперативная характеристика плана статистического приемочного контроля – это выраженная уравнением, графиком или таблицей зависимость вероятности приема партии от величины, характеризующей уровень качества этой продукции.

Если **N** – общее число изделий в партии и **M** - число дефектных изделий в ней, то характеристикой качества партии служит доля дефектных изделий (уровень дефектности) в партии

$$q = \frac{M}{N} \qquad (8.1)$$

Суждение об этом качестве всей партии выносится по выборке, в которой уровень дефектности

$$q_b = \frac{m}{n} \qquad (8.2)$$

Обозначим через **P** вероятность приема партии по результатам выборочного контроля. Эта вероятность зависит от уровня дефектности всей партии **q** и от плана контроля (объема и числа выборок, приемочного и браковочного чисел). При фиксированном плане контроля может быть установлена зависимость **P** = **P(q),** которая называется оперативной характеристикой данного плана контроля. *Оперативная характеристика* в виде графика **P** = **P(q)** приведена на рис. 8.1.

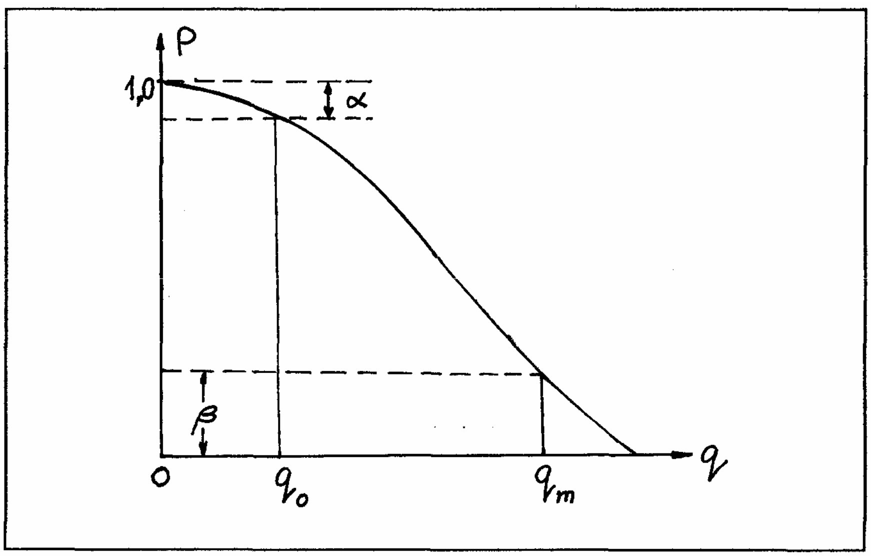


Рис. 8.1

Устанавливаются два уровня качества:

приемочный уровень качества, при котором $\mathbf{q = q_0}$;

браковочный уровень качества, соответствующий $\mathbf{q = q_m}$, причем $\mathbf{q_m > q_0}$

Если $\mathbf{q = 0}$ (бездефектная партия), то с вероятностью 1 партия принимается. Если $\mathbf{q = 1}$, (вся партия состоит из дефектных изделий), то вероятность приема партии равна нулю.

Поскольку заключение о годности партии производится на основе статистического материала, полученного в результате анализа случайной выборки, то, естественно, возможны ошибки.

Вероятность этих ошибок характеризуют *риском поставщика и риском заказчика.*

Риском поставщика α называется вероятность забракования партии изделий с приемлемым уровнем качества. Из графика на рис. 8.1. следует

$$\boldsymbol{\alpha = 1 - P(q_o)} \qquad (8.3)$$

Риском заказчика Р называется вероятность приемки партии изделий с браковочным уровнем качества

$$\boldsymbol{\beta = P(q_m)} \qquad (8.4)$$

В случае одноступенчатого контроля обязательно устанавливается число **С** – приемочное число. При выполнении условия $\mathbf{m \le C}$ партия изделий принимается, в противном случае – бракуется.

Таким образом, имеем следующий набор параметров: **n** – объем выборки, $\mathbf{q_0}$ – приемочный уровень качества, $\mathbf{q_m}$ – браковочный уровень качества, $\boldsymbol{\alpha}$ - риск поставщика, $\boldsymbol{\beta}$ - риск заказчика, **С** – приемочное число.

Взаимосвязь этих параметров определяется законом распределения. В общем случае статистические таблицы для решения задач контроля рассчитываются для гипергеометрического распределения. Практическое значение имеют биноминальное распределение и распределение Пуассона.

Основная совокупность задач, связанных с выборочным методом контроля состоит в следующем:

заданы $\mathbf{q_0}$, $\boldsymbol{\alpha}$ а и **Р** определить **n** и $\mathbf{q_m}$; заданы $\mathbf{q_0}$, $\mathbf{q_m}$, $\boldsymbol{\alpha}$ и $\boldsymbol{\beta}$ Р, определить **n** и **С**.

8.3. Статистический приемочный контроль с использованием биноминального распределения

Если объем выборки удовлетворяет условию

n < 0,1 N, (8.5)

то расчеты можно вести с помощью биноминального распределения.

Для случая биноминального распределения зависимость **P = P(q)** запишется в следующем виде:

$$P(q) = \sum_{k=0}^{c} C_n^k q^k (1-q)^{n-k} \tag{8.6}$$

Если **C = 0**, т.е. в выборке не должно быть ни одного браковочного изделия, то из формулы (8.6) непосредственно следует

$$\mathbf{P(q) = (1 - q)^n} \tag{8.7}$$

Риск поставщика будет определен, если положить $\mathbf{q = q_0}$, из (8.7) и (8 .3) получаем

$$\alpha = 1 - (1 - q_0)^n \tag{8.8}$$

или

$$l - \alpha = (l - q_0)^n \tag{8.9}$$

Соответственно для риска заказчика получаем, положив $\mathbf{q = q_m}$:

$$\beta = (l - q_m)^n \tag{8.10}$$

После логарифмирования (8.9)

$\mathbf{lg(l - \alpha) = nlg(l - q_0)}$ получаем

$$n = \frac{\lg(1-\alpha)}{\lg(1-q_0)} \tag{8.11}$$

После логарифмирования (8.10)

$\mathbf{lg\beta = nlg(l - q_m)}$ получаем

$$n = \frac{\lg\beta}{\lg(1-q_m)} \tag{8.12}$$

Формулы (8.9), (8.10), (8.11) и (8.12) устанавливают взаимосвязь параметров для случая биноминального закона в случае, когда **C = 0**.

Стандартной ситуацией является случай, когда α, β, $\mathbf{q_0}$ и **C = 0** и требуется определить **n** и $\mathbf{q_m}$.

8.4. Статистический приемочный контроль с использованием распределения Пуассона

Расчеты с использованием распределения Пуассона можно вести, если величины **q**, $\mathbf{q_0}$ и $\mathbf{q_m}$ удовлетворяют условию

$$q < 0{,}10 \tag{8.13}$$

Зависимость **P = P(q)** в этом случае записывается в следующем виде

$$P(q) = \sum_{k=0}^{c} \frac{1}{k!}(ng)^k \cdot l^{-nq} \tag{8.14}$$

При **C = 0** из этой формулы непосредственно следует

$$\mathbf{P(q) = e^{-nq}} \tag{8.15}$$

Для риска поставщика при **q** = **q₀** из 8.15 и 8.3 получаем

$$\alpha = 1 - l^{-nq_0} \tag{8.16}$$

Соответственно для риска заказчика при **q** = **q$_m$** из 8. 15 и 8.4 получаем

$$\beta = 1^{-nq_m} \tag{8.17}$$

Из выражений (8. 16) и (8. 17) после их логарифмирования можно получить

$$n = \frac{\ln(1-\alpha)}{q_0} \tag{8.18}$$

и

$$n = -\frac{\ln \beta}{q_m} \tag{8.19}$$

Соотношение браковочного уровня **q$_m$** к приемочному уровню **q$_0$** равно

$$E = \frac{q_m}{q_n} = \frac{\ln \beta}{\ln(1-\alpha)} \tag{8.20}$$

Формулы 8.16, 8.17, 8.18, 8.19 и 8.20 устанавливают взаимосвязь параметров для случая распределения Пуассона при **C=0**.

8.5. Метод последовательного анализа

Последовательный анализ – раздел математической статистики, характерной чертой которого является то, что число производимых наблюдений (момент остановки наблюдений) не фиксируется заранее, а выбирается по ходу наблюдений в зависимости от значений поступающих данных. Началу применения методов последовательного анализа в статистике положили работы А. Вальда [16]. Им было установлено, что в задаче различения (по результатам независимых наблюдений) двух простых гипотез последовательный критерий дает значительный выигрыш в среднем числе производимых наблюдений по сравнению с классическими способами различения с фиксированными объемами выборки и теми же вероятностями ошибочных решений.

Метод последовательного анализа применяется как для проверки партии продукции, так и для сравнения двух систем.

8.5.1. Проверка партии продукции

Пусть последовательно производится проверка каждого изделия

1, 2, 3,...m – 1, m,...,

причем, если i-е изделие дефектно, то принимается $\mathbf{x_i = 1}$, если изделие исправно (годно) – $\mathbf{x_i = 0}$.

В ходе проверки производится подсчет числа дефектных изделий нарастающим итогом.

$$d_m = \sum_{i=1}^{m} x_i \tag{8.21}$$

Эта величина сравнивается с приемочным a_m и браковочным r_m числами, которые также последовательно рассчитываются.

Значения же a_m и r_m зависят от ошибок 1-го и 2- го рода α и β и вероятностей $\mathbf{P_0}$ и $\mathbf{P_1}$, определяемых оперативной характеристикой **L(P)** (рис. 8.2).

Допустимый риск определяется из требования, чтобы вероятность забраковать партию не превышала заданной величины α, когда **P < P₀**, а вероятность принять партию не превышала β, когда P < **P₁**.

Принятие партии рассматривается как ошибка тогда и только тогда, когда **P < P₁**, а отказ принять партию, когда **P < P₀**. Если **P₀ < P < 1,** может быть принято любое решение.

Методика применения последовательного критерия состоит в следующем.

Задаются значения вероятностей **P₀** и **P₁** и величины ошибок первого и второго рода α и β.

Рассчитываются значения

$$q_1 = \frac{\beta}{1-\alpha}, \quad q_2 = \frac{1-\beta}{\alpha}, \quad \delta_0 = \frac{1-P_0}{1-P_1}, \quad \delta_1 = \frac{1-P_1}{1-P_0}, \quad \delta = \frac{P_1}{P_0} \tag{8.22}$$

Определяются значения приемочного a_m и браковочного r_m чисел

$$a_m = \frac{\ln q_1}{\ln\delta - \ln\delta_1} + m\frac{\ln\delta_0}{\ln\delta - \ln\delta_1} \tag{8.23}$$

$$ч_m = \frac{\ln q_2}{\ln - \ln\delta_1} + m\frac{\ln\delta_0}{\ln\delta - \ln\delta_1} \tag{8.24}$$

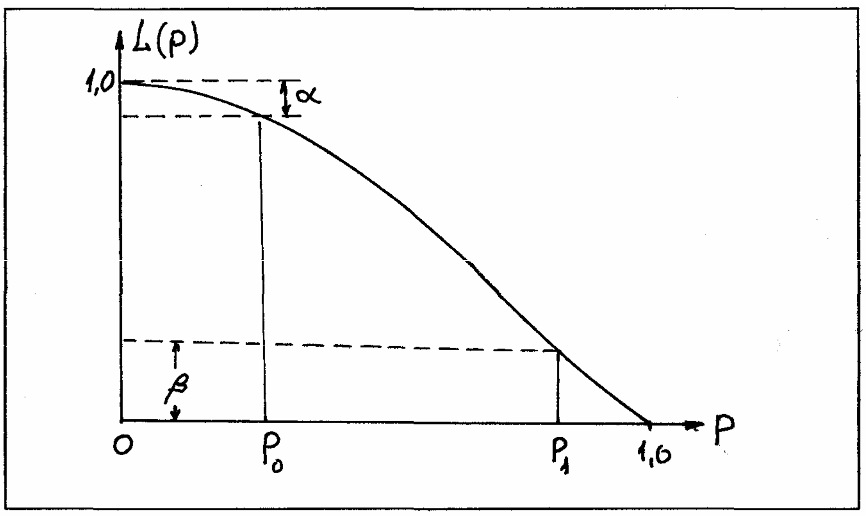


Рис. 8.2

По формуле (8.21) вычисляется число дефектных изделий $\mathbf{d_m}$ и проверяется условие $a_m < \mathbf{d_m} < r_m$.

Если $a_m < d_m < r_m$ – испытания продолжаются,

если $d_m \geq r_m$ – партия бракуется, (8.25)

если $d_m < a_m$ – партия принимается.

Практическая реализация изложенной методики может быть осуществлена табличным и графическим способом.

Табличный способ. Составляется таблица, в которую до начала испытаний заносятся величины **m = 1, 2,...;** $\mathbf{a_m}$ рассчитанные по формуле (8.23) и r_m, рассчитанные по формуле (8.24); по полученным в результате испытаний значениям x_i определяется величина

$$d_m = \sum_{i=1}^{m} x_i$$

Таблица 8.1

M	a_m	r_m		x_i	$d_m = \sum_{i=1}^{m} x_i$
1	a_{m1}	r_{m1}		0	0
2	a_{m2}	r_{m2}		1	1
3	a_{m3}	r_{m3}		1	2
.	.	.		.	.
.	.	.		.	.
.	.	.		.	.

В табл. 8.1 в качестве условного примера приведены значения $x_0 = 0$, $x_1 = 1$ и $x_e = 1$, это означает, что первое изделие исправно, а второе и третье – неисправны.

Решение о всей партии принимается по условиям (8.25).

Графический способ. Графический способ состоит в построении в координатах **m – d**$\mathbf{m - d_m}$ двух параллельных линий $\mathbf{L_0}$ и $\mathbf{L_1}$, являющихся границами зон решения (рис. 8.3).

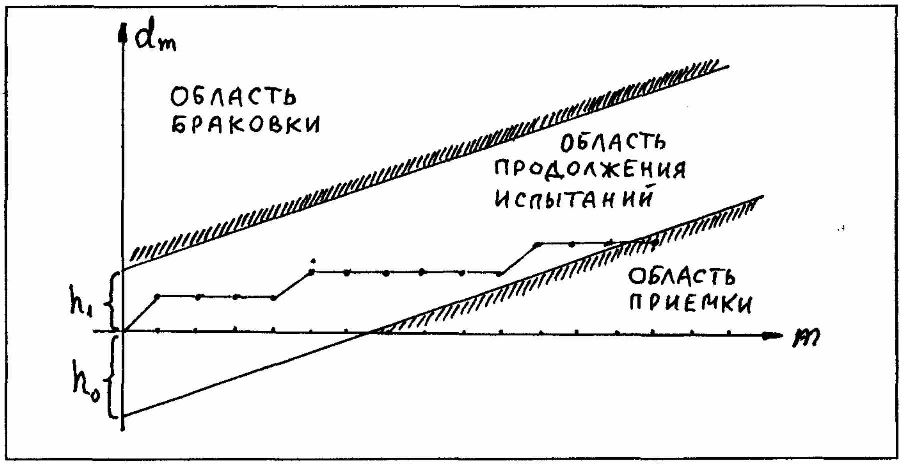


Рис. 8.3

Угловой коэффициент прямых $\mathbf{L_0}$ и $\mathbf{L_1}$ равен

$$S = \frac{\ln \dfrac{1-P_1}{1-P_0}}{\ln \dfrac{P_1}{P_0} - \ln \dfrac{1-P_1}{1-P_0}} \qquad (8.26)$$

Прямые $\mathbf{L_0}$ и $\mathbf{L_1}$ пересекают ось $\mathbf{d_m}$ в точках

$$h_0 = \frac{\ln \dfrac{\beta}{1-\alpha}}{\ln \dfrac{P_1}{P_0} - \ln \dfrac{1-P_1}{1-P_0}} \qquad (8.27)$$

$$h_1 = \frac{\ln \frac{1-\beta}{\alpha}}{\ln \frac{P_1}{P_0} - \ln \frac{1-P_1}{1-P_0}} \tag{8.28}$$

Рассчитанная для каждого значения **m** величина $\mathbf{d_m}$ откладывается на графике. Процесс прекращается, когда значение $\mathbf{d_m}$ попадает в область приемки или область браковки.

8.5.2. Сравнительная оценка эффективности двух систем

Метод последовательного анализа может быть применен для сравнительной оценки двух технических объектов (систем, устройств, агрегатов и т.п.), при сравнении двух методов, технологий, приемов и пр.

Для реализации метода проводятся две серии испытаний. Для системы **A:** $a_1, a_2, \ldots a_i \ldots$ и для системы **B:** $b_1, b_2, \ldots b_i$.,

где a_i и b_i означают результаты испытаний соответствующих систем.

Вероятность успеха системы **A – $\mathbf{P_1}$**, вероятность успеха системы **B – $\mathbf{P_2}$;** обе вероятности до испытаний неизвестны.

Имеем две гипотезы $\mathbf{P_1 \geq P_2}$ **и** $\mathbf{P_1 < P_2}$.

Успех испытания означает $a_i = 1$ (или $b_i = 1$), неуспех $a_i = 0$ (или $b_i = 0$).

Рассматриваются пары (a_i, b_i), такие, что $(a_i, b_i) =$ **(1,0) V (0,1)**.

Число пар **(1,0) – $\mathbf{t_1}$**, число пар **(0,1) – $\mathbf{t_2}$**.

Общее число наблюдений

$t = t_1 + t_2$.

Относительное превосходство системы (процесса) **B** над системой (процессом) **A** измеряется отношением их показателей эффективности $\mathbf{K_1}$ **и** $\mathbf{K_2}$.

$$u = \frac{K_2}{K_1} = \frac{\frac{P_2}{1-P_2}}{\frac{P_1}{1-P_1}} = \frac{P_2(1-P_1)}{P_1(1-P_2)} \tag{8.29}$$

Если $\mathbf{P_1 = P_2}$, то **U = 1**, т.е. системы равноценны. Если **U > 1**, то система **B** лучше, если **U < 1**, то лучше система **A.**

Для принятия решения о сравнительной оценке двух систем выбираются два значения $\mathbf{U_0}$ **и** $\mathbf{U_1}$ $\mathbf{(U_0 < U_1)}$, таких, что отказ от системы **A** в пользу системы **B** рассматривается как ошибка, имеющая практическое значение, когда истинное значение $\mathbf{U \leq U_0}$. Выбор системы **A** рассматривается как ошибка, когда $\mathbf{U \geq U_1}$.

Если $\mathbf{U_0 < U < U_1}$, то не важно, какое решение принято.

Величина допустимого риска определяется следующим образом. Вероятность отказа от системы **A** не должна превышать величины α, когда $\mathbf{U \leq U_0}$, а вероятность принять систему **A** не должна превышать величины β, когда $\mathbf{U \geq U_1}$.

Методика реализации последовательного анализа для выбора двух систем из сравнения их эффективности состоит в следующем.

1. Задаются значения величин $\mathbf{U_0}$, $\mathbf{U_1}$, α и β.

2. Рассчитываются значения

$$q_1 = \frac{\beta}{1-\alpha}, \quad q_2 = \frac{1-\beta}{\alpha}, \quad \delta = \frac{1+U_1}{1-U_0} \qquad (8.30)$$

3. Определяются значения приемочного a_t и r_t чисел.

$$a_t = \frac{\ln q_1}{\ln U_1 - \ln U_0} + t \cdot \frac{\ln \delta}{\ln U_1 - \ln U_0} \qquad (8.31)$$

$$ч_t = \frac{\ln q_2}{\ln U_1 - \ln U_2} + t \cdot \frac{\ln \delta}{\ln U_1 - \ln U_2}, \qquad (8.32)$$

где $t = t_1 + t_2$.

Если в ходе испытаний систем значения пар (a_i, b_i) = **(0,0),** или (a_i, b_i) **= (1,1),** т.е. обе системы показали одновременно неуспех или успех, то эти испытания не учитываются, т.к. они нечего не дают для анализа.

4. Условия принятия решения.

Если a_t < $\mathbf{t_2}$ < $\mathbf{r_t}$ – испытания продолжаются, если $\mathbf{t_2} \le a_t$ – принимается система **A**, если $\mathbf{t_2} \ge r_t$ – принимается система **B**. (8.33)

Практическая реализация изложенной методики, как и в случае приема партии изделий, может быть осуществлена табличным или графическим способом.

Табличный способ. Составляется таблица, в которую до начала испытаний заносятся величины **t = 1,2,...,** a_t и r_t, а в процессе испытаний – результат наблюдений (a_i, b_i) и подсчитывается $\mathbf{t_2}$, т.е. число пар, имеющих значение (0,1).

Таблица 5.2

t	a_t	r_t	a_i	b_i	(a_i, b_i)	$\mathbf{t_2}$
1	a_1	r_1	**0**	**0**	**(0,0)**	**-**
2	a_2	r_2	**1**	**0**	**(1,0)**	**0**
3	a_3	r_3	**0**	**1**	**(0,1)**	**1**
4	a_4	r_4	**1**	**1**	**(1,1)**	**1**
5	a_5	r_5	**0**	**1**	**(0,1)**	**2**
.	.	.	**.**	**.**	**.**	**.**
.	.	.	**.**	**.**	**.**	**.**
.	.	.	**.**	**.**	**.**	**.**

В табл. 5.2 в качестве условного примера приведены значения при t = 1 $(a_1, b_1) = (0,0)$, этот результат испытаний не учитывается; при t = 2 $(a_2, b_2) = (1,0)$, значение $t_2 = 0$; при t = 3 $(a_3, b_3) = (0,1)$, значение $t_2 = 1$; при t = 4 $(a_n, b_n) = (1,1)$, этот результат испытаний не учитывается, по-прежнему $t_2 = 1$; при t = 5 $(a_5, b_5) = (0,1)$, $t_2 = 2$..., т.е. в графе t_2 нарастающим итогом записываются результаты испытаний, дающие значения $(a_i, b_i) = (0,1)$.

Решение принимается по условиям 8.33.

Графический способ. Графический способ состоит в построении в координатах **t – $\mathbf{t_2}$** двух параллельных линий $\mathbf{L_0}$ и $\mathbf{L_1}$, являющихся границами зон решения (рис. 8.4)

Угловой коэффициент прямых $\mathbf{L_0}$ и $\mathbf{L_1}$ равен

$$S = \frac{\ln \delta}{\ln U_1 - \ln U_0} \qquad (8.34)$$

Прямые $\mathbf{L_0}$ и $\mathbf{L_1}$ пересекают ось $\mathbf{t_2}$ в точках

$$h_0 = \frac{\ln q_1}{\ln U_1 - \ln U_0} \tag{8.35}$$

$$h_1 = \frac{\ln q_2}{\ln U_1 - \ln U_0} \tag{8.36}$$

Рассчитанная для каждого значения **t** величина $\mathbf{t_2}$ откладывается на графике. Процесс прекращается, когда значение $\mathbf{t_2}$ попадает в область приемки системы **A** или системы **B**.

В том случае, когда при достаточно большом числе испытаний значения $\mathbf{t_2}$ не выходят из зоны продолжения испытаний, то это означает, что системы **A** и **B** равноценны и решение принимается или любое, или с привлечением других значимых для принятия решения параметров.

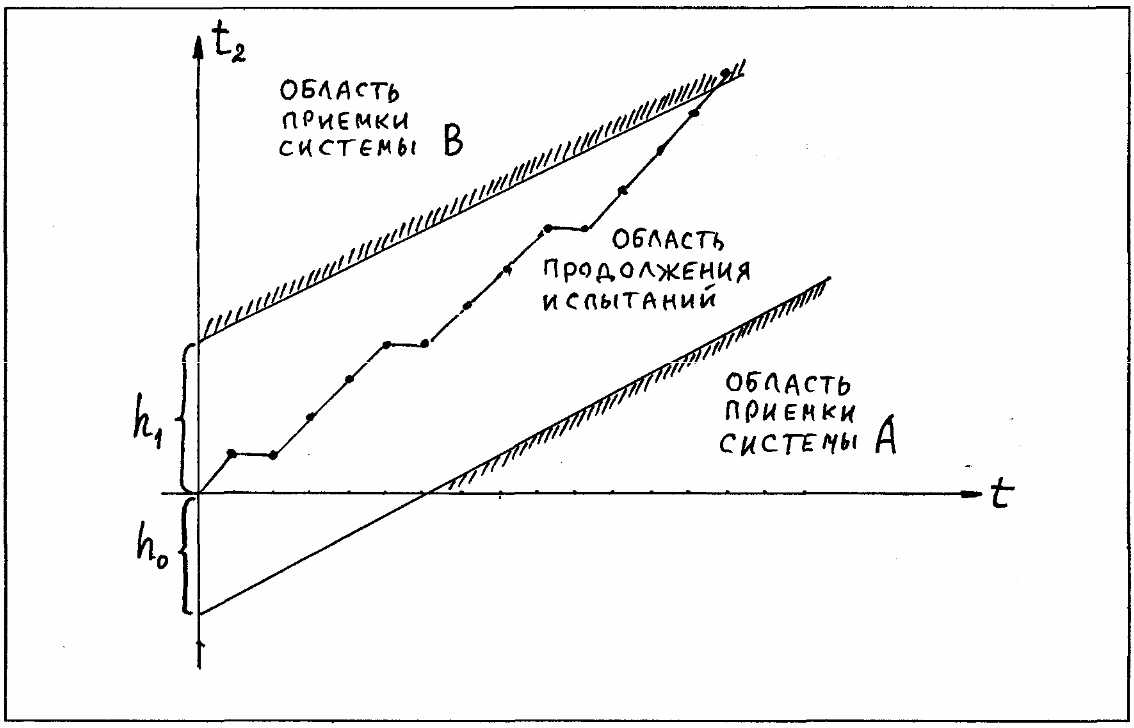


Рис. 8.4

9. МЕТОДЫ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ ДЛЯ АНАЛИЗА СИСТЕМ

9.1. Математическое программирование как инструмент решения задач выбора

Математическое программирование является инструментом решения разнообразных задач выбора. Математическое программирование в отличие от классических методов решения экстремальных задач основное внимание уделяет вопросам ограничений на область изменения переменных.

Само наименование «математическое программирование» связано с тем, что целью решения задач является выбор *программы* действий.

Формулировка задачи математического программирования состоит в следующем:

Вводится скалярная функция векторного аргумента **Ф(x)**. Эту функцию в зависимости от конкретного вида решаемой задачи называют *целевой функцией, функцией цели, критерием эффективности* или *показателем качества.*

Критерий эффективности **W** можно представить в виде целевой функции

$$\mathbf{W=\Phi(x)} \qquad (9.1)$$

На множестве

$$\mathbf{X = \{x: q_i(x) \geq 0,\ i = 1, 2, \ldots, k_j \quad h_j(x)=0,\ j = 1, 2, \ldots, m\}.} \qquad (9.2)$$

Здесь $\mathbf{q_i(x)}$ и $\mathbf{h_j(x)}$ – скалярные функции, определяющие ограничения на элементах **x** множества **X**.

Решение задачи состоит в отыскании такой точки **x*,** которая обеспечивает оптимальное (минимальное или максимальное) значение целевой функции **Ф(x),** при заданных функциями **q(x)** и **h(x)** ограничениях.

Если целевая функция **Ф(x)** и функция ограничения **q(x)** и **h(x)** линейны, то соответствующая задача является *задачей линейного программирования.*

Если хотя бы одна из упомянутых функций является нелинейной, то такая задача относится к *задаче нелинейного программирования.*

Задачи, решения которых достигаются в результате многих этапов, относятся к задачам динамического программирования.

В задачах целочисленного программирования искомые неизвестные могут принимать только целочисленные значения.

Если в целевой функции **Ф(x)** и функциях **q(x)** и **h(x),** определяющих ограничения, содержатся параметры с элементами неопределённости, то такая задача относится к задаче *стохастического программирования.*

9.2. Линейное программирование

9.2.1. Постановка задачи линейного программирования

Как следует из определения, целевая функция имеет следующий вид:

$$\mathbf{W = C_1x_1 + C_2x_2 + \ldots + C_jx_j + \ldots + C_nx_n} \qquad (9.3)$$

В зависимости от существа решаемой задачи ищется максимум или минимум этой функции.

Исходные ограничения также являются линейными функциями и могут задаваться в виде равенств или неравенств.

Пусть из общего числа **m** ограничений I заданы в виде равенств, а остальные **m - 1** в виде неравенств.

Получаем следующую систему уравнений:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + ... + a_{1j}x_j + ... + a_{1n}x_n = b_1 \\ .... \\ a_{i1}x_1 + a_{i2}x_2 + ... + a_{ij}x_j + ... + a_{in}x_n = b_i \\ .... \\ a_{l1}x_1 + a_{l2}x_2 + ... + a_{lj}x_j + ... + a_{\ln}x_n = b_l \\ a_{l+1,1}x_1 + a_{l+1,2}x_2 + ... + a_{l+1,j}x_j + ... + a_{l+l,n}x_n \le b_{l+1} \\ .... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mj}x_j + ... + a_{mn}x_n \le b_m \end{array}\right\} \qquad (9.4)$$

Чтобы закончить математическую формулировку задачи, необходимо записать, что **m < n** и что все неизвестные должны быть неотрицательными числами, т.е.

$$x_j > 0, \quad j = l, 2, ... n \qquad (9.5)$$

В системе (9.4) ограничений записаны только односторонние неравенства. В общем случае возможны такие случаи, когда в одной задаче имеются неравенства, направленные в разные стороны, т.е. неравенства противоположного смысла. В системе (9.4) все значения b_i положительны. Если какое – либо значение b_i будет отрицательным, то направление неравенства будет противоположным. Чтобы привести это уравнение к тому же направлению знака неравенства, достаточно все члены этого уравнения умножить на **-1**.

Совокупность значений $x_1 \geq 0$, $x_2 \geq 0, ... x_n \geq 0$, удовлетворяющих системе ограничений (9.4), образует многомерную область допустимых решений (ОДР). Значение же **max (min) W** $(x_1, x_2, ... x_n)$ может быть достигнуто (если оно существует) не при всех точках **ОДР**, а только при определенных, обеспечивающих условие **max (min)** целевой функции.

В случае, когда число переменных **n** равно числу ограничивающих уравнений **m**, то система (9.4) имеет единственное решение, которое может быть найдено с помощью матричной алгебры.

Если же **m < n**, т.е. когда число независимых уравнений, определяющих ограничения, меньше числа переменных, то существует бесчисленное множество решений, которое лежит в области допустимых решений (ОДР). При этом **n - m** переменным можно придавать производные значения *(свободные переменные),* а остальные **m** переменных выразятся через эти свободные переменные. Эти **m** переменных называют *базисными.*

9.2.2. Геометрическая интерпретация линейного программирования.

Геометрическую интерпретацию задачи линейного программирования можно наглядно представить, когда **n – m = 2**, т.е. когда две из общего числа **n** переменных, например, x_1 и x_2 выбраны в качестве свободных. Тогда остальные **m** переменных можно выразить через x_1 и x_2. В общем виде это запишется следующим образом:

$$\left.\begin{array}{l} x_3 = a_{31} \cdot x_1 + a_{32} \cdot x_2 + \beta_3 \\ x_4 = a_{41} \cdot x_1 + a_{42} \cdot x_2 + \beta_4 \\ .... \\ x_n = a_{n1} \cdot x_1 + a_{n2} \cdot x_2 + \beta_n \end{array}\right\} \qquad (9.6)$$

Условие неотрицательности всех n переменных запишется в этом случае в виде следующих неравенств

$$\left.\begin{aligned} x_1 \ge 0 \\ x_2 \ge 0 \end{aligned}\right\} \qquad (9.7)$$

$$\left.\begin{aligned} a_{31}x_1 + a_{32}x_2 + \beta_3 \ge 0 \\ a_{41}x_1 + a_{42}x_2 + \beta_4 \ge 0 \\ \dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \beta_n \ge 0 \end{aligned}\right\} \qquad (9.8)$$

Поскольку все условия выражены через две переменные $\mathbf{x_1}$ и $\mathbf{x_2}$, то мы получили двумерную задачу, которую можно наглядно изобразить на плоскости.

Свободные переменные $\mathbf{x_1}$ и $\mathbf{x_2}$ будем откладывать по осям $\mathbf{0x_1}$ и $\mathbf{0x_2}$. (рис. 9.1).

Условия (9.7) означают, что допустимые переменные лежат правее оси $\mathbf{0x_2}$ и выше оси $\mathbf{0x_1}$.

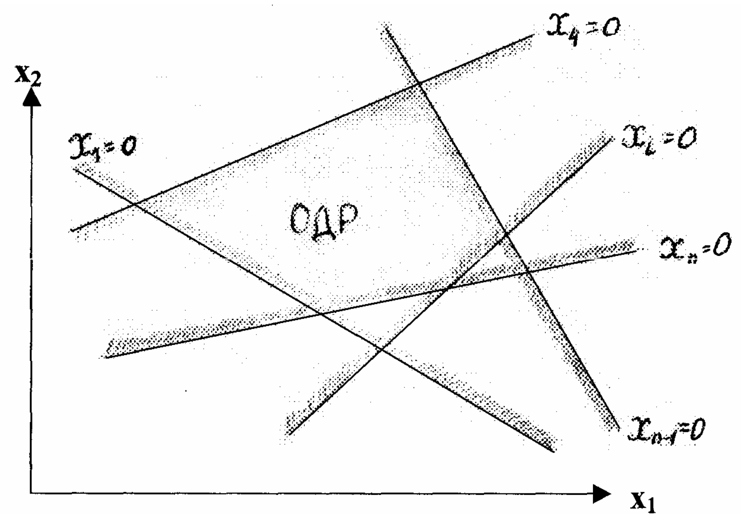


Рис. 9.1

Если положить величину $\mathbf{x_3}$ равной нулю, то получим уравнение

$$\alpha_{31} \cdot x_1 + \alpha_{32} \cdot x_2 + \beta_3 = 0,$$

которое является уравнением прямой линии. Штриховкой отметим ту сторону от прямой, где $\mathbf{x_3} > \mathbf{0}$.

Аналогичную процедуру можно провести и с оставшимися переменными, получив, кроме осей координат $\mathbf{n - 2}$ прямых, с обозначениями сторон, где каждая переменная $\mathbf{x_i} > \mathbf{0}$.

Многоугольник, внутри которого расположена область, принадлежащая одновременно всем полуплоскостям, удовлетворяющим условиям $\mathbf{x_i} > \mathbf{0}$, есть область допустимых решений (ОДР). На рис 9.1 ОДР помечена затемнениями. Заметим, что область допустимых решений представляет собой выпуклый многоугольник.

Рассмотрим далее геометрическую интерпретацию поведения целевой функции $\mathbf{W = C_1x_1 + C_2x_2 + \ldots + C_jX_j + \ldots + C_nx_n}$

Для тех же условий, т.е. когда n - m =2n в качестве свободных переменных выбраны $\mathbf{x_1}$ и $\mathbf{x_2}$. Так же как и раньше можно базисные переменные $x_3, x_4, \ldots, x_n$ выразить через $\mathbf{x_1}$ и $\mathbf{x_2}$, и для целевой функции получим выражение

$$\mathbf{W = \gamma_0 + \gamma_1 x_1 + \gamma_2 x_2} \qquad (9.9)$$

Здесь γ_0 - свободный член, который появился при переходе к переменным $\mathbf{x_1}$ и $\mathbf{x_2}$.

Линейная функция (9.9) имеет угловой коэффициент $-\frac{\gamma_1}{\gamma_2}$ и она достигает **max(min)** при тех значениях $\mathbf{x_1}$ и $\mathbf{x_2}$, что и функция

$$\mathbf{W^1} = \gamma_1 x_1 + \gamma_2 x_2\,, \tag{9.10}$$

Придавая величине $\mathbf{W^1}$ различные значения $\mathbf{C_1}$, $\mathbf{C_2}$..., получим на плоскости $\mathbf{x_1 0 x_2}$ различные прямые, параллельные между собой (рис. 9.2). Направление, в котором при параллельном перемещении прямой $\mathbf{W^1}$ величина $\mathbf{W^1}$ будет увеличиваться (уменьшаться), зависит от величины и знака величины γ_1 и γ_2.

На рис 9.2 показано направление, при котором величина $\mathbf{W^1}$ уменьшается, т.е. для показанных положений прямых $\mathbf{C_2} < \mathbf{C_1}$. При этом γ_1 и γ_2 будут отрицательными, а следовательно, угловой коэффициент, равный $-\dfrac{\gamma_1}{\gamma_2}$, будет тоже отрицательным, что и соответствует положению прямой $\mathbf{W^1}$.

Рассмотрим теперь взаимное расположение области допустимых решений и линейной функции $\mathbf{W^1}$, отражающей целевую функцию (рис. 9.3).

Если перемещать прямую $\mathbf{W^1}$ параллельно самой себе в направлении увеличения (уменьшения) целевой функции, то, в зависимости от конфигурации ОДР, могут быть следующие случаи.

1. Решение задачи линейного программирования единственно и оно соответствует одной из вершин многоугольника ОДР (рис. 9.3а)

2. Решение соответствует любой точке отрезка АВ, которая является одной из сторон многоугольника ОДР и она параллельна линейной целевой функции $\mathbf{W^1}$ (рис. 9.3б).

3. ОДР незамкнута, именно неограниченна в направлении перемещения прямой $\mathbf{W^1}$. В этом случае задача не имеет решения (рис. 9.3в).

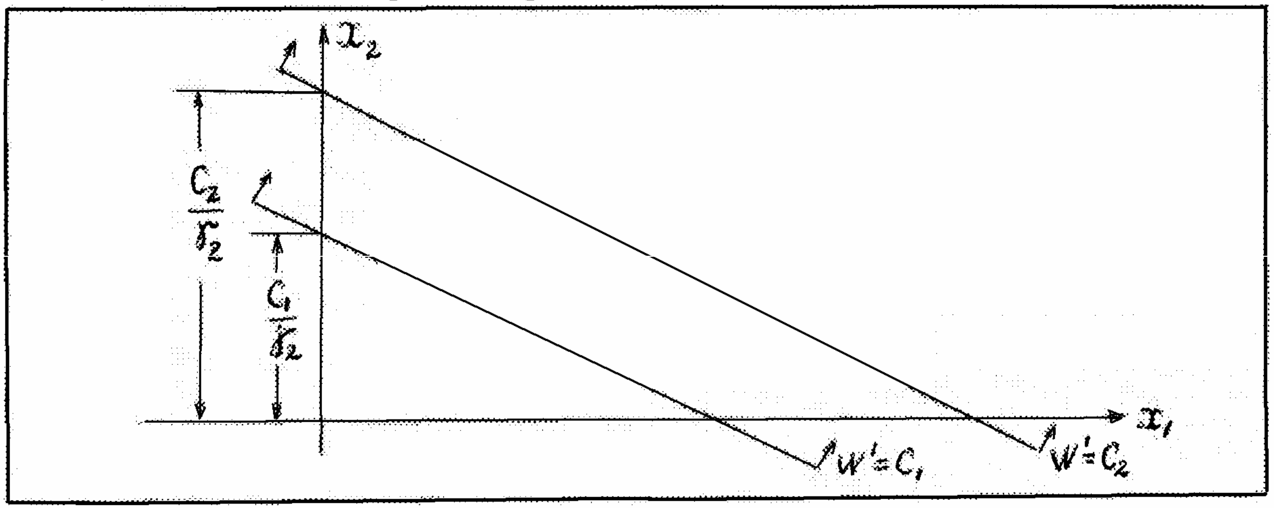


Рис. 9.2

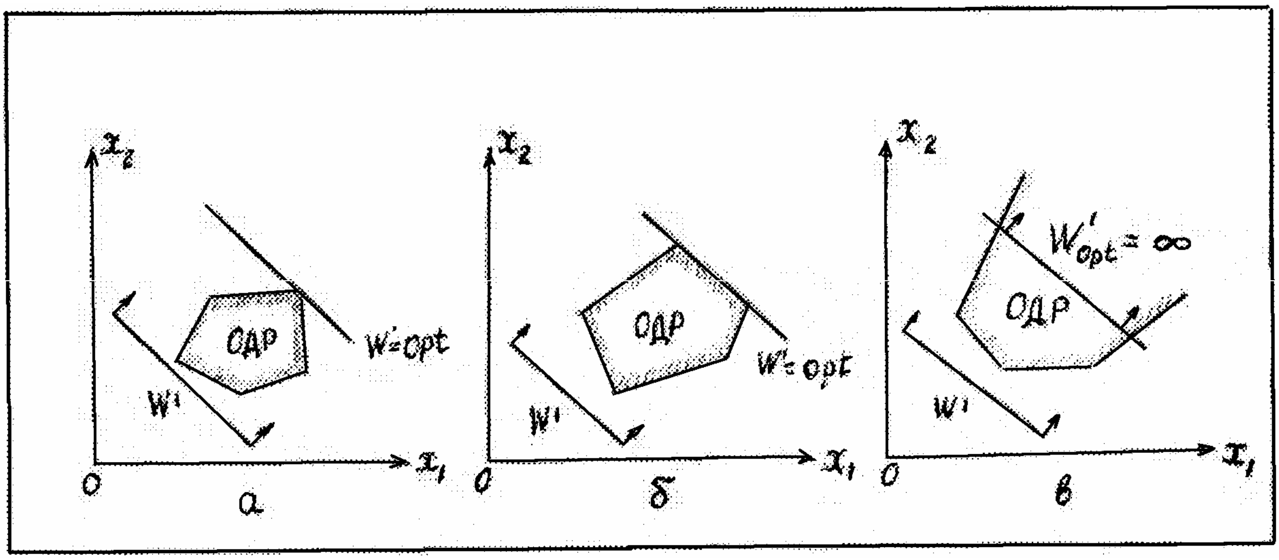


Рис. 9.3

9.3. Нелинейное и динамическое программирование

9.3.1. Задачи нелинейного программирования

Как уже упомянуто в п. 9.1 настоящей главы, к задачам нелинейного программирования относятся те задачи, в которых либо целевая функция, либо хотя бы одна из функций, определяющая ограничения, является нелинейной.

Если в задачах линейного программирования точка экстремума находится в вершине многогранника области допустимых решений (или, в некоторых случаях, совпадает с плоскостью многогранника), то в задачах нелинейного программирования она может лежать в вершине многогранника, на его ребре или грани и даже внутри ОДР. В том случае, когда нелинейным являются ограничения, то кроме основного, глобального оптимума могут существовать точки локальных оптимумов.

Не вдаваясь в общие теоретические аспекты, рассмотрим пример решения задачи нелинейного программирования.

Найти максимальное значение функции

$$\mathbf{W = x_2 - x_1^2 + 6x_1} \qquad (9.13)$$

При следующих ограничениях на переменные $\mathbf{x_1}$ и $\mathbf{x_2}$:

$$\left.\begin{aligned} x_1+2x_2 &\le 12 \\ 3x_1+2x_2 &\le 24 \end{aligned}\right\} \qquad (9.14)$$

$$\left.\begin{aligned} x_1, x_2 &\ge 0 \\ x_2 &\le 4 \end{aligned}\right\} \qquad (9.15)$$

Строим линии уровней $\mathbf{W = x_2 - x_1 + 6x_1 = h}$, где $\mathbf{h}$ - возможные значения функции $\mathbf{W(x_1,x_2)}$.

Для каждого фиксированного значения $\mathbf{h}$ эта функция имеет вид параболы, которая тем выше отдалена от оси $\mathbf{0x_1}$, чем больше значение $\mathbf{h}$. (рис. 9.4).

Областью допустимых решений является многогранник OABCD. Для нахождения решения задачи необходимо определить ту точку этого многоугольника, где функция (9.13) принимает максимальное значение.

Функция **W(x₁,x₂)** принимает экстремальное значение в точке касания одной из парабол с границей многоугольника OABCD. В нашем случае эта точка **E**. Координаты точки **E** можно найти из системы уравнений

$$\left.\begin{aligned} x_2 - x_1^2 + 6x_1 &= 13 \\ x_2 &= 4 \end{aligned}\right\} \qquad (9.16)$$

Решая эту систему, получаем $\mathbf{x_1^* = 3}$, $\mathbf{x_2^* = 4}$ **и** $\mathbf{W_{max} = 13}$ при этих значениях переменных.

Как видим, в задаче этого примера решение не является вершиной многоугольника ОДР. Поэтому метод перебора вершин, который используется при решении задач линейного программирования, в данном случае неприемлем.

Рис. 9.4

9.3.2. Динамическое программирование

В задачах линейного и нелинейного программирования, рассмотренных выше, процесс считался статическим, т.е. не зависящим от времени. Решение таких задач находилось в один этап или за один шаг. Такие задачи называют *одноэтапными* или *одношаговыми*.

Динамическое программирование (динамическое планирование) представляет собой математический метод оптимизации решений, специально приспособленный к *многошаговым* (многоэтапным) операциям (процессам, характеризующим систему).

Процесс, исследуемый методом динамического программирования, может явно зависеть от времени и естественно распадаться на ряд шагов или этапов. Например, процесс планирования хозяйственной деятельности системы предприятий, для которого естественным шагом является хозяйственный год. Примером такой деятельности могут быть разнообразные задачи распределения ресурсов.

В других случаях идущий во времени процесс является непрерывным (например, движение самолетов или ракеты). В этом случае этот процесс условно разбивают на этапы, каждый из которых занимает какой-то отрезок $\mathbf{\Delta t}$.

Наконец, метод динамического программирования может быть применен для решения задач системного анализа, в частности, формирования структуры системы. В этом случае разбиение на шаги производится не по времени, а по любому другому признаку, например, по элементам системы.

При решении задач динамического программирования используются показатели эффективности как аддитивные, так и мультипликативные.

В основе метода лежит принцип оптимальности, исходя из которого, планируя многошаговую операцию, выбирают управление на каждом шаге с учётом влияния принятого решения на функцию цели в последующих шагах.

Поэтому управление в ближайшем шаге выбирается так, чтобы оно, вместе с оптимальным управлением, на всех дальнейших шагах облегчило бы максимальный выигрыш на последующих шагах, включая данный.

Рассмотрим более общий подход к проблеме управления системой методом динамического программирования.

Пусть исходное состояние системы есть $\mathbf{S_0}$, а конечное состояние системы, в которое необходимо перевести систему, есть $\mathbf{S_k}$.

Предполагается, что система управляема, способ воздействия на систему - управление $\mathbf{U}$, должен быть таким, чтобы перевод из состояния $\mathbf{S_0}$ в состояние $\mathbf{S_k}$ был бы оптимальным (например, максимальным должен быть выигрыш). Поскольку выигрыш зависит от управления, то можно записать

$$W_{\max} = \max_{U} \{W(U)\} \qquad (9.17)$$

Запись эта означает "максимальное" из всех значений **W(U)** при всех возможных управлениях **U**.

Для любого промежуточного **i** - го шага принцип оптимальности (принцип Беллмана) может быть записан в виде следующего функционального рекуррентного соотношения

$$W_i(S) = \max_{U} \{W_i(S_1 U_i) + W_{i+1}(\varPhi_i(S_1 U_i))\} \qquad (9.18)$$

$\mathbf{W_i(S)}$ = - условный оптимальный выигрыш, получаемый на всех последующих шагах, начиная с **i** - го и до конца. Слово "условный" означает выигрыш при условии, что на **i**-м шаге система находится в состоянии **S**;

$\mathbf{W_i(S_1U_i)}$ - выигрыш на **i**-м шаге, который зависит от состояния **S** и принятого управления $\mathbf{U_i}$;

$\mathbf{\Phi_i(S_1U_i) = S}$ - состояние системы перед следующим **i + 1** шагом, которое естественно зависит от предыдущего состояния **S** и управления, принятого на **i**-м шаге.

$\mathbf{W_{i+1}\{\Phi_i(S_1U_i)\}}$ - условный оптимальный выигрыш, получаемый на всех шагах, начиная с **(i + 1)** - го шага. Отметим существующую особенность, заложенную в функциональном соотношении (9.18): функция $\mathbf{W_i(S)}$ **i** - го шага зависит от функции $\mathbf{W_{i+1}(S^1)}$ последующего шага, т.е. функцию $\mathbf{W_i(S)}$ можно определить, если известна следующая за ней по порядку функция $\mathbf{W_i(S)}$.

Эта особенность подсказывает следующий способ отыскания условных оптимальных выигрышей: сначала ищется условный оптимальный выигрыш для последнего шага, т.е. для того шага, который приводит к конечному желаемому состоянию $\mathbf{S_k}$, поскольку за этим последним шагом нет никакого другого.

Одно за другим строится вся цепочка условных оптимальных управлений. Действительно, зная $\mathbf{W_K(S)}$, можно по общей формуле (9.18), полагая в ней **i + 1 = m**, найти функцию $\mathbf{W_{m-1}(S)}$ и соответствующее условное оптимальное управление $\mathbf{U_{m-1}(S)}$; затем $\mathbf{W_{m-2}(S)}$ и $\mathbf{U_{m-2}(S)}$ и так далее вплоть до последнего от конца, т.е. до первого шага, для которого будут найдены функция $\mathbf{W_1(S)}$ и управление $\mathbf{U_1(S)}$.

На этом заканчивается первый этап оптимизации: найдены - условный оптимальный выигрыш и условное оптимальное управление для каждого шага.

Второй этап состоит в нахождении безусловного (окончательного) оптимального управления.

Исходное состояние $\mathbf{S_0}$ задано. Условный оптимальный выигрыш для первого шага был определен. Следовательно, для первого шага

$$\mathbf{W_{max} = W_1(S_0)} \qquad (9.19)$$

Одновременно находится и оптимальное управление на первом шаге

$$\mathbf{U_1 – U_1(S_0)}$$

По исходному состоянию $\mathbf{S_0}$ и управлению U_1 находят состояние после первого шага

$$\mathbf{S_1^1 = \Phi_1(S_0,U_1)} \qquad (9.20)$$

По состоянию $\mathbf{S_1^1}$ находим оптимальное управление на втором шаге $\mathbf{U_2=U_2(S_1^1)}$, затем $\mathbf{S_2^1 = \Phi_2(S_1^1,U_2)}$ и т.д. Образуется цепочка

$$\mathbf{S_0 \to U_1(S_0) \to S_1^1 \to U_2(S_1^1) \to \ldots \to S_{k-1}^1 \to U_k(S_{k-1}^1) \to S_k^1} \qquad (9.21)$$

10. АНАЛИЗ СИСТЕМ ПО СХЕМЕ МАРКОВСКИХ СЛУЧАЙНЫХ ПРОЦЕССОВ

10.1. Марковские случайные процессы

Среди множества видов случайных процессов большое теоретическое и практическое значение имеют процессы, обладающие следующим свойством:

Для каждого момента времени t_i вероятность любого состояния системы в будущем (при $\mathbf{t > t_i}$) зависит только от ее состояния в настоящем (при $\mathbf{t = t_i}$) и не зависит от того, когда и каким образом система пришла в это состояние (т.е. как развивался процесс в прошлом).

Впервые это свойство было сформулировано русским математиком Марковым А.А. в 1906 году, и впоследствии это свойство стало принято называть марковским свойством, а случайные процессы, обладающие этим свойством, стали называть Марковскими процессами.

Имеется немало определений марковского процесса, приведем одно из этих определений.

Случайный процесс $\mathbf{X(t)}$ называется марковским, если для любых $\mathbf{n}$ моментов времени $\mathbf{t_1 < t_2 < ... < t_n}$ из отрезка [О,Т] условная функция распределения «последнего» значения $\mathbf{X(t_n)}$ при фиксированных значениях $\mathbf{X(t_1), X(t_2), ..., X(t_{n-1})}$ зависит только от $\mathbf{X(t_{n-1})}$, т.е. при заданных значениях $\mathbf{x_1, x_2, \ldots, x_n}$ справедливо соотношение

$$\mathbf{P\{X(t_n) \le x_n/X(t_1) = x_1, X(t_2) = x_2, \ldots, X(t_{n-1}) = x_{n-1}\} = P\{X(t_n) \le x_n/X(t_{n-1}) = x_{n-1}\}} \quad (10.1)$$

Приведенное определение заимствовано из работы [62].

Марковский процесс называется процессом без последствия, т.е. будущее развитие марковского случайного процесса зависит только от настоящего состояния и не зависит от «предыстории» процесса, другими словами, для Марковского процесса при известном настоящем будущее не зависит от прошлого.

Применительно к случайным марковским процессам, различают марковские цепи, марковские последовательности, марковские процессы с конечным и бесконечным числом состояний, а также смешанные марковские процессы. Характер реализаций четырех основных видов марковских процессов приведен на рис. 10.1.

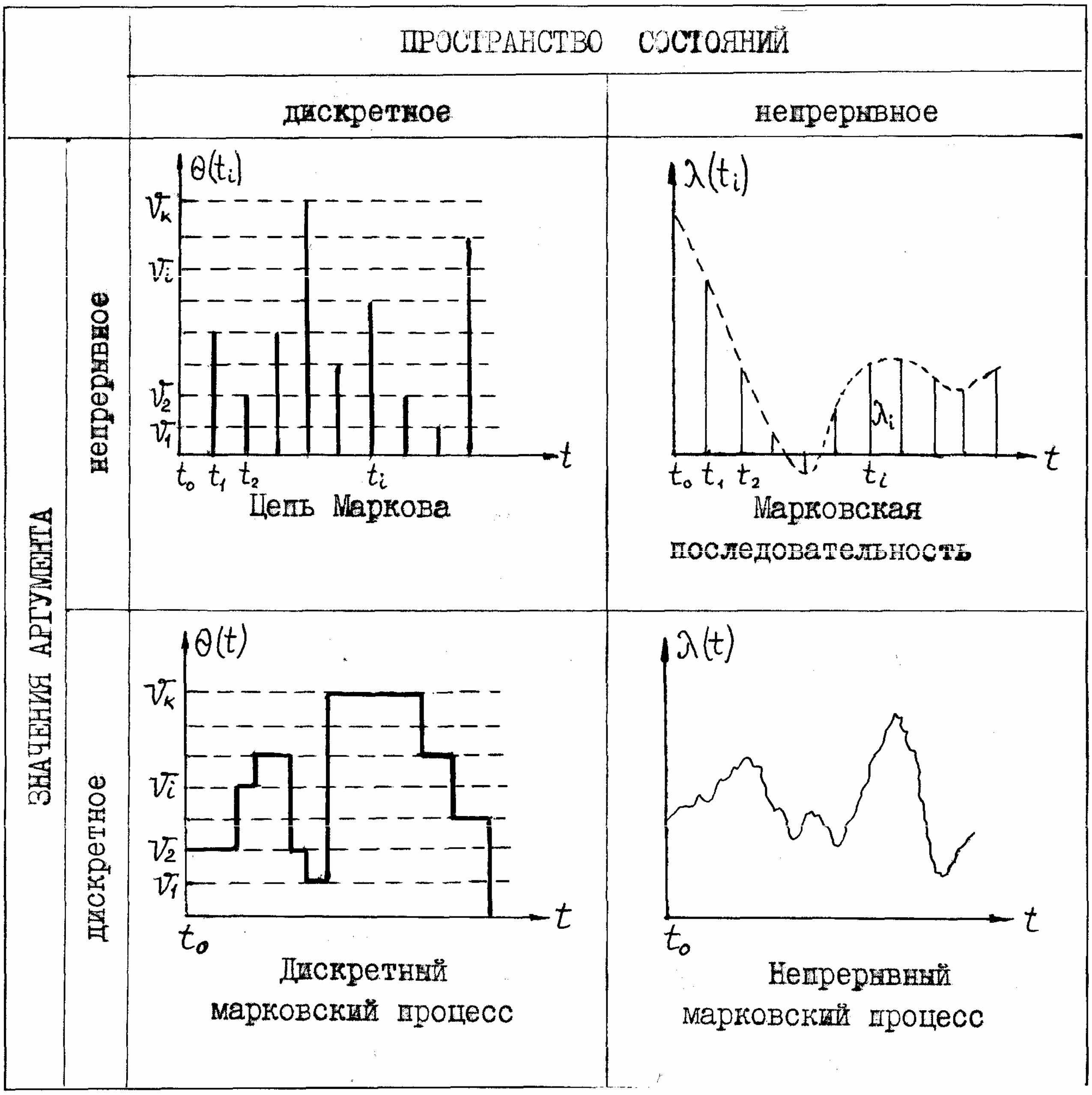


Рис. 10.1

10.2. Графы состояний

При анализе случайных марковских процессов с дискретными пространствами состояний удобно пользоваться наглядной геометрической схемой *графом состояний.* Графы состояний изображают возможные состояния системы с указанием (в виде стрелок) возможных переходов из состояния в состояние.

При этом для случая дискретного пространства состояний и дискретного времени (цепь Маркова) у стрелок проставляются соответствующие *вероятности переходов* (рис. 10.2). Такой граф состояний называют *размеченным,* а величины $\mathbf{P_{ij}}$ означают вероятности переходов из **i**-го состояния в **j**-е состояние.

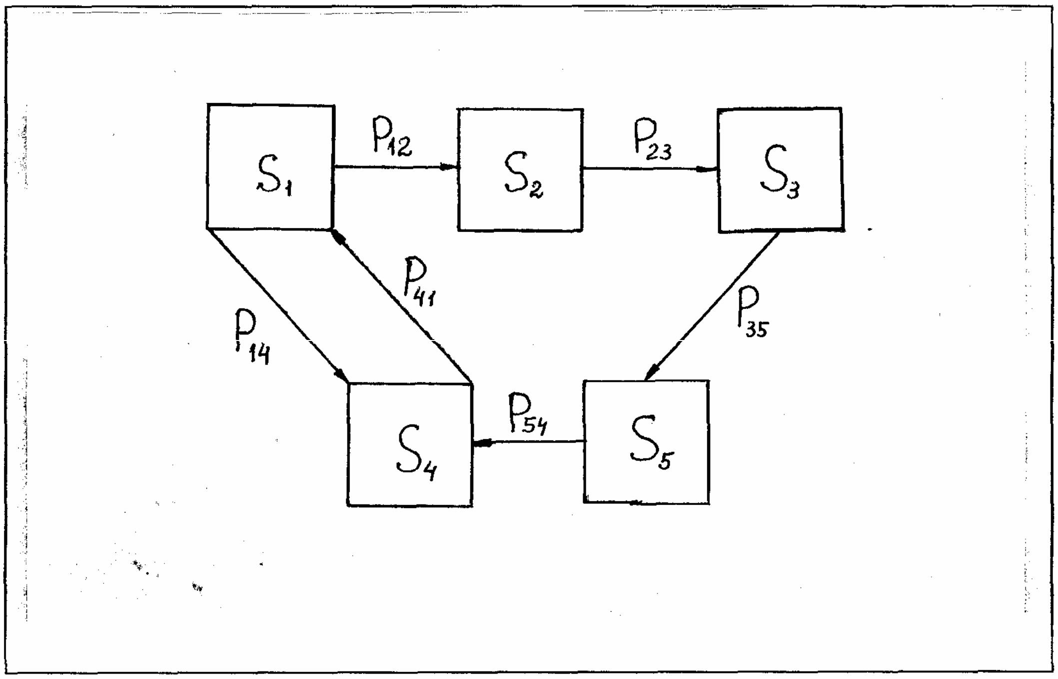


Рис. 10.2

Для случайных марковских процессов с дискретными состояниями и непрерывным временем вместо переходных вероятностей $\mathbf{P_{ij}}$ у стрелок указываются *плотность вероятностей переходов* $\boldsymbol{\lambda_{ij}}$. (рис. 10.3).

Плотность вероятности перехода $\boldsymbol{\lambda_{ij}}$ есть предел отношения вероятности перехода системы за время $\mathbf{\Delta t}$ из состояния $\mathbf{S_i}$ в состояние $\mathbf{S_j}$

$$\lambda_{ij} = \lim_{\Delta t \to 0} \frac{P_{ij}(\Delta t)}{\Delta t} \qquad (10.2)$$

С точностью до бесконечно малых больших порядков вероятность перехода $\mathbf{P_{ij}(\Delta t)}$ за время $\mathbf{\Delta t}$ равна

$$\mathbf{P_{ij}(\Delta t) = \boldsymbol{\lambda_{ij}}\Delta t} \qquad (10.3)$$

Если плотности вероятностей переходов $\boldsymbol{\lambda_{ij}}$ не зависят от времени $\mathbf{t}$, то такой марковский процесс называется *однородным*, при наличии зависимости от времени, т.е., если $\boldsymbol{\lambda_{ij} = \lambda_{ij}(t)}$, процесс называется *неоднородным.*

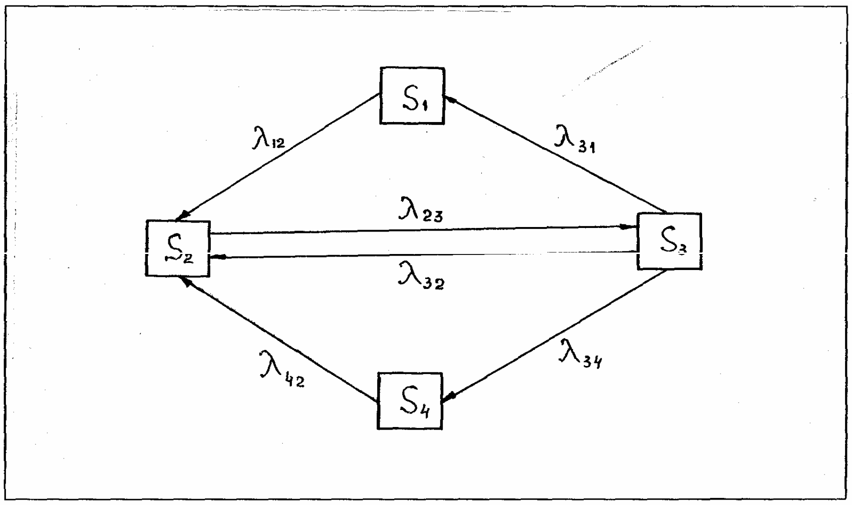


Рис. 10.3

10.3. Марковская цепь

Напомним, что марковской цепью называется случайный марковский процесс с дискретными состояниями и дискретным временем.

Марковская цепь называется *однородной,* если переходные вероятности не зависят от номера шага (не зависят от момента времени перехода). При зависимости переходных вероятностей от номера шага Марковская цепь называется *неоднородной.*

Для однородной марковской цепи для **n** состояний системы переходные вероятности $\mathbf{P_{ij}}$ могут быть записаны в виде квадратной матрицы

$$\left\|P_{ij}\right\| = \left\|\begin{matrix} P_{11} & P_{12} & \dots & P_{1j} & \dots & P_{1n} \\ P_{12} & P_{22} & \dots & P_{2j} & \dots & P_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ P_{i1} & P_{i2} & \dots & P_{ij} & \dots & P_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ P_{n1} & P_{n2} & \dots & P_{nj} & \dots & P_{nn} \end{matrix}\right\| \qquad (10.4)$$

Характерной особенностью матрицы $\mathbf{P_{ij}}$ является то, что сумма членов, стоящих в каждой строке, равна единице, т.е.

$$\sum_{j=1}^{n} P_{ij} = 1 \qquad (10.5)$$

Квадратная матрица, элементы которой неотрицательны и сумма элементов, стоящих в каждой строке (или столбце) равна единице, называется *стохастической* матрицей.

10.4. Марковские процессы с дискретными состояниями и непрерывным временем

Для случайных Марковских процессов с дискретными состояниями и непрерывным временем в графе состояний у стрелок указывается плотность вероятностей переходов $\boldsymbol{\lambda_{ij}}$. Поток событий, переводящий систему из одного состояния в другое, является пуассоновским, т.е. распределение времени между событиями перехода, подчиняется экспоненциальному распределению.

Пусть система имеет конечное число состояний $\mathbf{S_1, S_2, \dots, S_i, \dots, S_n}$.

Вероятность этих состояний для любого момента времени **t**

$$\mathbf{P_1(t), P_2(t), \dots, P_i(t), \dots, P_n(t)} \qquad (10.5)$$

Для любого времени t

$$\sum_{i=1}^{n} P_i(t) = 1 \qquad (10.6)$$

Для нахождения вероятностей (10.5) необходимо решить систему дифференциальных уравнений (уравнений Колмогорова)

$$\frac{dP_i(t)}{dt} = \sum_{j=1}^{n} \lambda_{ij} P_j(t) - P_j(t) \cdot \sum_{i=1}^{n} \lambda_{ij} \qquad \mathbf{I = 1, 2, \dots, n} \qquad (10.7)$$

Уравнения (10.7) легко составить с использованием размеченного графа состояний, пользуясь следующим мнемоническим правилом: производная вероятности каждого состояния равна сумме всех потоков вероятностей, идущих на данное состояние, минус сумма всех потоков вероятностей, идущих из данного состояния.

Возникает вопрос, как будет вести себя система при $t \to \infty$? Существуют ли пределы функций $\mathbf{P_1(t), P_2(t), \dots, P_n(t)}$**?**

Если эти пределы существуют, то соответствующие вероятности состояний называются *предельными* вероятностями состояний (или «финальными», т.е. «конечными»).

Если предельные вероятности существуют, то в этом состоянии имеет место установившийся режим, для которого производные будут равны нулю. В этом случае система дифференцированных уравнений Колмогорова превращается в *систему алгебраических уравнений*. Совместно с нормировочным условием $\sum_{i=1}^{n} P_i(t) = 1$ эти уравнения дают возможность вычислить все предельные вероятности состояний $\mathbf{P_1, P_2, ..., P_n}$.

10.5. Полумарковские процессы

Рассмотрим процесс с дискретными состояниями и непрерывным временем, схема которого изображена на рис.10.4. $\mathbf{S_1, S_2, ..., S_i, ..., S_k, ..., S_e, ..., S_n}$ - состояния системы. Исходное состояние системы - $\mathbf{S_i}$. На схеме изображена следующая последовательность состояний:

$\mathbf{S_i \to S_K \to S_2 \to S_1 \to S_i \to S_e \to ...}$

Следует иметь в виду, что в данном случае это одна из возможных последовательностей реализации процесса. Одношаговые вероятности переходов изображенной на рисунке реализации процесса суть $\mathbf{P_{ik}, P_{k2}, P_{21}, P_{1i}, P_{i1}}$, ... Время $\mathbf{T_{k2}}$ есть время ожидания в состоянии $\mathbf{S_K}$ до перехода в состояние $\mathbf{S_2}$, соответственно: $\mathbf{T_{21}}$ – время ожидания в состоянии $\mathbf{S_2}$ до перехода в состояние $\mathbf{S_1}$, $\mathbf{T_{1i}}$ - в состоянии $\mathbf{S_1}$ до перехода в состояние $\mathbf{S_i}$, $\mathbf{T_{i1}}$ - в состоянии $\mathbf{S_i}$ до перехода в состояние $\mathbf{S_1}$.

Каждое из этих времен ожидания имеет какой-то закон распределения.

Если игнорировать случайный характер времени ожидания и рассматривать только моменты перехода, то процесс **S(t)** будет представлять собой однородную цепь Маркова. Если же учитывать пребывание процесса в разных состояниях в течение случайного отрезка времени, то здесь возможны, в зависимости от характера распределения этого времени, два различных характера процесса.

Если распределение случайного времени ожидания является экспоненциальным, т.е. поток, переводящий систему из состояния в состояние, является пуассоновским, то имеет место чисто марковский процесс с дискретными состояниями и непрерывным временем.

В этом случае справедливы все уравнения, приведенные в п. 10.4, т.е. дифференциальные уравнения Колмогорова и система алгебраических уравнений для предельных вероятностей состояний.

Если же распределение времени ожидания является любым другим (кроме экспоненциального), в том числе это время может быть и некоторым постоянным числом, то процесс не является марковским в чистом виде. Марковский процесс является только и только в момент перехода, при этом вероятности переходов между состояниями $\mathbf{P_{ik}}$ определяются матрицей вероятностей переходов как и для случая марковской цепи (п. 10.3)

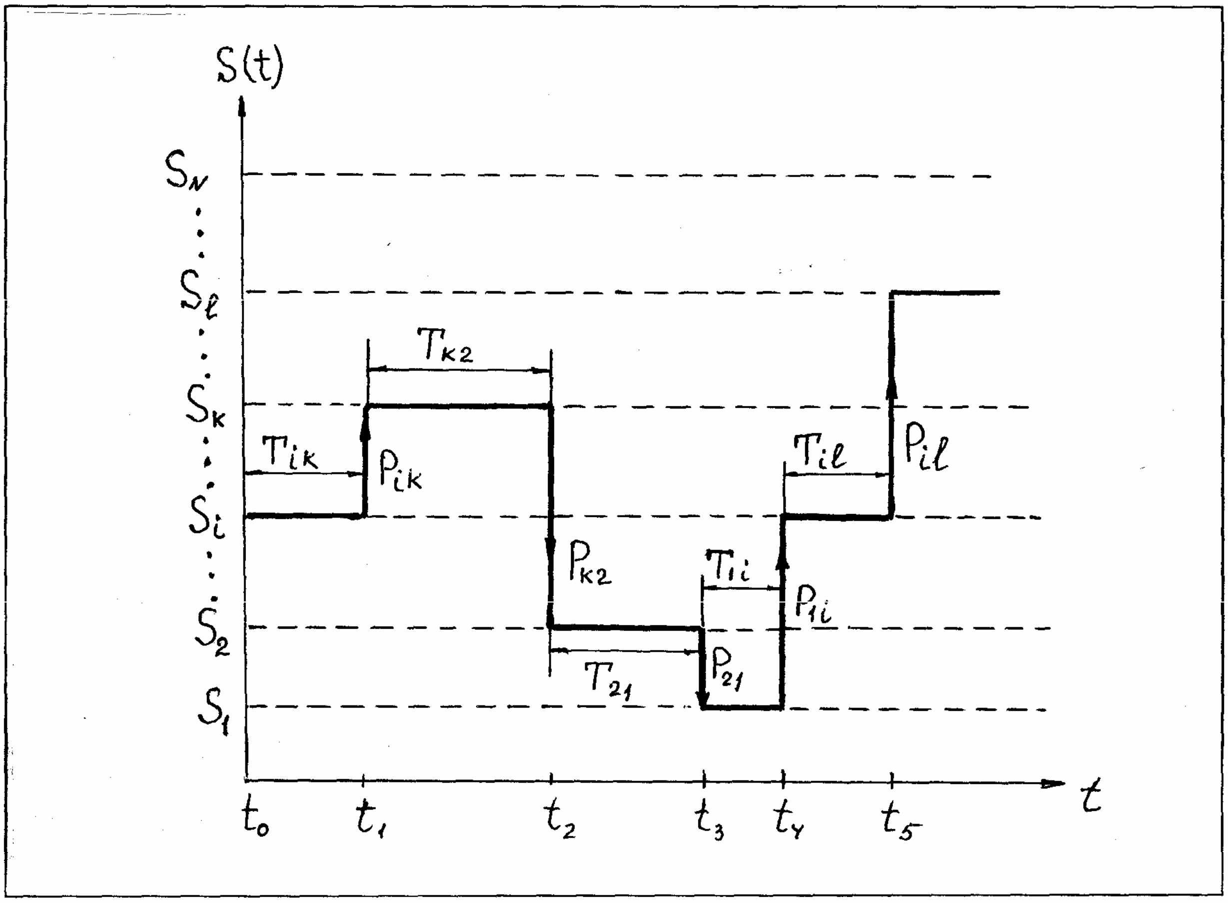


Рис. 10.4

Случайный процесс, при котором переходы между состояниями являются Марковскими, а времена нахождения в любом из состояний описываются произвольной функцией распределения, называется *полумарковским*. В научной литературе применяются также термины *вложенная цепь Маркова* или *вложенный марковский процесс*. Смысл этих терминов состоит в том, что Марковский процесс перехода между состояниями системы происходит внутри другого процесса (не Марковского), вложен в этот другой процесс.

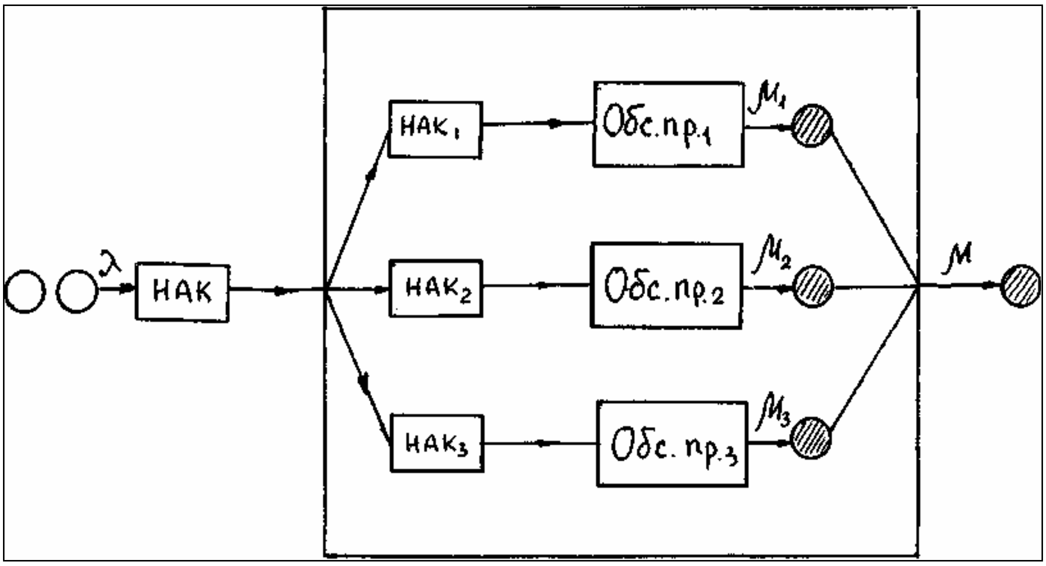
11. АНАЛИЗ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ

11.1. Компоненты систем массового обслуживания

Системы массового обслуживания можно изобразить обобщенной условной схемой, приведённой на рис.11.1.

Основными компонентами системы являются следующие элементы:

- Входящий поток заявок, который может быть охарактеризован интенсивностью потока — λ .
- Приборы (каналы) обслуживания, которых в системе может быть один (одноканальная система) или несколько (многоканальная система).
- Накопители (устройства для обеспечения ожидания обслуживания), которые могут располагаться как перед всей системой, так и перед каждым каналом обслуживания.
- Выходящий поток обслуженных заявок, который может быть охарактеризован интенсивностью обслуживания —μ .



О О поступающие заявки О О обслуженные заявки

Рис. 11.1

Структура системы массового обслуживания определяется количеством и типом обслуживающих приборов, а также наличием накопителей.

Характер обслуживания состоит в порядке обслуживания, который принято называть дисциплиной обслуживания, и величиной μ — интенсивностью обслуживания.

11.2. Классификация систем массового обслуживания

Система массового обслуживания классифицируется по характеру входящего потока требований $П_{вх}$, распределению времени обслуживания $В_{об}$, по числу обслуживающих приборов $N_{пр}$, и ёмкости накопителя (длине очереди) $Е_{нак.}$

В соответствии с наиболее распространённой и общепринятой классификацией Кендалла любая система массового обслуживания характеризуется этими четырьмя параметрами в виде следующей записи

$П_{вх}$ / $В_{об}$ / $N_{пр}$/ $Е_{нак}$

Характер входящего потока требований принято обозначать следующими символами:

M (Markovian) - входящий поток требований является Пуассоновским, т.е. распределение времени между поступающими заявками подчинено экспоненциальному закону;

E (Erlangian) - входящий поток является Эрланговским;

D (Determenistic) - детерминированный постоянный поток;

G (General) - произвольный рекуррентный поток.

Те же символы применяются и для обозначения распределения времени обслуживания:

M - распределение по экспоненциальному закону;

E - распределение по закону Эрланга;

D - время обслуживания постоянная величина;

G - произвольное распределение времени обслуживания.

Число обслуживающих приборов - равно или больше единицы. При $N_{пр} = 1$ систему принято называть одноканальной СМО; при $N_{пр} > 1$ -многоканальной.

Ёмкость накопителя $Е_{нак}$ может варьироваться от $Е_{нак} = 0$ до $Е_{нак} = \infty$. При $Е_{нак} = 0$ поступившая заявка в случае, если все каналы заняты, теряется (получат отказ в обслуживании). Такие системы принято называть - система с потерями (с отказами).

При $Е_{нак} > 1$ система является системой с ожиданием (с очередью). В таких системах заявка, поступившая в момент, когда все каналы заняты, отправляется в накопитель (становится в очередь) и ожидает освобождения хотя бы одного канала.

Обслуживание в системе с ожиданием может быть самым разнообразным. Заявки могут обслуживаться в порядке поступления или в случайном порядке. В некоторых СМО применяется "обслуживание с приоритетом", при котором некоторые заявки имеют "льготы" и по тем или иным признакам обслуживаются в первую очередь.

В системах, в которых входящий поток является пуассоновским, а распределение времени обслуживания - экспоненциальное, процессы обслуживания являются Марковскими процессами. Анализ таких систем может быть сделан методами, изложенными в предыдущей главе.

Приведём условие обозначения наиболее распространенных систем этого типа.

- M/M/1/0 - одноканальная СМО с отказами;
- M/M/n/0 - многоканальная СМО с отказами;
- M/M/1/m - одноканальная СМО с ожиданием (ёмкость накопителя равна m);
- M/M/n/m - многоканальная СМО с ожиданием, но с возможностью отказа (число каналов - n, ёмкость накопителя равна m);
- M/M/1/∞ - одноканальная СМО с ожиданием без отказа (ёмкость накопителя равна ∞).

11.3. Показатели качества обслуживания СМО

В зависимости от типа СМО при оценке качества её функционирования могут применяться различные показатели. В основном - это те показатели, которые представляют интерес для пользователя. Рассмотрим основные показатели качества работы СМО.

Вероятность потери заявки (вероятность отказа) - $P_{отк}$. Заявка получает отказ, когда все каналы обслуживания заняты, а свободных мест ожидания (если есть накопитель) нет.

Вероятность простоя - P_o, это вероятность того, что поступившая заявка немедленно будет обслужена, другими словами - это есть то, что СМО свободна.

Приведённая интенсивность потока заявок - $\rho = \frac{\lambda}{\mu}$. Величина ρ представляет собой среднее число заявок, приходящих в СМО за среднее время обслуживания одной заявки.

Абсолютная пропускная способность - среднее число заявок, которое может обслужить СМО за единицу времени - А.

Относительная пропускная способность - средняя доля поступивших заявок, обслуживаемая системой. Другими словами это отношение среднего числа заявок, обслуживаемых системой в единицу времени, к среднему числу поступающих за это время заявок - q.

Величины А и q связаны соотношением:

$$A = \lambda q \qquad (11.1)$$

Для СМО с неограниченным ожиданием каждая поступившая заявка рано или поздно будет обслужена, поэтому q=l и A=λ.

Для многоканальной системы её характеристикой может быть среднее число занятых каналов - $\bar{z}$.

Среднее число заявок под обслуживанием - ϖ для многоканальной системы без очереди или с ограниченной длиной очереди совпадает со средним значением занятых каналов. Для СМО с m=∞ при ρ >1 $\varpi = \rho$.

Для СМО с очередью важным параметром для пользователя может быть среднее время ожидания в очереди - $t_{ож}$. Для любой системы существенное значение имеет общее время пребывания в системе -$t_{сист}$ (в очереди и под обслуживанием).

11.4. Анализ систем массового обслуживания с отказами

11.4.1. Система M/M/1/0 - одноканальная СМО с отказами

В системе рассматриваемого типа поток заявок является пуассоновским, интенсивность этого потока - λ. Поток обслуживания - μ тоже является простейшим.

Анализ системы начнём с установления возможных состояний системы. Этих состояний всего два: So - канал свободен и S1 - канал занят. Граф состояний приведён на рис.11.2.

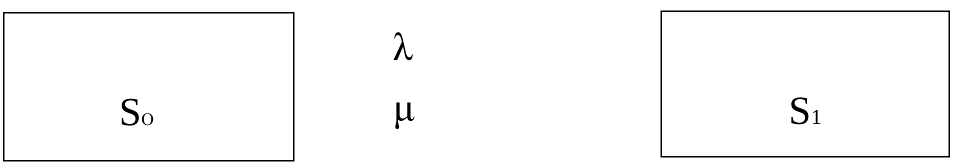


Рис. 11.2

Для случая постоянного потока заявок, т.е. для случая λ =const вероятность того, что канал свободен, выражается зависимостью

$$P_O(t)= \frac{\mu}{\lambda+\mu}+\frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t} \qquad (11.2)$$

Величина P1(t) дополняет величину Po(t) до единицы в соответствии с нормировочным условием P_0 (t) + P_1 (t) = 1. Зависимости графически изображены на рис.11.3.

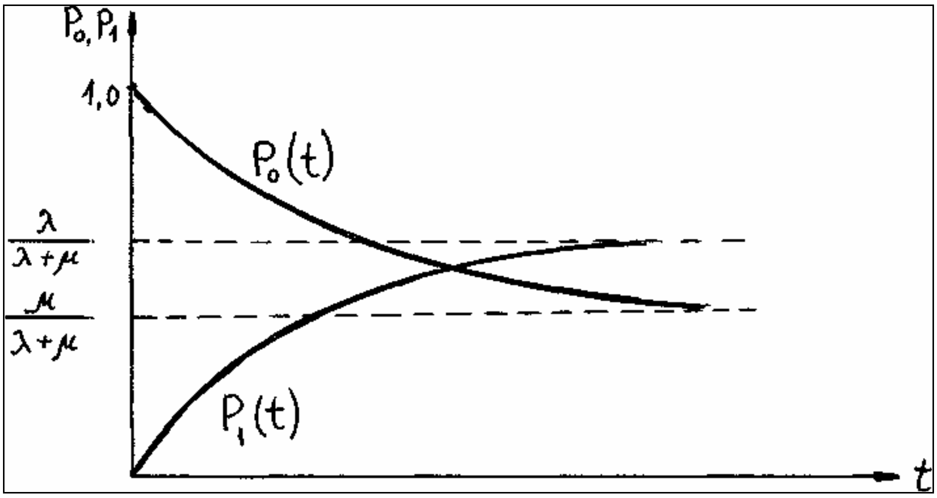


Рис.11.3

В начальный момент канал заведомо свободен и $P_0(t)$ = l. С увеличением t величина P_0 (t) уменьшается и при $t \to \infty$ P_0 (t) = $\frac{\mu}{\lambda+\mu}$. Соответственно для начального момента P_1 (t)= 0 и в пределе $P_0(t)=\frac{\mu}{\lambda+\mu}$.

Для одноканальной системы с отказами величина ро есть не что иное, как относительная пропускная способность

$$q= P_0 (t)=\frac{\mu}{\lambda+\mu} \qquad (11.3)$$

Абсолютная пропускная способность в соответствии с формулой (11.1) есть

$$A= \lambda *q=P_0 (t) = \lambda\frac{\mu}{\lambda+\mu} \qquad (11.4)$$

Вероятность P_0 (t) есть не что иное, как вероятность отказа $P_{отк}$ = P_1 = 1 - P_0 =q. В пределе (при $t \to \infty$)

$$P_{отк}(t)= 1-\frac{\mu}{\lambda+\varpi} = \frac{\lambda}{\lambda+\mu} \qquad (11.5)$$

11.4.2. Система M/M/n/o – многоканальная СМО с отказами

Состояния n - канальной системы:

- S_0 - система полностью свободна;
- S_1 - занят один канал, остальные каналы свободны;
- S_2- занято два канала, остальные каналы свободны;
- ....;
- S_i - занято i каналов, остальные каналы свободны;
- ....;
- S_n - заняты все n каналов.

Размеченный граф состояний рассматриваемой системы представлен на рис. 11.4.

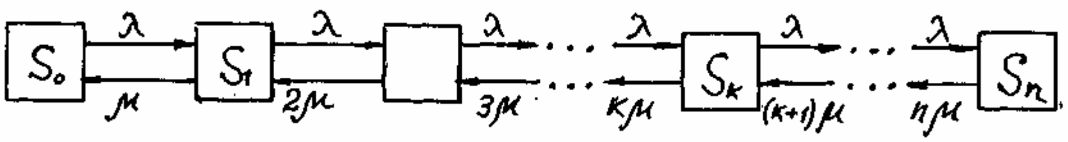


Рис.11.4

У стрелок, идущих слева направо, везде стоит интенсивность λ потока заявок. Интенсивности же потоков обслуживания увеличиваются от состояния к состоянию.

Поскольку входной поток и поток обслуживания являются пуассоновскими, то процесс является марковским и, пользуясь правилом, изложенным в п.10.4, можно составить систему дифференциальных уравнений для вероятностей состояний

$$\left.\begin{aligned}
&\frac{dP_0(t)}{dt} = -\lambda \cdot P_0(t) + \mu \cdot P_1(t) \\
&\frac{dP_1(t)}{dt} = -(\lambda + \mu) \cdot P_1(t) + \lambda \cdot P_0(t) + 2\mu \cdot P_2(t) \\
&\frac{dP_2(t)}{dt} = -(\lambda + 2\mu) \cdot P_2(t) + \lambda \cdot P_1(t) + 3\mu \cdot P_3(t) \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\frac{dP_k(t)}{dt} = -(\lambda + k\mu) \cdot P_k(t) + \lambda \cdot P_{k-1}(t) + (k+1)\mu \cdot P_{k+1}(t) \\
&\frac{dP_n(t)}{dt} = -n\mu \cdot P_n(t) + \lambda \cdot P_{n-1}(t)
\end{aligned}\right\} \qquad (11.6)$$

Начальными условиями для решения этих уравнений являются следующие значения вероятностей состояний при t=0:

$$P_0(0) = 1 \; ; \; P_1(0) = P_2(0) = \ldots\ldots = P_k(0) = \ldots\ldots = P_n(0) = 0$$

Аналитическое решение для получения значений вероятностей состояний в функции времени довольно сложно. При необходимости целесообразно воспользоваться имеющимися стандартными программами на ЭВМ и получить значения с требуемой точностью и в интересующие нас моменты времени.

Практическое значение имеют предельные вероятности состояний P_0, P_1, P_2, ... P_k, ... P_n. Для этого систему дифференциальных уравнений (11.6) преобразуем в систему алгебраических уравнений.

Выражение для вероятности того, что рассматриваемая система свободна, имеет вид:

$$P_0 = \left[1 + \frac{\lambda/\mu}{1!} + \frac{\left(\lambda/\mu\right)^2}{2!} + \ldots + \frac{\left(\lambda/\mu\right)^n}{n!} \right]^{-1} \tag{11.7}$$

Для любых других вероятностей состояний получаем

$$P_k = \frac{\lambda^k}{\mu \cdot 2\mu \cdot \ldots \cdot k\mu} P_0 = \frac{\left(\lambda/\mu\right)^k}{k!} P_0 \quad \mathbf{(k=1,2,\ldots,n)} \tag{11.8}$$

В формулы (11.7) и (11.8) входит отношение интенсивности потока заявок λ к интенсивности потока обслуживания (одного канала) μ. В последующем будем обозначать это соотношение так $\frac{\lambda}{\mu} = \rho$.

Заявка получает отказ, если приходит в момент, когда все n каналов заняты. Таким образом, вероятность отказа $P_{отк}$ равна

$$P_{отк} = P_n = \frac{\rho^n}{n!} P_o \tag{11.9}$$

Относительная пропускная способность q есть вероятность того, что заявка будет принята к обслуживанию, т.е. численно дополняет $P_{отк}$ до единицы

$$q = 1 - P_{отк} \tag{11.10}$$

Абсолютная пропускная способность

$$A = \lambda \cdot q = \lambda \cdot (1 - P_{отк}) \tag{11.11}$$

Среднее число занятых каналов $\bar{z}$ (в данном случае оно совпадает со средним числом заявок, находящихся в системе) может быть определено из следующих соображений. Ранее введённая абсолютная пропускная способность A есть среднее число заявок, обслуживаемых в единицу времени. Один занятый канал обслуживает в среднем за единицу времени μ заявок. Таким образом, среднее число занятых каналов будет равно частному от деления A на μ, т.е.

$$\bar{z} = \frac{A}{\mu} = \frac{\lambda \cdot (1 - P_{\text{отк}})}{\mu} = \rho \cdot (1 - P_{\text{отк}})$$

(11.12)

11.5. Анализ систем массового обслуживания с ожиданием

11.5.1. Системы M/M/1/m и M/M/1/∞ - одноканальные системы с ожиданием

Рассмотрим сначала систему M/M/1/m - одноканальную систему с ограниченной очередью.

Возможные состояния этой системы:

- S_0 - канал свободен;
- S_1 - канал занят, очереди нет;
- S_2 - канал занят, одна заявка стоит в очереди;
- ……………………………………………………..
- S_i - канал занят , i-1 заявок стоит в очереди;
- ……………………………………………………..
- S_m+i - канал занят, m заявок стоит в очереди (накопитель полностью загружен).

Заявка, приходящая в момент, когда система находится в состоянии S $_m$+i, получает отказ.

Граф состояний рассматриваемой системы приведен на рис. 11.5.

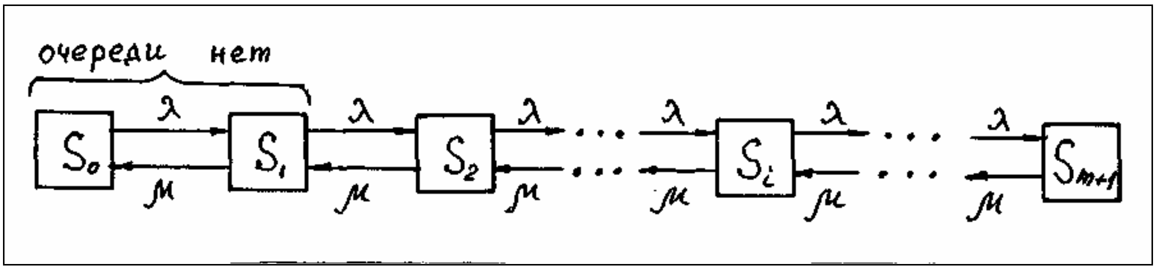


Рис. 11.5

Можно записать для предельных вероятностей состояний

$$\left.\begin{aligned} P_1 &= \rho \cdot P_0 \\ P_2 &= \rho^2 \cdot P_0 \\ &\ldots\ldots\ldots\ldots\ldots \\ P_k &= \rho^k \cdot P_0 \\ &\ldots\ldots\ldots\ldots\ldots \\ P_{m+1} &= \rho^{m+1} \cdot P_0 \end{aligned}\right\}$$

(11.13)

где, как и для случая многоканальной системы с отказами, но при **n=1**

$$P_0=\left[1+\rho+\rho^2+\ldots+\rho^{m+1}\right]^{-1} \qquad (11.14)$$

Здесь выражение в скобках представляет собой геометрическую прогрессию с первым членом 1 и знаменателем прогрессии, равным р. Суммируя эту прогрессию, находим

$$P_0=\left[\left(1-\rho^{m+2}\right)\Big/\left(1-\rho\right)\right]^{-1}=\frac{1-\rho}{1-\rho^{m+2}}$$

(11.15)

Формула (11.15) справедлива только при $\rho \neq 1$. При $\rho=1$ получаем неопределенность вида $\frac{0}{0}$, раскладывая которую по правилу Лопиталя, получим

$$P_0=\frac{1}{m+2} \qquad (11.16)$$

Вероятность отказа $P_{отк}$ определим из того, что отказ-заявку получаем в том случае, когда занят единственный канал обслуживания и заняты все m мест в очереди, т.е. как следует из формул 11.13 и 11.15

$$P_{отк}=p_{m+1}=\rho^{m+1}\cdot P_0=\frac{\rho^{m+1}(1-\rho)}{1-\rho^{m+2}}$$

(11.17)

Абсолютная пропускная способность

$$q=1-P_{отк}=1-\frac{\rho^{m+1}(1-\rho)}{1-\rho^{m+2}}=\frac{1-\rho^{m+1}}{1-\rho^{m+2}}$$

(11.18)

Относительная пропускная способность

$$A=\lambda\cdot q \qquad (11.19)$$

Среднее число заявок $\bar{r}$, находящихся в очереди, определим как математическое ожидание дискретной случайной величины R - числа заявок, находящихся в очереди. Величину $\bar{r}$ можно получить, умножая число заявок в очереди на соответствующие вероятности и производя суммирование

$$\bar{r}=\mathbf{1*P_2+2*P_3+\ldots+(k-l)*P_k+\ldots+m*P_m}$$

Учитывая выражение для $\mathbf{P_k}$ (**k** = 2, **3, ...m+1)** и для $\mathbf{P_0}$, после соответствующих преобразований можно получить

$$\bar{r}=\frac{\rho^2\left[1-\rho^m(m+1-m\rho\rho\right]}{\left(1-\rho^{m+2}\right)(1-\rho)}$$

(11.20)

Общее число заявок в системе

$\bar{k}=\bar{r}+\bar{w}$, где $\bar{w}$ - математическое ожидание числа заявок, находящихся под обслуживанием

$$\overline{w} = 0 \cdot P_0 + 1 \cdot (1 - P_0) = \frac{\rho - \rho^{m+2}}{1 - \rho^{m+2}} \tag{11.21}$$

т.е.

$$\overline{k} = \overline{r} + \frac{\rho - \rho^{m+2}}{1 - \rho^{m+2}}$$

где $\overline{r}$ определяем по зависимости (11.20).

Среднее время ожидания $\bar{t}_{\text{ож}}$ будет очевидно равно средней длине очереди, умноженной на среднее значение времени в потоке заявок. Для пуассоновского потока заявок среднее время $t_{\text{ср}} = \frac{1}{\lambda}$. Таким образом имеем $\bar{t}_{\text{ож}} = \frac{\bar{r}}{\lambda}$.

Среднее время обслуживания одной заявки $\bar{t}_{\text{ср}} = \frac{1}{\mu}$. Математическое ожидание времени обслуживания равно $q\frac{1}{\mu}$, общее среднее время пребывания заявки в системе

$$\bar{t}_{\text{сист}} = \frac{\bar{r}}{\lambda} + \frac{q}{\mu} \tag{11.22}$$

Для системы М/М/1/∞, т.е. одноканальной системы с бесконечной очередью, установившийся режим возможен только при ρ<1. При ρ>1 очередь растёт до бесконечности. Вероятности состояний системы М/М/1/∞ для случая ρ<1 можно получить из выражений для рассмотренной выше системы М/М/1/m путём предельного перехода $m \to \infty$.

Получим:

$$\left.\begin{array}{l} P_0 = 1 - \rho \\ P_1 = \rho \cdot (1 - \rho) \\ P_2 = \rho^2 \cdot (1 - \rho) \\ \dots\dots\dots\dots\dots\dots \\ P_k = \rho^k \cdot (1 - \rho) \\ \dots\dots\dots\dots\dots\dots \end{array}\right\} \tag{11.23}$$

Пользуясь тем же приёмом предельного перехода, получим выражения для показателей качества:

$$\left.\begin{aligned}
&P_{отк} = 0\\
&q = 1, \quad A = q\lambda = \lambda\\
&\bar{r} = \frac{\rho^2}{1-\rho}\\
&\bar{k} = \frac{\rho}{1-\rho}\\
&\bar{t}_{ож} = \frac{1}{\mu}\cdot\frac{\rho}{1-\rho} = \frac{\rho^2}{\lambda(1-\rho)}\\
&\bar{t}_{сист} = \frac{1}{\mu}\cdot\frac{\rho}{1-\rho}
\end{aligned}\right\} \qquad (11.24)$$

11. 5.2. Система M/M/n/m - многоканальная система с ожиданием

На n - каналов системы M/M/n/m поступает поток заявок с интенсивностью λ, интенсивность обслуживания одного канала - μ, число мест в очереди - m.

Возможные состояния системы:

- S_0 - система полностью свободна;
- S_1 - занят один канал, остальные каналы свободны;
- ....;
- S_k - занято k каналов, остальные каналы свободны;
- ....;
- S_n - заняты все n каналов;
- S_{n+1} - заняты все n каналов, одна заявка в очереди;
- ....;
- S_{n+r} - заняты все n каналов, r заявок в очереди;
- ....;
- S_{n+m} - заняты все n каналов, заняты все m мест в очереди.

В состояниях So $\div S_n$ - очереди нет.

Граф состояний представлен на рис. 11.6 изображает знакомую уже нам схему размножения и гибели.

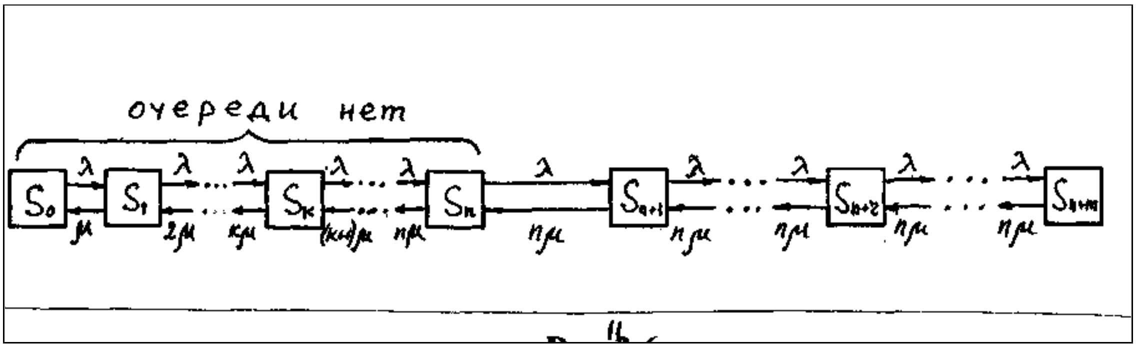


Рис.11.6

Воспользуемся полученными ранее для этой схемы решениями.

$$P_0 = \left[1 + \rho + \frac{\rho^2}{2} + \ldots + \frac{\rho^n}{n!} + \frac{\rho^{n+1}}{n \cdot n!} + \frac{\rho^{n+2}}{n^2 \cdot n!} + \ldots + \frac{\rho^{n+m}}{n^m \cdot n!}\right]^{-1}$$

(11.25)

Последние подчеркнутые члены образуют геометрическую прогрессию со знаменателем $\frac{\rho}{n}$. Определив сумму этих членов,

получаем следующее выражение для **P_0**

$$P_0 = \left[1 + \rho + \frac{\rho^2}{2} + \ldots + \frac{\rho^n}{n!} + \frac{\rho^n}{n!} \cdot \frac{\frac{\rho}{n} - \left(\frac{\rho}{n}\right)^{m+1}}{1 - \frac{\rho}{n}} \right]^{-1}$$

(11.26)

Для остальных вероятностей состояний, используя уже имеющийся принцип решения, получаем:

$$\left.\begin{aligned}
&P_1 = \frac{\rho}{1!} P_0 \\
&P_2 = \frac{\rho^2}{2!} P_0 \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
&P_n = \frac{\rho^n}{n!} P_0 \\
&P_{n+1} = \frac{\rho^{n+1}}{n \cdot n!} P_0 \\
&P_{n+2} = \frac{\rho^{n+2}}{n^2 \cdot n!} P_0 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&P_{n+m} = \frac{\rho^{n+m}}{n^m \cdot n!} P_0
\end{aligned}\right\}$$

(11.27)

Запишем далее выражения для показателей качества рассматриваемой системы.

$$\mathbf{P_{отк}} = \mathbf{P_{n+m}} = \frac{\rho^{n+m}}{n^m \cdot n!} \cdot \mathbf{P_o} \qquad (11.28)$$

Это выражение означает, что заявка получает отказ, если все каналы и все места в очереди заняты.

Относительная пропускная способность дополняет вероятность отказа до единицы.

$$q = 1 - P_{отк} = 1 - \frac{\rho^{n+m}}{n^m \cdot n!} P_0$$

(11.29)

Абсолютная пропускная способность

$$A = \lambda \cdot q = \lambda \cdot \left(1 - \frac{\rho^{n+m}}{n^m \cdot n!} P_0 \right)$$

(11.30)

Среднее число занятых каналов

$$\bar{z} = \frac{A}{\mu} \qquad (11.31)$$

Среднее число заявок в очереди

$$\bar{r} = 1 \cdot P_{n+1} + 2 \cdot P_{n+2} + \ldots + m \cdot P_{n+m} =$$
$$= \frac{\rho^{n+1}}{n \cdot n!} \cdot \left[1 + 2\frac{\rho}{n} + 3\left(\frac{\rho}{n}\right)^2 + \ldots + m \cdot \left(\frac{\rho}{m}\right)^{m+1} \right] P_0 =$$
$$= \frac{\rho^{n+1}}{n \cdot n!} \cdot \frac{1 - (m+1)\left(\frac{\rho}{n}\right) + m\left(\frac{\rho}{n}\right)^{m+1}}{\left(1 - \frac{\rho}{m}\right)^2} \cdot P_0$$

(11.32)

Общее число заявок в системе

$$\bar{k} = \bar{z} + \bar{r} \qquad (11.33)$$

Время ожидания

$$\bar{t}_{ож} = \frac{\bar{r}}{\lambda} \qquad (11.34)$$

Время пребывания в системе

$$\bar{t}_{сист} = \bar{t}_{ож} + \frac{q}{} \qquad (11.35)$$

Выражения для параметров системы типа M/M/m/∞ получим из предыдущих формул путём предельного перехода **m**$\to \infty$**.** Полученные ниже выражения справедливы только при $\frac{\rho}{n} < 1$.

Из выражения (11.26) получим

$$P_0 = \left(1 + \rho + \frac{\rho^2}{2} + \ldots + \frac{\rho^n}{n!} + \frac{\rho^{n+1}}{n!(n-\rho)}\right) \quad (11.36)$$

Далее получим

$P_{отк} = 0$, q=1, A=$\lambda \cdot q = q$

Величину $\bar{r}$ получаем из (11.32) при m$\to \infty$

$$\bar{r} = \frac{\rho^{n+1}}{n \cdot n! \left(1 - \frac{\rho}{n}\right)^2} P_0 \quad (11.37)$$

Среднее число занятых каналов

$$\bar{z} = \frac{A}{\mu} = \frac{\lambda}{\mu} = \rho \quad (11.38)$$

$\bar{t}_{сист}$ и $\bar{t}_{ож}$ определяются по формулам (11.34) и (11.35).

12. МЕТОД СЕТЕВОГО ПЛАНИРОВАНИЯ

Среди большого многообразия различных систем имеются такие системы, в которых существенную роль играют последовательность отдельных этапов работ и их взаимосвязей. Примерами таких систем являются системы технического обслуживания, строительство большого промышленного объекта и т.п. Характерным для таких систем является то, что этапы работ, которые выполняются как бы независимо друг от друга, фактически взаимно обуславливают друг друга, так что выполнение некоторых работ не может быть начато раньше, чем будут завершены некоторые другие.

Для решения задач планирования работ в таких системах используют метод сетевого планирования или сетевое планирование управления.

12.1. Структурные таблицы и сетевые графики

Основным исходным материалом для сетевого планирования является список или перечень работ, в котором указана их взаимная обусловленность. Если в такой таблице перечень работ не упорядочен, то производится его упорядочение. Работы подразделяются на ранги. Работы первого ранга - это такие работы, для выполнения которых не требуется выполнение никаких работ (имеется в виду данный комплекс работ). Работы второго ранга - это такие работы, которые выполняются, которые обусловлены (опираются) одной или несколькими работами первого ранга, и т.д.

Пусть задана некоторая упорядоченная структурная таблица (таблица 12.1).

В ней обозначено, что работа a_4 опирается на работы a_1 и a_2 ; работа a_5 на работы a_1 и a_2 и т.п. В этой же таблице указана продолжительность работ t_{ai}, в некоторых условных единицах. Более наглядное представление о взаимозависимости всех работ дают сетевые графики. На рис 12.1 изображен сетевой график, отображающий взаимосвязь работ, перечисленных в таблице 12.1

Таблица 12.1

Работа	Опирается на работы	t_{ai}
a_1	—	10
a_2	—	5
a_3	—	15
a_4	a_1, a_2	18
a_5	a_2, a_6	19
a_6	a_4	18
a_7	a_5, a_6	8
a_8	$a_3 a_5$, a_6	25
a_9	a_7	30
a_{10}	a_5, a_8	10

При построении сетевого графика применяют два основных понятия: работа и событие. Работа - это процесс, приводящий к определённому результату. Событие - означает факт завершения предшествующего комплекса

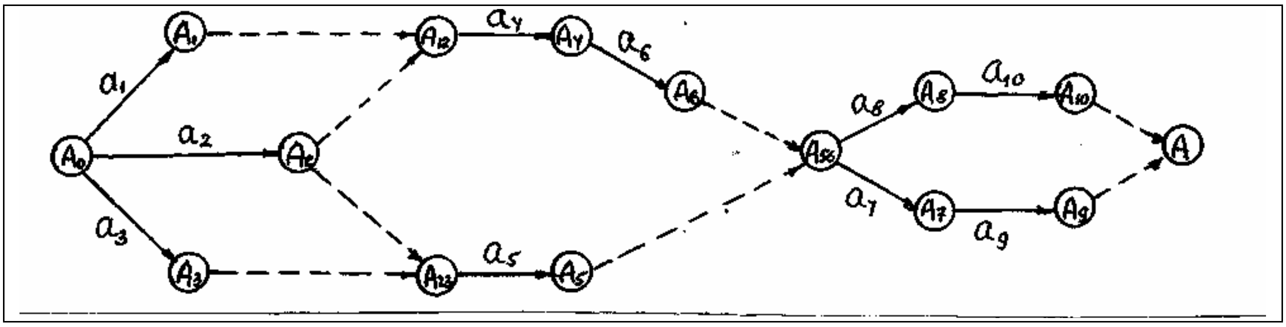


Рис 12.1

работ. Событие в отличие от работы не имеет продолжительности и не требует затрат материальных ресурсов.

Различают следующие виды работ:

- действительные работы, которые сопровождаются затратами времени и ресурсов (на рис. 12.1 изображены сплошными линиями);
- фиктивные работы, которые не требуют затрат ресурсов, но показывают взаимосвязь начала какой-либо работы от окончания другой (на рис. 12.1 изображены пунктирными линиями).

Основными характеристиками каждой работы являются ресурсы, необходимые для её выполнения: время, количество специалистов, материальные ресурсы (оборудование, запчасти, сырьё и т.п.).

Разновидностью действительной работы является ожидание - процесс, требующий только затрат времени (например, простой специалистов в ожидании освобождения салона самолёта).

На рис.12 1 приняты следующие обозначения:

A_0- исходное событие;

a_1, a_2, a_3 - работы первого ранга;

A_1, A_2, A_3 , - события, означающие завершение работ a_1, a_2, a_3;

A_{12} - событие, означающее завершение работ и A_1 и A_2 (a_4 опирается на работы a_1 и a_2);

A_{23} - событие, означающее завершение работ и A_2 и A_3 (a_5 опирается на работы a_2 и a_3);

A_{56} - событие, означающее завершение работ и A_5 и A_6 (a_5 опирается на работы a_2 и a_3);

А - событие, означающее завершение всех работ.

Основные правила составления сетевого графика:

- на графике не должно быть событий, кроме завершающего, с которых не начинается ни одной работы;
- не должно быть событий, кроме исходного, в которое не входит ни одной работы;
- не должно быть замкнутых контуров работы;
- при наличии между двумя событиями нескольких работ, выполняемых параллельно, для определённости вводят дополнительные события и фиктивные работы (фиктивные работы вводят также для обозначения зависимости отдельных работ в сети).

12.2. Временные сетевые графики

В рассмотренном выше сетевом графике нет указаний на время начала и окончания работ. Если соединить сетевой график с осью времени, то мы получим временной сетевой график (рис.12.2). На этом графике проекция длины каждой стрелки будет соответствовать времени выполнения этой работы, условные значения времён t_{ai} взяты из таблицы 12.1.

Общее время выполнения работ - время от события A_0 до события A_9 = А соответствуют Т = 84 единицам. Это время представляет собой сумму времён исполнения не всех работ, а только некоторых из них:

$$\mathbf{T - t_{a1} + t_{a4} + t_{a6} + t_{a7} + t_{a9} = 10 + 18 + 18 + 8 + 30 = 84}$$

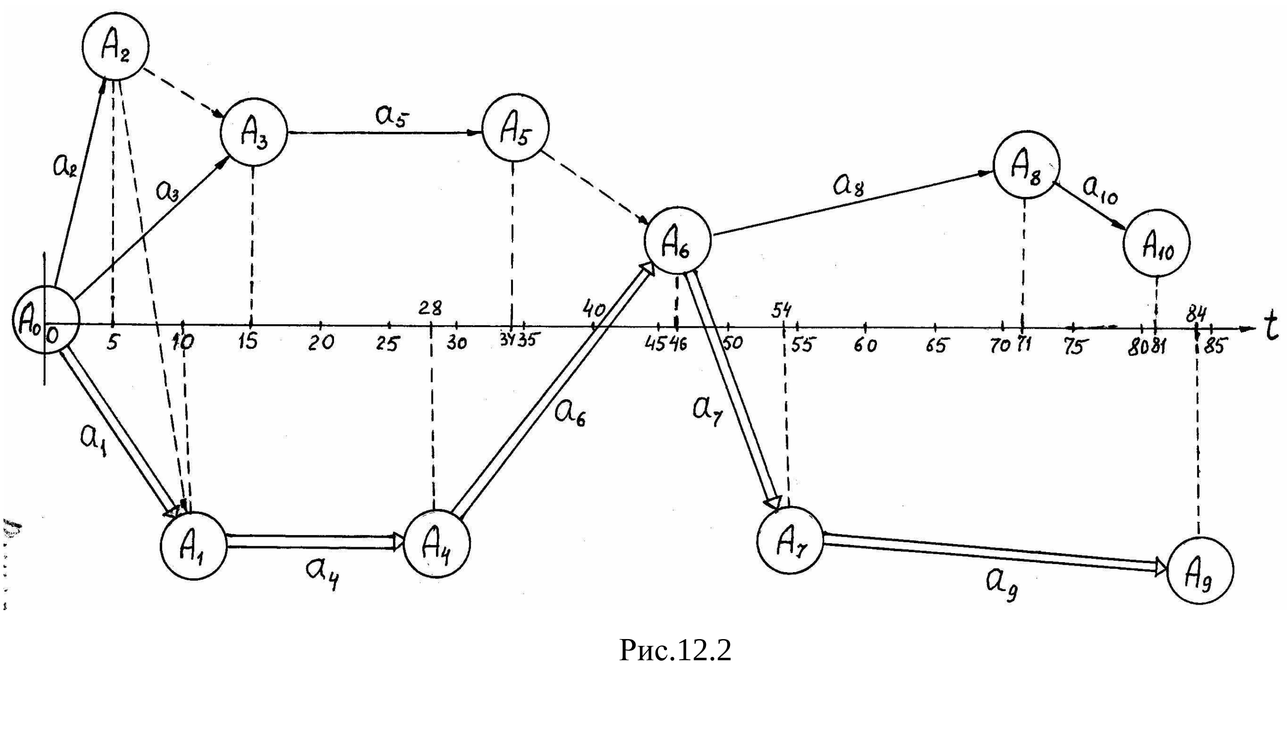


Рис.12.2

Работы a_1, a_4, a_6,a_7 и a_9 называются критическими работами, а цепочка, обозначенная двойными стрелками, является критическим путём.

Особенность критических работ состоит в следующем. Для того, чтобы было соблюдено минимальное время комплекса работ, каждая из работ должна начинаться точно в момент, когда закончена последняя из работ, на которые она опирается, и продолжаться не более того времени, которое отведено ей по плану. Малейшее запаздывание в выполнении каждой из критических работ приводит к задержке выполнения плана в целом.

Таким образом, критический путь на сетевом графике - это совокупность наиболее уязвимых "слабых мест" плана, которые должны укладываться во временной план с наибольшей чёткостью.

Что касается остальных ("некритических") работ (в нашем случае это a_2, a_3, a_5,a_8 и a_{10}), то с ними дело обстоит не так плохо. Каждая из этих работ имеет известные временные резервы и может быть закончена с некоторым запозданием без ущерба для срока выполнения всего комплекса работ.

По временному сетевому графику могут быть определены резервы, соответствующие некритическим работам.

Рассмотрим "некритические дуги" - совокупность некритических работ, начинающихся и кончающихся на критическом пути.

В нашем случае это следующие "некритические дуги":

1 . A_0 - a_2 - A_2- A_1 (одна некритическая работа a_2);

2. A_0 - a_3 - A_3 - a_5 - A_5 - A_6 (две некритические работы a_3 и a_5);

3. A_0 - a_2 - A_2 - A_3- a_5 - A_5 - A_6 (две некритические работы a_2 и a_5);

4. A_6 – a_8 - A_8 - a_{10} - A_{10} (две некритические работы a_8 и a_{10}).

На первой некритической дуге некритическая работа a_2; на замыкающем отрезке одна некритическая работа a_1. Резерв времени, приходящийся на работу a_2, равен $R_2 = t_1 - t_2 = 10 - 5 = 5$, т.е. без ущерба для общего срока выполнения работы, работа a_2 может быть задержана на 5 единиц времени.

Для второй некритической дуги имеем две некритические работы a_3 и a_5 , а на замыкающем участке три некритические работы a_1, a_4 и a_6. Резерв времени на работы a_3 и a_5 $R_{3,5} = t_1 + t_4 + t_6 - (t_3 + t_5) = 10+18+18-(5+19) = 22$. Но уже известно, что $R_2 = 5$, а для 3s не больше 15, так что третья некритическая дуга не даёт ничего нового. На четвёртой некритической дуге резерв времени равен $R_{8,10} = t_7 + t_9 - (t_8 + t_{10}) = 8+30-(25+8) = 5$. Этот резерв может быть распределён между a_8 и a_{10}.

Из изложенного выше следует, что знание критического пути полезно в двух отношениях.

Во-первых, оно позволяет выявить совокупность наиболее "угрожаемых" работ, наблюдать за ними и, в случае необходимости, их форсировать.

Во-вторых, это значение даёт возможность ускорить выполнение всего комплекса работ за счёт привлечения ресурсов, скрытых в некритических работах, если удаётся за счёт их "безвредного" замедления перебросить часть сил и средств на более важные критические работы.

12.3. Пример составления сетевого графика

Имеется структурная таблица обслуживания кратковременной стоянки в аэропорту (таблица 12.2).

Таблица 12.2

Условное обозначение работы	Содержание работы	Опирается на работы	t_{ai}
a_1	Руление самолёта к аэровокзалу после приземления	—	8
a_2	Разгрузка самолёта	a_1	50
a_3	Выход пассажиров из самолёта	a_1	3
a_4	Техническое обслуживание самолёта	a_1	42
a_5	Заправка самолёта топливом	a_3	29
a_6	Загрузка самолёта	a_2	29
a_7	Посадка пассажиров в самолёт	a_5 , a_6	15
a_8	Подсоединение к самолёту тягача	a_5 , a_6	6
a_9	Выруливание самолёта на предварительный старт	a_4 , a_7	10
a_{10}	Взлёт самолёта	a_9	10

На рис 12.3 показан временной сетевой график рассматриваемых работ.

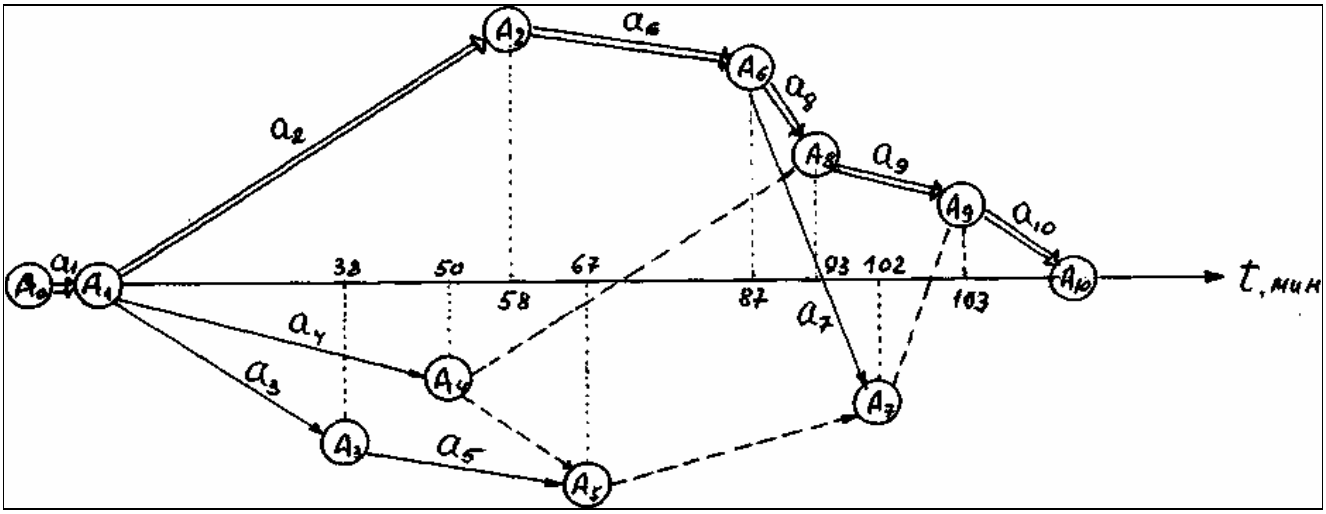


Рис.12.3

Критический путь: A_0- a_1 - A_1 - a_2 - A_2 - a_6 - A_6 - a_8 - A_8 - a_9 - A_9 - a_{10} - A_{10}

Общая продолжительность работ равна сумме критических работ.

$T = t_{a1} + t_{a2} + t_{a6} + t_{a8} + t_{a9} + t_{a10} = 80 + 50 + 29 + 60 + 10 + 10 = 113$ мин.

Некритические дуги:

1. A_1 - a_3 - A_3- a_5 - A_5 - A_6;

Некритические работы a_3 и a_5. Резерв времени $(a_2 + a_6) - (a_3 + a_5) = (50 + 29) - (30 + 20) = 20$ мин. Этот резерв может быть использован между выходом пассажиров из самолёта (работа a_3) и заправкой самолёта (работа a_5).

2. A_1 - a_4-A_4 - A_6;

Некритическая работа a_4. Резерв времени $(a_2 + a_6 + a_8) – a_4 = (50 + 29 + 6) -42 = 43$ мин. Это время может быть использовано для технического обслуживания самолёта (работа a_4).

3. A_6-a_7-A_{78} -a_9;

Некритическая работа a_7 . Резерв времени $(a_8 + a_9) - a_9 = (6 + 10) - 15 = 1$ мин. Этот резерв может быть использован для посадки пассажиров в самолёт (работа a_7).

СОДЕРЖАНИЕ