

1. ОСНОВЫ АЛГОРИТМИЧЕСКОГО СТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ

1.1. ПОНЯТИЕ МОДЕЛИ

В настоящее время нельзя назвать область человеческой деятельности, в которой в той или иной степени не использовались бы методы моделирования. Между тем общепризнанного определения понятия модели не существует.

На наш взгляд, заслуживает предпочтения следующее определение:

модель - объект любой природы, который создается исследователем с целью получения новых знаний об объекте-оригинале и отражает только существенные (с точки зрения разработчика) свойства оригинала.

Анализируя содержание этого определения, можно сделать следующие выводы:

- 1) любая модель субъективна, она несет на себе печать индивидуальности исследователя;
- 2) любая модель **гомоморфна**, т. е. в ней отражаются не все, а только существенные свойства объекта-оригинала;
- 3) возможно существование множества моделей одного и того же объекта-оригинала, отличающихся целями исследования и степенью адекватности.

Модель считается **адекватной** объекту-оригиналу, если она с достаточной степенью приближения на уровне понимания моделируемого процесса исследователем отражает закономерности процесса функционирования реальной системы во внешней среде.

1.2. КЛАССИФИКАЦИЯ МОДЕЛЕЙ

По форме представления объектов модели можно разделить на две группы: *материальные* и *идеальные*. Материальные модели, в свою очередь, делятся на *физические* и *аналоговые** В *физических моделях* обеспечивается аналогия физической природы и модели (примером может служить аэродинамическая труба). В *аналоговых моделях* добиваются сходства процессов, протекающих в оригинале и модели (так с помощью гидроинтегратора моделируется передача тепла).

Идеальные модели можно разделить на *знаковые* (семиотические) и *интуитивные* (мысленные). *Знаковые модели* делятся на логические, геометрические и математические.

Математические модели можно разделить на *аналитические*, *алгоритмические* и *комбинированные*.

Для *аналитического* моделирования характерно то, что для описания процессов функционирования системы используются системы алгебраических, дифференциальных, интегральных или конечно-разностных уравнений. Аналитическая модель может быть исследована следующими методами:

- а) аналитическим, когда стремятся получить в общем виде явные зависимости

для искомых характеристик;

б) численным, когда, не умея решать уравнения в общем виде, стремятся получить числовые результаты при конкретных начальных данных;

в) качественным, когда, не имея решения в явном виде, можно найти некоторые свойства решения (например, оценить устойчивость решения).

Желая использовать аналитический метод, часто идут на существенные упрощения первоначальной модели, чтобы иметь возможность изучить хотя бы общие свойства системы. Аналитические модели бывают *детерминированные* и *статистические*. Численный метод проведения аналитических расчетов с помощью датчиков случайных чисел получил название метода статистических испытаний, или метода Монте-Карло. При *алгоритмическом моделировании* описывается процесс функционирования системы во времени, причем имитируются элементарные явления, составляющие процесс, с сохранением их логической структуры и последовательности протекания во времени. *Алгоритмические* модели также могут быть детерминированными и статистическими. В последнем случае в модели с помощью датчиков случайных чисел имитируется действие неопределенных и случайных факторов. Такой метод моделирования получил название метода статистического моделирования. В настоящее время этот метод считается наиболее эффективным методом исследования сложных систем, а часто и единственным практически доступным методом получения информации о поведении гипотетической системы на этапе ее проектирования.

Комбинированное моделирование позволяет объединить достоинства аналитического и алгоритмического моделирования. При построении комбинированных моделей производится предварительная декомпозиция процесса функционирования модели на составляющие подпроцессы. Для тех из них, где это возможно, используются аналитические модели, а для остальных процессов строятся алгоритмические модели.

1.3. ПОСЛЕДОВАТЕЛЬНОСТЬ РАЗРАБОТКИ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ

В процесс разработки и машинной реализации математической модели входят следующие этапы:

- построение концептуальной модели;
- разработка алгоритма модели системы;
- разработка программы модели системы;
- проведение машинных экспериментов с моделью системы.

1.3.1. Построение концептуальной модели

Построение концептуальной модели включает следующие подэтапы:

- постановку задачи моделирования;
- определение требований к исходной информации и ее сбор;
- выдвижение гипотез и предположений;
- определение параметров и переменных модели;
- обоснование выбора показателей и критериев эффективности системы;
- составление содержательного описания модели.

При постановке задачи моделирования дается четкая формулировка целей и задач исследования реальной системы, обосновывается необходимость машинного моделирования, выбирается методика решения задачи с учетом имеющихся ресурсов, определяется возможность разделения задачи на подзадачи.

При сборе необходимой исходной информации необходимо помнить, что именно от качества исходной информации об объекте моделирования зависит как адекватность модели, так и достоверность результатов моделирования.

Гипотезы при построении модели системы служат для заполнения «пробелов» в понимании задачи исследователем. Предположения дают возможность провести упрощение модели. В процессе работы с моделью системы возможно многократное возвращение к этому подэтапу в зависимости от полученных результатов моделирования и новой информации об объекте.

При определении параметров и переменных составляется перечень входных, выходных и управляющих переменных, а также внешних и внутренних параметров системы.

Выбранные показатели и критерии эффективности системы должны отражать цель функционирования системы и представлять собой функции переменных и параметров системы.

Разработка концептуальной модели завершается составлением содержательного описания, которое используется как основной документ, характеризующий результаты работы на первом этапе.

1.3.2. Разработка алгоритма модели

Разработка алгоритма модели включает следующие подэтапы:

- построение логической схемы алгоритма;
- получение математических соотношений;
- проверку достоверности алгоритма.

Вначале создается укрупненная (обобщенная) схема моделирующего алгоритма, которая задает общий порядок действий при моделировании исследуемого процесса. Затем разрабатывается детальная схема, каждый элемент которой впоследствии превращается в оператор программы.

Для комбинированных моделей разрабатывается аналитическая часть в виде явных функций и имитационная часть в виде моделирующего алгоритма.

Проверка достоверности алгоритма должна дать ответ на вопрос, насколько алгоритм отражает замысел моделирования, сформулированный на этапе разработки концептуальной модели.

1.3.3. Разработка программы

Разработка программы для ЭВМ включает следующие подэтапы:

- выбор вычислительных средств;
- проведение программирования;
- проверку достоверности программы.

Прежде всего выбираются тип ЭВМ (компьютера) и язык программирования. Создание программы по детально разработанному алгоритму может осуществить программист без участия и помощи разработчика модели.

После составления программы производится проверка ее достоверности на контрольном примере. На этом подэтапе необходимо оценить затраты машинного времени для расчета одной реализации моделируемого процесса, что позволит разработчику модели правильно сформулировать требования к точности и достоверности результатов моделирования.

1.3.4. Проведение машинных экспериментов с моделью системы

На этом этапе проводятся серийные расчеты по составленной и отлаженной программе. Этап включает следующие подэтапы:

- планирование машинного эксперимента;
- проведение рабочих расчетов;
- представление результатов моделирования;
- интерпретацию результатов моделирования;
- выдачу рекомендаций по оптимизации режима работы реальной системы.

Перед проведением рабочих расчетов на ЭВМ должен быть составлен план проведения эксперимента с указанием комбинаций переменных и параметров, для которых должно проводиться моделирование системы. Задача заключается в разработке оптимального плана эксперимента, реализация которого позволяет при сравнительно небольшом числе испытаний модели получить достоверные данные о закономерностях функционирования системы.

Результаты моделирования могут быть представлены в виде таблиц, графиков, диаграмм, схем и т. п. В большинстве случаев наиболее простой формой считаются таблицы, хотя графики более наглядно иллюстрируют результаты моделирования системы.

Целесообразно предусмотреть вывод результатов на экран дисплея и на принтер.

Интерпретация результатов моделирования имеет целью переход от информации, полученной в результате машинного эксперимента с моделью, к выводам, касающимся процесса функционирования объекта-оригинала.

На основании анализа результатов моделирования принимается решение о том, при каких условиях система будет функционировать с наибольшей эффективностью.

1.4. ТИПОВЫЕ МАТЕМАТИЧЕСКИЕ СХЕМЫ

В процессе создания математической модели, реализуемой на ЭВМ, происходит переход от содержательного описания к формальному алгоритму. Промежуточным звеном между ними может служить *математическая схема*.

Существует ряд типовых математических схем, которые могут лечь в основу разрабатываемого конкретного моделирующего алгоритма.

К ним относятся следующие схемы (модели):

- непрерывно-детерминированные модели (D-схемы);
- дискретно-детерминированные модели (F-схемы);
- дискретно-стохастические модели (P-схемы);
- непрерывно-стохастические модели (Q-схемы).

К *непрерывно-детерминированным моделям* относятся модели, описываемые системами обыкновенных дифференциальных уравнений или уравнений в частных производных. В качестве независимой переменной, от которой зависят неизвестные искомые функции, обычно служит время. Тогда вектор-функция искомых переменных будет непрерывной. Математические схемы такого вида отражают динамику изучаемой системы и поэтому называются D-схемами (англ. dynamic).

К *дискретно-детерминированным моделям* относятся так называемые конечные автоматы. Автомат можно представить как некоторое устройство, на которое подаются входные сигналы и снимаются выходные и которое может иметь некоторые внутренние состояния. У конечного автомата множество входных сигналов и внутренних состояний является конечным множеством. Название F-схема происходит от английских слов finite automata.

К *дискретно-стохастическим моделям* относятся вероятностные (стохастические) автоматы или по-английски probabilistic automat. Отсюда название - P-схема. В общем виде вероятностный автомат можно определить как дискретный потактный преобразователь информации с памятью, функционирование которого в каждом такте зависит только от состояния памяти в нем и может быть описано стохастически.

Примером типовой схемы *непрерывно-стохастического типа* может служить схема

системы массового обслуживания (СМО) или по-английски queueing system. Отсюда название -Q-схема.

В качестве процесса обслуживания в СМО могут быть представлены различные по физической природе процессы функционирования экономических, производственных, технических и других систем, например потоки товаров, потоки продукции, потоки деталей, потоки клиентов и т. п.

Для любой системы массового обслуживания характерно наличие трех отличительных свойств:

- объектов, у которых может возникнуть потребность в удовлетворении некоторых заявок;
- агрегатов, предназначенных для удовлетворения заявок на обслуживание;
- специальной организации приема в систему заявок и их обслуживания.

Схема системы массового обслуживания показана на рис. 1.1.



Рис. 1.1. Схема системы массового обслуживания

Совокупность заявок рассматривают как *поток событий*, т. е. последовательность событий, происходящих в случайные моменты времени. Время обслуживания заявки также считается случайной величиной.

Из-за совместного действия этих двух случайных факторов количество обслуженных заявок в заданном интервале времени является величиной случайной.

Исследование моделей СМО ставит целью установление параметров случайных величин, характеризующих процесс обслуживания заявок.

Существует несколько разновидностей СМО:

- 1) по числу каналов обслуживания СМО делятся на одноканальные и многоканальные?
- 2) по числу фаз (последовательно соединенных агрегатов) СМО делятся на однофазные и многофазные;
- 3) по наличию обратной связи СМО делятся на разомкнутые (с бесконечным числом заявок) и замкнутые (с конечным числом заявок);
- 4) по наличию очереди СМО делятся на системы без очередей (с потерями заявок), системы с неограниченным ожиданием (по времени или длине очереди) и системы с

ограниченным ожиданием (по времени или длине очереди);

5) по принципу формирования очередей СМО делятся на системы с общей очередью и системы с несколькими очередями;

6) по наличию отказов СМО делятся на системы с отказами и системы без отказов;

7) по виду приоритета СМО делятся на системы со статическим приоритетом (обслуживание в порядке поступления заявок) и системы с динамическим приоритетом, который, в свою очередь, имеет три разновидности:

- **относительный приоритет** (заявка высокого приоритета ожидает окончания обслуживания заявки с более низким приоритетом);

- **абсолютный приоритет** (заявка высокого приоритета при поступлении немедленно вытесняет заявку с более низким приоритетом);

- **смешанный приоритет** (если заявка с низшим приоритетом обслуживалась в течение времени, меньше критического, то используется абсолютный приоритет, в противном случае используется относительный приоритет).

1.5. ДАТЧИКИ СЛУЧАЙНЫХ ЧИСЕЛ С РАВНОМЕРНЫМ РАСПРЕДЕЛЕНИЕМ

Алгоритмическое моделирование - это численный метод исследования систем и процессов с помощью моделирующего алгоритма.

Каждый раз, когда на ход моделируемого процесса оказывает влияние случайный фактор, его действие имитируется с помощью специально организованного розыгрыша (жеребья). Таким образом строится одна случайная реализация моделируемого

явления, представляющая собой как бы один результат опыта. По одному опыту, конечно, нельзя судить о закономерностях изучаемого процесса. Но при большом числе реализаций средние характеристики, вырабатываемые моделью, приобретают свойство устойчивости, которое усиливается с увеличением числа реализаций.

Бросание жребия можно осуществить вручную (выбором из таблицы случайных чисел), но удобнее это делать с помощью специальных программ, входящих в состав программного обеспечения ЭВМ. Такие программы называют **датчиками** или **генераторами** случайных чисел.

В трансляторах почти всех алгоритмических языков имеются стандартные процедуры или функции, которые генерируют случайные (точнее, псевдослучайные) величины с равномерным распределением.

В трансляторе языка Visual Basic имеется стандартная функция *RND*, (в Delphi – Random), возвращающая случайные вещественные числа одинарной точности в интервале

(0,1).

Обращение к этой функции может иметь вид: $z = RND$, где z - возможное значение случайной величины, равномерно распределенной в интервале (0,1).

1.6. МОДЕЛИРОВАНИЕ СЛУЧАЙНЫХ СОБЫТИЙ

1.6.1. Моделирование простого события

Пусть имеется событие A , вероятность наступления которого равна P_A . Требуется выработать правило, при многократном использовании которого частота появления события стремилась бы к его вероятности. Выберем с помощью датчика случайных чисел, равномерно распределенных в интервале (0,1) некоторое число z и определим вероятность того, что $z < P_A$. Для случайной величины z с равномерным распределением справедлива следующая зависимость:

$$P(z < P_A) = \int_0^{P_A} f(x)dx = P_A.$$

Таким образом, вероятность попадания случайной величины в интервал $(0, P_A)$ равна величине P_A . Поэтому если при розыгрыше число z попало в этот интервал, то следует считать, что событие A произошло. Противоположное событие (не A) произойдет с вероятностью $(1 - P_A)$ в том случае, если $z > P_A$.

Процедура моделирования простого события в имитационной модели описывается алгоритмом, схема которого показана на рис. 1.2.

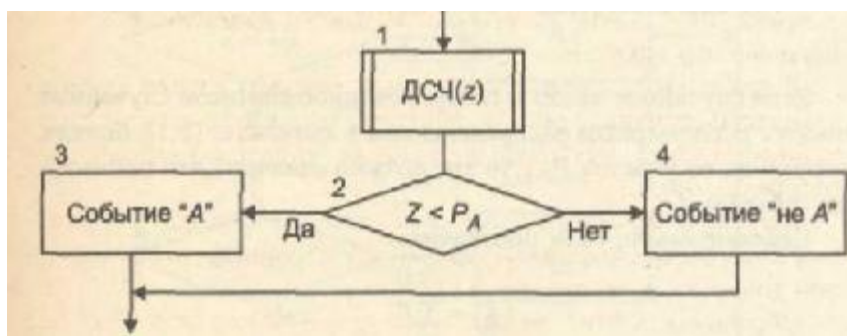


Рис. 1.2. Моделирование простого события

Оператор 1 обращается к датчику случайных чисел, генерирующему случайную величину z . Оператор 2 проверяет условие $z < P_A$. Если оно выполняется, считается, что произошло событие A . В противном случае считается, что произошло противоположное событие (не A).

1.6.2. Моделирование полной группы несовместных событий

Пусть имеется полная группа несовместных событий (ПГНС) A_1, A_2, \dots, A_K с

$$\sum_{i=1}^k P_i = 1$$

вероятностями P_1, P_2, \dots, P_k . При этом выполняется условие:

Разделим интервал $(0, 1)$ на k отрезков, длины которых составляют $P_1; P_2; \dots, P_k$ (рис. 1.3).

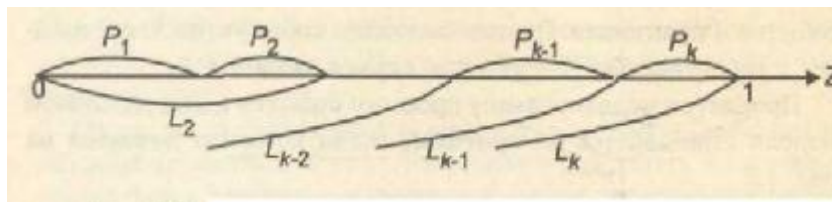


Рис.1.3. Моделирование полной группы несовместных событий

Если случайное число z , генерированное датчиком случайных чисел с равномерным распределением в интервале $(0,1)$, попало, например, на участок P_{k-1} то это должно означать, что произошло событие A_{k-1} .

Действительно, если обозначить

$$L_j = \sum_{i=1}^j P_i,$$

то окажется справедливым выражение:

$$P(L_{k-2} < z < L_{k-1}) = \int_{L_{k-2}}^{L_{k-1}} 1 \cdot dz = P_{k-1}.$$

Следовательно, произойдет событие, которое имеет вероятность P_{k-1}

Процедура моделирования полной группы несовместных событий описывается алгоритмом, схема которого показана на рис. 1.4.

Оператор 1 обращается к датчику случайных чисел с равномерным распределением в интервале $(0,1)$. Условный оператор 2 проверяет условие попадания случайной величины z в интервал $(0, L_1)$. Если это условие выполняется, то считается, что произошло событие A_1 . Если условие в операторе 2 не выполняется, то алгоритм осуществляет проверку условий попадания случайной величины в другие интервалы. Одно из событий A_1, A_2, \dots, A_k обязательно произойдет.

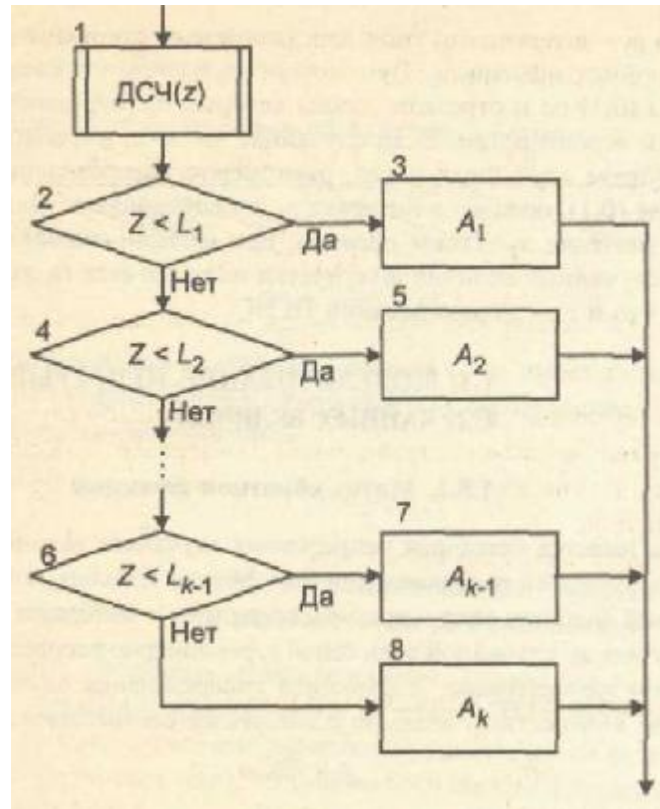


Рис. 1.4. Схема алгоритма моделирования ПГНС

1.7. МОДЕЛИРОВАНИЕ ДИСКРЕТНОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Дискретная случайная величина может быть задана табличной зависимостью:

X	x_1	x_2	\dots	x_n
p	P_1	P_2	\dots	P_n

Здесь P_j — вероятность того, что дискретная случайная величина X примет значение x_j . При этом $p_1 + p_2 + \dots + p_n = 1$. Разделим интервал $(0,1)$ на n отрезков, длины которых пропорциональны заданным вероятностям. Если случайное число z , вырабатываемое датчиком случайных чисел, равномерно распределенных в интервале $(0,1)$, попадет в интервал p_k то случайная величина X примет значение x_k . Таким образом, при моделировании дискретных случайных величин фактически используется та же процедура, что и при моделировании ПГНС.

1.8. МОДЕЛИРОВАНИЕ НЕПРЕРЫВНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

1.8.1. Метод обратной функции

Пусть имеется некоторая непрерывная случайная величина x , заданная функцией распределения $F(x)$. Можно доказать, что значения этой функции равномерно распределены в интервале $(0,1)$. Поэтому между случайной величиной z , равномерно распределенной в том же интервале, и функцией распределения случайной величины x существует взаимно

однозначное соответствие, т. е.

$$z = F(x). \quad (1.1)$$

Отсюда следует, что

$$x = F^{-1}(z), \quad (1.2)$$

где F^{-1} - обратная функция.

Следовательно, если уравнение (1.1) имеет аналитическое решение, то для моделирования случайной величины x можно использовать датчик случайных чисел, генерирующий величину z , и затем осуществить расчет по формуле (1.2).

Основы вероятностных методов анализа и моделирования систем

1.1. Элементарные понятия о случайных событиях, величинах и функциях

Под событием понимается всякий факт, который может произойти в данных условиях. Теория вероятностей рассматривает события в тесной связи с теми условиями, в которых они наступают. Совокупность условий, в которых рассматривается данное событие, называют **комплексом условий**, а реализацию этого комплекса условий на практике — **испытанием**. В зависимости от связи между событиями и соответствующими комплексами условий различают достоверные, невозможные и случайные события.

Достоверным называется такое событие, которое наступает каждый раз при реализации данного комплекса условий. Достоверное событие будем обозначать через U .

Невозможным называется событие, которое никогда не наступает при реализации данного комплекса условий. Невозможное, событие будем обозначать символом \downarrow .

Случайным называется событие, которое может либо наступить при реализации данного комплекса условий, либо не наступить. Достоверное и невозможное события могут рассматриваться как крайние частные случаи случайных событий. Случайные события будем обозначать через $A, B, C...$

Согласно теоретико-множественному подходу при рассмотрении понятия «случайное событие» вводится понятие «элементарное событие». **Элементарное событие** — это один из нескольких возможных, но несовместных исходов того или иного опыта (испытания). Совокупность или множество их составляют пространство **элементарных** событий.

В общем случае пространство элементарных событий может быть любой природы: конечным и бесконечным, дискретным и непрерывным. Пространство элементарных событий является синонимом достоверного события, так как один из его элементов непременно наступит.

Всякому событию при данном комплексе условий соответствует определенная степень возможности. Более возможные события при многократных испытаниях в среднем наступают чаще, а менее возможные — реже. **Частотой события** называется отношение

числа испытаний, в которых появилось данное событие, и общего числа испытаний. Частота события A равна:

$$P^*(A) = \frac{m(A)}{n} \quad (1.6)$$

где n - общее число проведенных испытаний;

$m(A)$ — число испытаний, в которых наступило событие A .

Частота достоверного события U равна единице:

$$P^*(U) = \frac{n}{n} = 1.$$

Частота невозможного события равна нулю:

$$P^*(0) = \frac{0}{n} = 0.$$

Частота случайного события A находится в интервале $[0;1]$: $0 < P^*(A) < 1$.

Следует отметить, что частота случайного события обладает устойчивостью, что доказывается и формулируется в *теореме Я. Бернулли*, относящейся к закону больших чисел.

Свойство устойчивости частоты случайного события отражает связь между комплексом условий и возможностью наступления событий при данном комплексе. Количественной мерой степени возможности появления события для заданного комплекса условий является *вероятность события*. Чем более возможно появление случайного события, тем больше его вероятность. Наоборот, чем менее возможно появление события, тем меньше его вероятность.

Вероятность и частота события тесно связаны между собой. Зная частоту, вычисленную при достаточно большом числе испытаний, есть все основания считать ее близкой к соответствующей вероятности и полагать, что

$$P(A) = P^*(A) \cong \frac{m(A)}{n}.$$

Такой способ определения вероятности события $P(A)$ называется **статистическим**.

Случайные события могут быть представлены через случайные величины. Понятие «случайная величина» расширяет область применения вероятностных методов в решении практических задач, позволяет исследовать более сложные случайные явления. **Случайной** называется такая величина, которая в результате испытания (реализации определенного комплекса условий) может принять то или иное значение, причем до испытания неизвестно, какое именно. Если повторять испытания, то результатом каждого будет какое-либо одно значение случайной величины из множества возможных.

Случайные величины подразделяются на дискретные и непрерывные.

Множество значений **дискретной случайной величины** конечно или счетно, например:

- количество отказов автомобилей автопредприятия в течение рабочей смены;
- число рабочих, пришедших в бухгалтерию завода в течение одного часа получать заработную плату, и т. д.

Множество значений **непрерывной случайной величины** представляет собой множество всех точек, принадлежащих какому-либо интервалу числовой оси, например:

- расход топлива на километр пробега;
- время безотказной работы автомобиля;
- время обслуживания заявки в очереди.

Кроме дискретной и непрерывной случайных величин встречаются **случайные величины смешанного типа**, для которых наряду с участками непрерывных значений имеются отдельные, дискретные значения.

Для того чтобы задать случайную величину, прежде всего необходимо задать множество значений, которые она может принимать. Однако одного перечня значений случайной величины еще недостаточно для каких-либо существенных выводов. Нужно еще знать, как часто, т. е. с какой вероятностью, она принимает эти значения. Ответ на поставленный вопрос дает исчерпывающая характеристика случайной величины — закон ее распределения.

Закон распределения представляет собой соотношение, позволяющее определить вероятность появления случайной величины в любом интервале (и, в частности, вероятности любых значений случайной величины).

Основными формами закона распределения являются: **ряд распределения, функция распределения и плотность распределения**.

Ряд распределения представляет собой таблицу, в которой перечислены возможные значения случайной величины и соответствующие им вероятности:

X	x_1	x_2	\dots	x_n
p	P_1	P_2	\dots	P_n

В таблице x_i — i -е значение случайной величины X ; p_i — вероятность появления i -го значения случайной величины X . При этом

$$\sum_{i=1}^n P_i = 1.$$

Эмпирический ряд распределения представляет собой таблицу, в которой перечислены наблюдаемые значения (фактические реализации) случайной величины и соответствующие им частоты:

x_i	x_1	x_2	\dots	x_n	x_i
m_i	m_1	m_2	\dots	m_n	m_i

В таблице x_i — i -я фактическая (наблюдаемая) реализация случайной величины X ; m_i — количество появлений (частота) величины x_i

Ряд распределения не может служить характеристикой непрерывной случайной величины, поскольку значения этой **величины** нельзя перечислить, так как множество их несчетно. Кроме того, вероятность отдельного значения непрерывной случайной величины равна нулю.

Для характеристики непрерывной случайной величины определяют вероятность появления значения случайной величины меньшего x , где x — текущая переменная, т. е. определяют вероятность события $X < x$. Вероятность этого события зависит от x , т. е. является функцией x . Эта функция называется **функцией распределения случайной величины** X и обозначается $F(x)$: Таким образом, функцией распределения случайной величины X называется функция аргумента x , равная вероятности того, что случайная величина X примет любое значение, меньшее x .

Вероятность попадания случайной величины в полузамкнутый интервал $[a, b)$ равна разности значений функции распределения в точках b и a :

$$P(a \leq x < b) = F(b) - F(a). \quad (1.18)$$

Функция распределения есть неубывающая функция, значения которой начинаются с нуля и доходят до единицы, причем в отдельных случаях функция может иметь скачки — разрывы. Функцию распределения дискретной случайной величины можно определить, зная ее ряд распределения, по формуле:

$$F(x) = \sum_{x_i < x} P(x_i), \quad (1.19)$$

где суммирование распространяется на значения x_i которые меньше x .

Следует отметить, что функция распределения дискретной случайной величины увеличивается скачками каждый раз, когда X при своем изменении проходит через какое-нибудь из возможных значений x_i , причем величина скачка равна вероятности этого значения. Между двумя соседними значениями величины X функция $F(x)$ постоянна.

Поскольку для непрерывной случайной величины нельзя использовать в качестве

характеристики вероятности появления ее отдельных значений, то определяют вероятность появления случайной величины в пределах малого интервала $[x, x + \Delta x]$, примыкающего к x . Разделив эту вероятность на длину интервала Δx , находят среднюю плотность вероятности и при неограниченном уменьшении длины интервала переходят к пределу, который является плотностью распределения в точке x :

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x \leq X < x + \Delta x)}{\Delta x}.$$

Плотность распределения $f(x)$ есть предел отношения вероятности попадания случайной величины на малый участок и длины этого участка при ее неограниченном уменьшении.

Вероятность попадания случайной величины на произвольный участок $[a, b]$ равна:

$$P(a \leq X < b) = \int_a^b f(x) \cdot dx.$$

Интеграл в бесконечных пределах от плотности распределения равен единице, т. е.

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

График плотности распределения называется кривой распределения, лежащей в верхней полуплоскости. Кривая распределения совместно с осью абсцисс ограничивает площадь, равную единице (рис. 1.1).

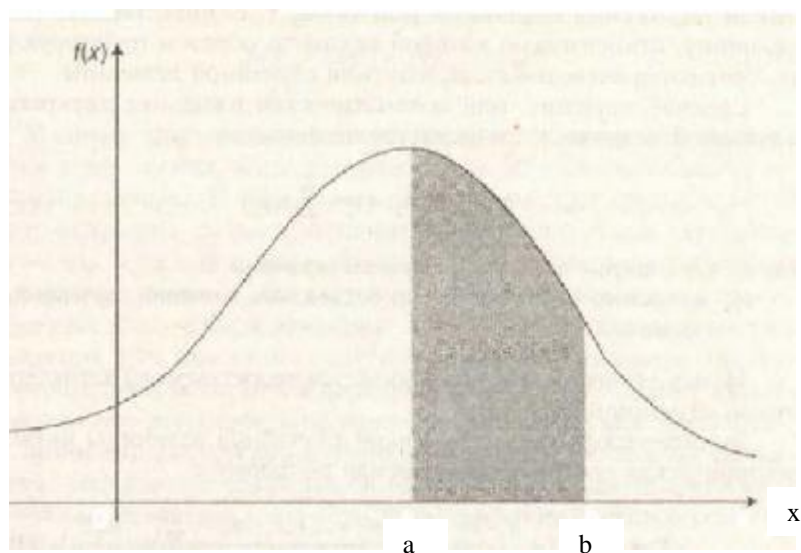


Рис. 1.1. График плотности распределения (кривая распределения)

Вероятность попадания на участок $[a, b)$ равна площади ограниченной кривой распределения, опирающейся на участок $[a, b)$ (на рис. 1.1, — заштрихованная площадь).

Плотность распределения есть производная функции распределения. С другой стороны:

$$F(x) = P(X < x) = P(-\infty < X < x),$$

откуда

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

Величину $F(x)$ называют интегральной функцией распределения величины X . Величина $f(x)$ — дифференциальная функция распределения случайной величины X .

1.2. Числовые характеристики случайных величин

При решении многих практических задач часто достаточно указать отдельные числовые характеристики, определяющие особенности того или иного распределения случайной величины. Это прежде всего среднее значение, которое принадлежит к характеристикам положения случайной величины, т. е. представляет такую величину, относительно которой каким-то образом группируются, рассеиваются всевозможные значения случайной величины.

Среднее значение, или **математическое ожидание дискретной случайной величины**, вычисляется по формуле

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

где x_i , — возможные значения случайной величины X ;

p_i — вероятность появления i -го возможного значения случайной величины X .

Математическое ожидание является теоретической характеристикой случайной величины.

Эмпирической характеристикой случайной величины является эмпирическая средняя, вычисляемая по формуле

$$\bar{X} = \sum_{i=1}^n x_i \cdot \frac{m_i}{N}$$

где $\frac{m_i}{N}$ - частота значений x_i , - при N наблюдениях (испытаниях); $N = \sum_{i=1}^n m_i$.

m_i — количество появлений значений x_i , - при N наблюдениях.

Эмпирическая средняя случайной величины по мере увеличения испытаний (наблюдений) приобретает тенденцию стабилизироваться относительно постоянной величины — математического ожидания.

Для **непрерывной случайной** величины X математическое ожидание определяется интегралом:

$$M[X] = \int_{-\infty}^{+\infty} x \cdot f(x) \cdot dx.$$

(Кроме математического ожидания на практике иногда применяются и другие характеристики положения, в частности медиана и мода случайной величины.

Медианой Me случайной величины называется такая величина, относительно которой равновероятно получение большего или меньшего значения случайной величины:

$$P(X > Me) = P\{X < Me\}.$$

Медиану применяют в качестве характеристики ряда распределения в тех случаях, когда имеются очень большие колебания случайной величины. В этом случае на эмпирическую среднюю $M^*[X]$ будут оказывать сильное влияние крайние значения случайной величины, а медиана менее чувствительна к крайним значениям случайной величины. На медиану влияет не столько колебание в значениях случайной величины X , сколько колебания в частоте появления того или иного значения случайной величины. Медиану необходимо вычислять в дополнение к математическому ожиданию в случае распределений, имеющих большую скошенность¹.

Модой Mo дискретной случайной величины называется ее значение, обладающее наибольшей вероятностью. Для непрерывной случайной величины мода есть такое значение, которое отвечает максимальной плотности распределения.

В общем случае математическое ожидание, медиана и мода не совпадают. В частном случае при симметричном распределении все три характеристики положения случайной величины совпадают.)

Для оценки степени разброса, рассеивания значений случайной величины относительно среднего вычисляют следующие характеристики:

- дисперсию;
- среднее квадратическое отклонение;
- коэффициент вариации.

Дисперсией называется математическое ожидание квадрата отклонений случайной

величины от своего математического ожидания:

$$D_x = \sigma_x^2 = M[(X - m_x)^2]$$

Чем больше дисперсия, тем в среднем больше отклонение значений случайной величины относительно математического ожидания, т. е. будет больше рассеивание случайной величины.

Дисперсия дискретной случайной величины вычисляется по формуле:

$$\sigma_x^2 = \sum_{i=1}^n (x_i - m_x)^2 \cdot p(x_i).$$

Дисперсия непрерывной случайной величины равна:

$$\sigma_x^2 = \int (x - m_x)^2 \cdot f(x) \cdot dx$$

Наряду с дисперсией случайной величины, в качестве характеристики рассеивания случайной величины используется среднее квадратическое отклонение, которое равно положительному значению корня квадратного из дисперсии.

Среднее квадратическое отклонение имеет одинаковую размерность со случайной величиной, в этом состоит ее преимущество относительно дисперсии.

Эмпирические значения характеристик рассеивания вычисляют по формулам:

- дисперсия:

$$\sigma_u^2 = \sum (x_i - \bar{x})^2 \cdot \frac{m_i}{N};$$

- среднее квадратическое отклонение:

$$\sigma_x = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \frac{m_i}{N}}$$

Для малых выборок, если число испытаний (наблюдений) $N \leq 30$, то характеристики рассеивания вычисляются по формулам:

- дисперсия:

$$\sigma_u^2 = \sum (x_i - \bar{x})^2 \cdot \frac{m_i}{N-1};$$

- среднее квадратическое отклонение:

$$\sigma_x = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \frac{m_i}{N-1}}$$

Величины σ_x^2 и σ_x показывают абсолютное отклонение от среднего значения случайной величины, что недостаточно характеризует уровень ее рассеивания. Относительной характеристикой рассеивания является коэффициент вариации, вычисляемый как отношение среднего квадратического отклонения и эмпирической средней:

$$V = \frac{\sigma_x}{\bar{x}} \cdot 100\%$$

или

$$V = \frac{\sigma_x}{\bar{x}}.$$

Коэффициент вариации может использоваться для сравнения меры рассеивания случайных величин, имеющих различную размерность.

Выбор теоретического закона распределения случайной величины

В любом статистическом распределении присутствуют элементы случайности, и, как следствие, экспериментальные точки гистограммы распределения обычно колеблются от опыта к опыту около неизвестной кривой истинного распределения.

При наличии числовых характеристик случайной величины {математического ожидания, дисперсии, коэффициента вариации) законы ее распределения могут быть определены в первом приближении по таблице.

Законы распределения случайной положительной величины в зависимости от коэффициента вариации

Таблица

Пределы изменения коэффициента вариации V	Закон распределения случайной величины X
$V \leq 0,3$	Нормальный
$0,3 < V < 0,4$	Гамма-распределение
$0,4 \geq < 1$	Вейбулла
$V=1$	Экспоненциальный, Пуассона

Для более точного определения теоретического закона распределения проводят дополнительную статистическую обработку данных. При обработке статистических данных решают вопрос о том, как подобрать для исходного статистического ряда теоретическую

кривую распределения, которая выражала бы лишь существенные черты статистического материала, но не случайности, обусловленные недостаточным объемом выборки экспериментальных данных. Под построением теоретической кривой распределения понимается такая обработка статистических данных, когда обеспечивается подбор наиболее подходящего теоретического закона распределения, который может быть задан либо функцией распределения $F(x)$, либо плотностью распределения $f(x)$.

Моделирование случайных величин

Для моделирования случайной величины необходимо знать закон её распределения.

От последовательности случайных чисел, равномерно распределенных в интервале $[0,1]$, нетрудно перейти к последовательности случайных чисел с произвольным заданным законом распределения. Равномерно распределенные в интервале $[0,1]$ последовательности случайных чисел можно получить, например, используя функцию `random` в Delphi

Существует основное соотношение, связывающее случайные числа с заданным законом распределения и случайные числа с равномерным законом распределения в интервале $[0,1]$. Суть его состоит в том, что для преобразования последовательности случайных чисел с равномерным законом распределения в интервале $[0,1]$ в последовательность случайных чисел с заданной функцией распределения $F(x)$ необходимо из совокупности случайных чисел с равномерным законом распределения в интервале $[0,1]$ выбрать случайное число z и решить уравнение:

$$F(x) = z$$

относительно x .

Решение уравнения представляет собой случайное число из совокупности случайных чисел, имеющих функцию распределения $F(x)$.

В случае когда вместо функции распределения $F(x)$ задана плотность вероятности $f(x)$, необходимо решить уравнение:

$$\int_{-\infty}^x f(x)dx = z.$$

Для ряда законов распределения, наиболее часто встречающихся в реальной жизни, получено аналитическое решение последнего уравнения.

Моделирование случайных величин с показательным законом распределения

Непрерывная случайная величина X имеет показательное распределение, если ее

плотность распределения выражается формулой:

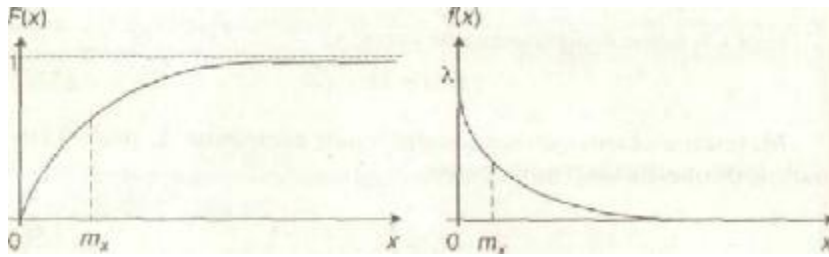
$$f(x) = \lambda \cdot e^{-\lambda x}, \quad x > 0.$$

Положительная величина λ является параметром показательного распределения.

Функция распределения случайной величины X выглядит следующим образом:

$$F(x) = 1 - e^{-\lambda x}, \quad \lambda > 0, \quad 0 \leq x < \infty.$$

Графики функции и плотности показательного распределения приведены на рис.



а) функция показательного распределения б) плотность показательного распределения

Рис.. Графики показательного распределения

Математическое ожидание случайной величины X , имеющей показательное распределение, обратно его параметру, т. е.

$$m_x = \frac{1}{\lambda}.$$

Дисперсия случайной величины X имеющей показательное распределение, равна

$$D_x = \frac{1}{\lambda^2}.$$

Отсюда $\sigma_x = \frac{1}{\lambda} = m_x$.

Коэффициент вариации случайной величины X , имеющей показательное распределение, равен 1:

$$V_x = \frac{\sigma_x}{m_x} = 1.$$

Поскольку

$$z = F(x) = 1 - e^{-\lambda x},$$

тогда для случайной величины x можно записать:

$$x = -\frac{1}{\lambda} \ln(1 - z).$$

Поскольку случайная величина $(1-z)$ имеет также равномерное распределение в интервале $(0,1)$, то последнее соотношение можно заменить соотношением

$$x = -\frac{1}{\lambda} \ln(z)$$

Это выражение будем ниже использовать при моделировании времени обслуживания заявок в СМО и времени поступления заявок в очередь.

Моделирование случайных величин с нормальным распределением

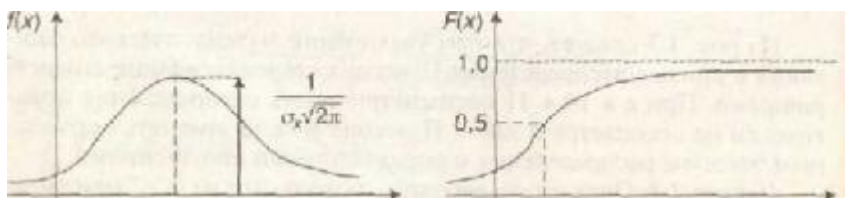
Наиболее известным непрерывным распределением является нормальное. Плотность нормального распределения определяется по формуле:

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}}$$

Непрерывная случайная величина X принимает значения от $-\infty$ до $+\infty$.
Соответствующая функция распределения равна:

$$F(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-m_x)^2}{2\sigma_x^2}} dx.$$

Типичные графики плотности вероятности $f(x)$ и функции нормального распределения приведены на рис.



а) плотность вероятности б) функция нормального распределения

Рис. Графики кривых нормального распределения

Кривой плотности вероятности $f(x)$ нормального распределения является плавная колоколообразная симметричная кривая.

Перечислим основные свойства нормального распределения.

Нормальное распределение полностью характеризуется математическим ожиданием и дисперсией.

Кривая плотности вероятности $f(x)$ нормального распределения симметрична относительно математического ожидания m_x . Максимум плотности распределения соответствует абсциссе, равной m_x .

Метод обратной функции для нормального распределения неприменим, так как после подстановки соответствующей функции распределения выражение для случайной величины не имеет аналитического решения. Поэтому приведем экспериментальный результат для случайной величины с нормальным законом распределения:

$$x = \bar{x} + \sigma \cdot \left(\sum_{i=1}^{12} z_i - 6 \right)$$

где \bar{x} — известное математическое ожидание случайной величины x ;

σ — известное квадратичное отклонение случайной величины x .

Моделирование случайных величин с гамма - распределением

Неотрицательная случайная величина X имеет гамма-распределение, если ее плотность распределения вычисляется по формуле:

$$f_k(x) = \frac{\lambda^k \cdot x^{k-1} \cdot e^{-\lambda x}}{\Gamma(k)} \text{ при } x > 0,$$

где $\lambda > 0$ и $k > 0$;

$\Gamma(k)$ — гамма-функция:

$$\Gamma(k) = \int_0^{\infty} e^{-t} \cdot t^{k-1} \cdot dt,$$

если k — целое неотрицательное число, то

$$\Gamma(k) = k!$$

Математическое ожидание случайной величины X , подчиненной гамма-распределению, равно:

$$m_x = \frac{k}{\lambda}.$$

При этом дисперсия величины X определяется по формуле:

$$D_x = \frac{k}{\lambda^2}.$$

При целом $k > 1$ гамма-распределение превращается в распределение Эрланга k -го порядка, т. е.

$$f_k(x) = \frac{\lambda \cdot (\lambda x)^{k-1} \cdot e^{-\lambda x}}{(k-1)!} \quad (x > 0; k = 1, 2, \dots).$$

Закону Эрланга k -го порядка подчинена сумма независимых случайных величин $x_1 + x_2 + \dots + x_k$, каждая из которых распределена по показательному закону с параметром λ .

При $k = 1$ гамма-распределение превращается в показательное с параметром λ .

Формула для моделирования случайных величин, имеющих гамма-распределение выглядит так:

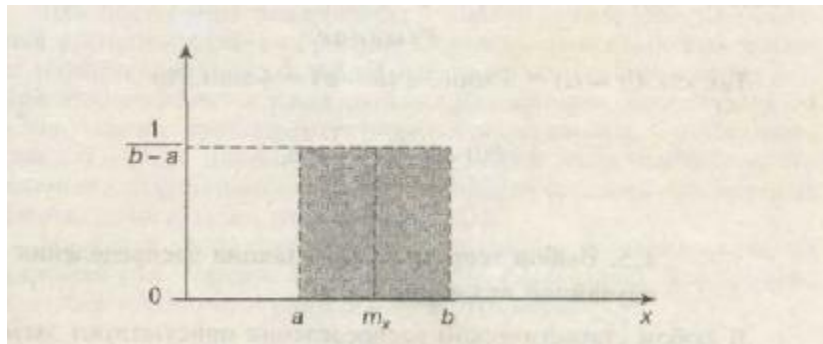
$$x = -\frac{1}{\lambda} \sum_{i=1}^k \ln(1 - z_i), \text{ где } k - \text{целое число.}$$

Моделирование случайных величин с равномерным распределением - распределением

Непрерывная случайная величина X имеет равномерное распределение на отрезке (a, b) , если на этом отрезке плотность распределения постоянна, а вне его — равна нулю.

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

Кривая равномерного распределения показана на рис.



Значения $f(x)$ в крайних точках a и b участка (a, b) не указываются, так как вероятность попадания в любую из этих точек для непрерывной случайной величины X равна нулю.

Кривая равномерного распределения имеет вид прямоугольника,, опирающегося на участок $[a, b]$.

Математическое ожидание случайной величины X , имеющей равномерное распределение на участке $[a, b]$ равно:

$$m_x = \frac{a+b}{2}.$$

Дисперсия случайной величины X , имеющей равномерное распределение на участке $[a, b]$, вычисляется по формуле:

$$D_x = \frac{(b-a)^2}{12}.$$

Отсюда

$$\sigma_x = \frac{b-a}{2\sqrt{3}}.$$

Вероятность попадания равномерно распределенной случайной величины X на участок $[\alpha, \beta]$ выразим формулой:

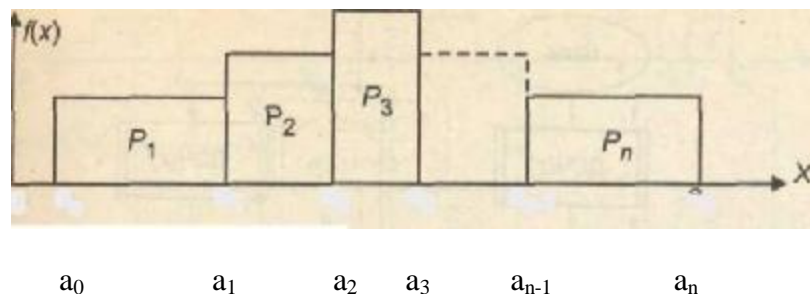
$$P(\alpha < x < \beta) = \frac{\beta - \alpha}{b - a}.$$

Формула для моделирования случайных величин с равномерным распределением

$$x = a + z(b - a).$$

Моделирование случайных величин с произвольным распределением

Пусть случайная величина x : задана в интервале (a_0, a_n) кусочно-постоянной функцией $f(x)$. Это значит, что интервал разбит на n частичных интервалов и плотность распределения $f(x)$ на каждом из них постоянна.



Плотность распределения произвольной функции

Целесообразно выбрать величины a_k так, чтобы вероятности попадания в любой частичный интервал P_k были одинаковы, т. е.

$$\int_{a_{k-1}}^{a_k} f(x)dx = \frac{1}{n} \quad (k = 1, 2, \dots, n).$$

Из условия постоянства функции на каждом частичном интервале следует, что случайная величина x может быть определена по формуле

$$x = a_{k-1} + z(a_k - a_{k-1}) \quad (k = 1, 2, \dots, n). \quad (1.4)$$

где z - возможное значение (реализация) случайной величины, равномерно

распределенной в интервале $(0,1)$; a_{k-1} - левая граница частичного интервала; a_k -

правая граница частичного интервала.

Попадание в любой частичный интервал можно рассматривать как событие, входящее в полную группу несовместных событий. Поэтому процедура моделирования в общем случае состоит в следующем.

С помощью датчика случайных чисел с равномерным распределением, вырабатывающего величину z , моделируют дискретную случайную величину - номер интервала k .

Вторично разыгрывают случайную величину z и определяют возможное значение случайной величины x по формуле (1.4).

1. Теория массового обслуживания. Основные положения.

1.1. Предмет и задачи теории массового обслуживания.

Теория массового обслуживания опирается на теорию вероятностей и математическую статистику.

На первичное развитие теории массового обслуживания оказали особое влияние работы датского ученого А.К. Эрланга (1878-1929).

Теория массового обслуживания - область прикладной математики, занимающаяся анализом процессов в системах производства, обслуживания, управления, в которых однородные события повторяются многократно, например, на предприятиях бытового обслуживания; в системах приема, переработки и передачи информации; автоматических линиях производства и др.

Предметом теории массового обслуживания является установление зависимостей между характером потока заявок, числом каналов обслуживания, производительностью отдельного канала и эффективным обслуживанием с целью нахождения наилучших путей управления этими процессами.

Задача теории массового обслуживания - установить зависимость результирующих показателей работы системы массового обслуживания (вероятности того, что заявка будет обслужена; математического ожидания числа обслуженных заявок и т.д.) от входных показателей (количества каналов в системе, параметров входящего потока заявок и т.д.). Результирующими показателями или интересующими нас характеристиками СМО являются - показатели эффективности СМО, которые описывают способна ли данная система справляться с потоком заявок.

Задачи теории массового обслуживания носят оптимизационный характер и в конечном итоге включают экономический аспект по определению такого варианта системы, при котором будет обеспечен минимум суммарных затрат от ожидания обслуживания, потерь времени и ресурсов на обслуживание и простоев каналов обслуживания.

1.2. Система массового обслуживания.

Система обслуживания считается заданной, если известны:

- 1) поток требований, его характер;
- 2) множество обслуживающих приборов;
- 3) дисциплина обслуживания (совокупность правил, задающих процесс обслуживания).

Каждая СМО состоит из какого-то числа обслуживающих единиц, которые называются каналами обслуживания. В качестве каналов могут фигурировать: линии связи, различные приборы, лица, выполняющие те или иные операции и т.п.

Всякая СМО предназначена для обслуживания какого-то потока заявок, поступающих в какие-то случайные моменты времени. Обслуживание заявок продолжается какое-то случайное время, после чего канал освобождается и готов к приему следующей заявки. Случайный характер потока заявок и времен обслуживания приводит к тому, что в какие-то периоды времени на входе СМО скапливается излишне большое число заявок (они либо становятся в

очередь, либо покидают СМО не обслуженными); в другие же периоды СМО будет работать с недогрузкой или вообще простаивать.

Процесс работы СМО представляет собой случайный процесс с дискретными состояниями и непрерывным временем; состояние СМО меняется скачком в моменты появления каких-то событий (или прихода новой заявки, или окончания обслуживания, или момента, когда заявка, которой надоело ждать, покидает очередь).

1.3. Классификация СМО.

Для облегчения процесса моделирования используют классификацию СМО по различным признакам, для которых пригодны определенные группы методов и моделей теории СМО, упрощающие подбор адекватных математических моделей

1.4. Характеристики СМО.

Перечень характеристик систем массового обслуживания можно представить следующим образом:

- среднее время обслуживания;
- среднее время ожидания в очереди;
- среднее время пребывания в СМО;
- средняя длина очереди;
- среднее число заявок в СМО;
- количество каналов обслуживания;
- интенсивность входного потока заявок;
- интенсивность обслуживания;
- интенсивность нагрузки;
- коэффициент нагрузки;
- относительная пропускная способность;
- абсолютная пропускная способность;
- доля времени простоя СМО;
- доля обслуженных заявок;
- доля потерянных заявок;
- среднее число занятых каналов;
- среднее число свободных каналов;
- коэффициент загрузки каналов;
- среднее время простоя каналов.