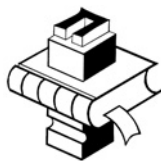


И.А. Гарькина, А.М. Данилов, А.П. Прошин, Ю.А.Соколова

# **ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА. ОБРАБОТКА ОПЫТНЫХ ДАННЫХ**

Под редакцией доктора технических наук,  
профессора А.М. Данилова



**Издательство ПАЛЕОТИП**

Москва 2005

**УДК 519:691**  
**ББК 22.172в6**  
**Г21**

**Рецензенты:**

зав. кафедрой математики и инженерной графики ПАИИ,  
д-р физ.-мат. наук, проф. О.А. Голованов;  
кафедра технологии строительного производства  
Мордовского государственного университета им. Н.П. Огарева  
(зав. кафедрой – член-корр. РААСН,  
д-р техн. наук, проф. В.Д. Черкасов)

**Гарькина, И.А.**

**Г21** Планирование эксперимента. Обработка опытных данных /  
И.А. Гарькина, А.М. Данилов, А.П. Прошин, Ю.А. Соколова ;  
под ред. д-ра техн. наук, проф. А.М. Данилова. – М.: Изда-  
тельство «Палеотип», 2005. – 272 с.

**ISBN 5-94727-117-6**

Настоящее издание содержит теоретические сведения и  
нормативные основы планирования эксперимента и обработки  
опытных данных.

Для аспирантов и научных работников, а также студен-  
тов старших курсов высших технических учебных заведений.

**УДК 519:691**  
**ББК 22.172в6**

**ISBN 5-94727-117-6**

© Колл. авторов, 2005  
© Издательство «Палеотип», 2005

*Когда идет караван, надо иногда останавливаться и ждать, пока осядет пыль, чтобы видеть, куда ты идешь.*

*Арабская пословица*

## **ПРЕДИСЛОВИЕ**

В основе предлагаемого издания лежат лекции, прочитанные в течение ряда лет аспирантам Пензенского государственного университета архитектуры и строительства, обучающимся по специальности 052305 – «Строительные материалы и изделия», в рамках их подготовки к проведению исследовательских работ, планированию эксперимента и обработке опытных данных.

Приводятся основные сведения из курса теории вероятностей и математической статистики, а также корреляционного анализа, включая ранговую корреляцию.

Излагаются элементы теории функций многих переменных: безусловный и условный экстремумы, метод наименьших квадратов, градиентные методы поиска экстремума.

Значительное место уделяется задачам оптимизации, линейному, нелинейному, квадратичному программированию, планированию экспериментальных исследований методом Бокса-Уилсона.

Находят отражение вопросы сглаживания экспериментальных зависимостей по методу наименьших квадратов и определения эмпирических формул, а также аппроксимации функций многих переменных.

Приводятся нормативные основы планирования эксперимента и обработки опытных данных.

Авторы выражают искреннюю признательность членам-корреспондентам РААСН, д.т.н., проф. Черкасову В.Д., Селяеву В.П., д.т.н., проф. Данилову А.А., д.ф.-м.н., проф. Голованову О.А. за участие в обсуждении книги.

Будем признательны за любые замечания и предложения по содержанию книги и излагаемым в ней вопросам.

*Авторский коллектив*

# 1. ЭЛЕМЕНТЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

**Предмет теории вероятностей.** Теория вероятностей – математическая наука, изучающая закономерности в случайных явлениях.

*Случайное явление* – это явление, которое при неоднократном воспроизведении одного и того же опыта протекает каждый раз несколько по-иному. Так, при стрельбе из орудия фактическая траектория снаряда несколько отклоняется от теоретической за счет совокупного влияния многих факторов: ошибки изготовления снаряда, отклонения веса снаряда от номинала, ошибки установки ствола в заданное положение, метеоусловий и т.д.

Исследование природы и структуры случайных воздействий, действующих на систему и влияющих на конечный результат, представляет чрезвычайно важную задачу.

Как правило, на изучаемую систему непрерывно воздействуют случайные помехи. Их наличие приводит к тому, что система поставленную задачу решает с ошибкой, в ряде случаев выходящей за пределы допустимой. Важно знать, как часто будут появляться такие ошибки и какие следует принять меры для того, чтобы практически исключить их возможность.

Элемент неопределенности, сложности, многопричинности, присущий случайным явлениям, требует создания специальных методов для их изучения. Такие методы и разрабатываются в теории вероятностей. Ее предметом являются специфические закономерности, наблюдаемые в случайных явлениях.

*Вероятностный, или статистический, метод* в науке не противопоставляет себя классическому, обычному методу точных наук, а является его дополнением, позволяющим глубже анализировать явления с учетом присущих ему элементов случайности.

Теория вероятностей находит широкое применение в строительном материаловедении, разработке и проектировании строительных объектов, в теории стрельбы и бомбометания, теории боеприпасов, теории прицелов и приборов управления огнем, теории распознавания образов, теории электронно-вычислительных машин, в распределении сырьевых ресурсов и т.д.

## 1.1. Основные понятия теории вероятностей

Под *событием* понимается всякий факт, который в результате опыта может произойти или не произойти.

*Примеры:*

*A* – появление герба при бросании монеты,

*B* – появление трех гербов при трехкратном бросании монеты,

*C* – попадание в цель при выстреле,

*E* – обнаружение объекта при одном цикле обзора радиолокационной станции,

*D* – появление туза при вынимании карты из колоды.

Очевидно, событие *A* более возможно, чем событие *B* или *D*.

Чтобы *количественно сравнить* между собой события *по степени их возможности*, с событием связывают *определенное число – вероятность события*, которое тем больше, чем более возможно событие.

Более вероятными считаются те события, которые происходят чаще. Таким образом, понятие вероятности события связано с опытным, практическим понятием *частоты* события.

В качестве *единицы измерения возможности событий* естественно принять вероятность *достоверного события*, то есть события, которое в результате опыта непременно должно произойти (например, выпадение не более 6 очков при бросании игральной кости). Противоположным по отношению к достоверному событию является *невозможное*, которое в данном опыте произойти не может.

Достоверному событию приписывается вероятность, равная 1, невозможному – 0. Тогда все остальные возможные события будут иметь вероятность между 0 и 1.

**Классическое определение вероятности.** Предварительно введем несколько вспомогательных понятий.

*Полная группа событий* – совокупность таких событий, что в результате опыта должно появиться хотя бы одно из них (например, выпадение герба или надписи при бросании монеты; попадание или промах при выстреле; появление 1, 2, 3, 4, 5, 6 очков при бросании игральной кости).

*Несовместные события* – события, которые не могут появиться вместе в данном опыте (например, выпадение герба и надписи; попадание и промах).

*Равновозможные события* – такие события, для которых по условиям симметрии есть основания считать, что ни одно из них не является объективно более возможным, чем другие (например, появление 1, 2, 3, 4, 5, 6 очков при бросании игральной кости).

*Случаи* – несовместные, равновозможные события, образующие полную группу. Случай называется *благоприятным (благоприятствующим)* некоторому событию, если появление этого случая влечет за собой появление данного события. Например, при бросании игральной кости возможны 6 случаев: выпадение 1, 2, 3, 4, 5, 6 оч-

ков. Если событие  $A$  – появление четного числа очков, то этому событию благоприятны 3 случая: выпадение 2, 4 и 6 очков.

*Единственно возможные события* – такие события, если появление одного и только одного события является достоверным событием.

*Вероятностью события  $A$  называют отношение числа благоприятствующих этому событию исходов к общему числу всех несовместных, единственно возможных и равновозможных элементарных исходов испытания:*

$$P(A) = \frac{m}{n},$$

где  $m$  – число благоприятствующих событию  $A$  исходов;  $n$  – число всех возможных элементарных исходов.

Приведенная формула используется при непосредственном подсчете вероятностей и пригодна тогда и только тогда, когда имеется симметрия возможных исходов.

*Пример.* В урне 2 белых и 3 черных шара. Наугад вынимается один шар. Вероятность того, что шар будет белым,  $P(A) = \frac{2}{5}$ .

**Относительная частота** – отношение числа испытаний, в которых событие появилось, к общему числу фактически произведенных испытаний:

$$W(A) = \frac{M}{n},$$

где  $M$  – число появлений события  $A$ ;  $n$  – общее число испытаний.

Заметим, при определении вероятности не требуется, чтобы испытания в действительности проводились, а при определении относительной частоты это обязательно. *Вероятность вычисляют до опыта, а относительную частоту – после опыта.*

Как показали длительные наблюдения, относительная частота колеблется около некоторого *постоянного* числа, то есть проявляется ее **устойчивость**. Оказалось, что это постоянное число – есть вероятность появления события. Ниже этот вопрос будет рассмотрен подробнее.

**Ограниченность классического определения вероятности** заключается в том, что здесь *число исходов* предполагается *конечным*. К тому же, часто нельзя события предполагать равновозможными. В связи с этим *наряду* с классическим определением используется *статистическое определение вероятности*, где за *вероятность события* принимается *относительная частота* или число, близкое к

ней (так, если в результате достаточно большого числа испытаний оказалось, что относительная частота близка к 0,4, то это число можно принять за статистическую вероятность события).

## 1.2. Основные теоремы теории вероятностей

На практике обычно требуется определить вероятность событий, непосредственное экспериментальное воспроизведение которых затруднено (дорого или просто невозможно, например, при проектировании новой техники). Поэтому для определения вероятностей применяются не непосредственные *прямые методы*, а *косвенные*, позволяющие по известным вероятностям одних событий определять вероятности других, связанных с ними событий.

Вся теория вероятностей, в основном, и представляет собой систему таких косвенных методов, позволяющих свести потребность экспериментальных исследований к минимуму.

При применении косвенных методов в той или иной мере пользуются двумя *основными теоремами* теории вероятностей: *сложения вероятностей и умножения вероятностей*.

Введем два вспомогательных понятия.

*Сумма двух событий  $A$  и  $B$*  – это событие  $C$ , состоящее в выполнении события  $A$  или  $B$ , или обоих вместе (то есть выполнение хотя бы одного из событий  $A$  и  $B$ ). Например, если  $A$  – попадание в цель при первом выстреле,  $B$  – при втором, то  $C = A + B$  – попадание в цель при двух выстрелах.

*Сумма нескольких событий – это событие, состоящее в появлении хотя бы одного из них.*

*Произведение двух событий  $A$  и  $B$*  – это событие  $C$ , состоящее в совместном выполнении  $A$  и  $B$ . Например, если  $A$  – появление туза,  $B$  – появление бубновой масти, то  $C = AB$  – появление бубнового туза.

*Произведение нескольких событий* – это событие, состоящее в совместном их появлении.

**Теорема сложения вероятностей.** Вероятность появления одного из двух несовместных событий (то есть суммы), безразлично какого, равна сумме вероятностей этих событий

$$P(A + B) = P(A) + P(B).$$

Действительно, пусть  $n$  – общее число возможных исходов испытания;  $m_1$ ,  $m_2$  – числа исходов, благоприятствующих появлению событий  $A$  и  $B$  соответственно.

Тогда  $m_1 + m_2$  – число исходов, благоприятствующих  $A + B$ , и

$$P(A + B) = \frac{m_1 + m_2}{n} = \frac{m_1}{n} + \frac{m_2}{n} = P(A) + P(B).$$

Теорема методом полной индукции легко обобщается на случай конечного числа несовместных событий.

**Следствие 1.** Если  $A_1, A_2, \dots, A_n$  – полная группа событий, то

$$\sum_{i=1}^n P(A_i) = 1.$$

Действительно, так как  $A_1, A_2, \dots, A_n$  – полная группа, то появление хотя бы одного из них есть достоверное событие, то есть  $P(A_1 + A_2 + \dots + A_n) = 1$ . Откуда  $P(A_1) + P(A_2) + \dots + P(A_n) = 1$ .

Противоположными называются два несовместных события  $A$  и  $\bar{A}$ , образующих полную группу.

**Следствие 2.** Сумма вероятностей противоположных событий равна 1:

$$P(A) + P(\bar{A}) = 1.$$

Событие  $A$  называется *независимым* от  $B$ , если вероятность события  $A$  не зависит от того, произошло событие  $B$  или нет.

Событие  $A$  называется *зависимым* от  $B$ , если вероятность события  $A$  зависит от того, произошло событие  $B$  или нет.

*Условная вероятность события  $A$*  – это вероятность события  $A$ , вычисленная при условии, что имело место событие  $B$ ; обозначается  $P_B(A)$  или  $P(A | B)$ .

**Теорема умножения вероятностей.** Вероятность произведения двух событий равна произведению одного из них на условную вероятность другого, вычисленную при условии, что первое имело место:

$$P(AB) = P(A)P(B | A).$$

Действительно, пусть  $m$  – число случаев, благоприятствующих событию  $A$ ;  $l$  – число случаев, благоприятствующих событию  $AB$ .

$$\text{Тогда } P(AB) = \frac{l}{n}, \quad P(A) = \frac{m}{n}.$$

Вычислим  $P(B | A)$ . Так как событие  $A$  произошло, то из ранее возможных  $n$  случаев остаются возможными только те  $m$ , которые благоприятствовали этому событию  $A$ . Из этих  $m$  случаев всего  $l$  благоприятствуют событию  $B$ . Так что  $P(B | A) = \frac{l}{m}$ .



Подставляя полученные значения в формулу теоремы умножения, получим тождество

$$\frac{l}{n} \equiv \frac{m}{n} \frac{l}{m},$$

откуда следует справедливость утверждения теоремы.

**Следствие 1.** Если событие  $A$  не зависит от события  $B$ , то  $B$  не зависит от  $A$ .

Действительно, так как  $A$  не зависит от  $B$ , то  $P(A) = P_B(A)$ .

Имеем  $P(AB) = P(A)P_A(B)$  или  $P(AB) = P(B)P_B(A)$ . Тогда  $P(A)P_A(B) = P(B)P_B(A)$ . Откуда при  $P(A) \neq 0$  следует  $P(B) = P_A(B)$ , то есть событие  $B$  не зависит от события  $A$ .

Таким образом, зависимость или независимость событий взаимны.

**Следствие 2.** Вероятность произведения двух независимых событий равна произведению вероятностей этих событий

$$P(AB) = P(A)P_A(B) = P(A)P(B).$$

Методом полной индукции легко получим:

$$P(A_1 A_2 \dots A_n) = P(A_1)P(A_2 | A_1)P(A_3 | A_1 A_2) \dots P(A_n | A_1 A_2 \dots A_{n-1}).$$

**Теорема.** Вероятность появления хотя бы одного из событий  $A_1, A_2, \dots, A_n$ , независимых в совокупности, равна разности между единицей и произведением вероятностей противоположных событий.

Действительно, пусть  $A$  – событие, состоящее в появлении хотя бы одного из  $A_1, A_2, \dots, A_n$ . Тогда  $A$  и  $\bar{A}_1 \bar{A}_2 \dots \bar{A}_n$  (ни одно из событий не появилось) противоположны, то есть

$$P(A) + P(\bar{A}_1 \bar{A}_2 \dots \bar{A}_n) = 1;$$

$$P(A) = 1 - P(\bar{A}_1 \bar{A}_2 \dots \bar{A}_n);$$

$$P(A) = 1 - q_1 q_2 \dots q_n, \text{ где } q_i = P(\bar{A}_i), \quad i = \overline{1, n}.$$

#### Следствия теорем сложения и умножения

**Теорема.** Вероятность появления хотя бы одного из двух совместных событий равна сумме вероятностей этих событий без вероятности их совместного появления:

$$P(A + B) = P(A) + P(B) - P(AB).$$

Действительно, событие  $A$  состоит в наступлении одного из двух несовместных событий  $A\bar{B}$  или  $AB$ :

$$A = A\bar{B} + AB.$$

По теореме сложения вероятностей несовместных событий  $P(A) = P(\overline{AB}) + P(AB)$  и  $P(\overline{A}B) = P(A) - P(AB)$ .

Аналогично  $P(B) = P(\overline{AB}) + P(AB)$  и  $P(\overline{A}B) = P(B) - P(AB)$ .

Событие  $A + B$  будет осуществлено, если наступит одно из трех несовместных событий  $\overline{A}\overline{B}$ ,  $\overline{A}B$ ,  $AB$ . Так что по теореме сложения вероятностей несовместных событий имеем:

$$P(A + B) = P(\overline{A}\overline{B}) + P(\overline{A}B) + P(AB) = \\ [P(A) - P(AB)] + [P(B) - P(AB)] + P(AB).$$

Откуда  $P(A + B) = P(A) + P(B) - P(AB)$ , что и требовалось доказать.

**Формула полной вероятности.** Вероятность события  $A$ , которое может наступить лишь при условии появления одного из несовместных событий (гипотез)  $H_1, H_2, \dots, H_n$  равна сумме произведений вероятностей каждого из  $H_i$  на соответствующую условную вероятность события  $A$ , то есть

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A | H_i).$$

Здесь требуется определить вероятность события  $A$ , которая может наступить при условии появления одного из несовместных событий  $H_1, H_2, \dots, H_n$ , образующих полную группу несовместных событий. Поскольку заранее неизвестно, какое из событий  $H_i$  наступит, то  $H_i$  называются гипотезами.

Действительно,  $H_1, H_2, \dots, H_n$  составляют полную группу, так что  $A$  может появиться только в комбинации с какой-либо из гипотез  $H_i$ , то есть  $A = H_1A + H_2A + \dots + H_nA$ .

Здесь  $H_i$  несовместны, то и  $H_iA$  несовместны. Откуда по Теореме сложения вероятностей  $P(A) = \sum_{i=1}^n P(H_iA)$ . Но по теореме умножения вероятностей  $P(H_iA) = P(H_i)P(A | H_i)$ .

Откуда следует  $P(A) = \sum_{i=1}^n P(H_i) \cdot P(A | H_i)$ , что и требовалось доказать.

**Формула Байеса (теорема гипотез).** Она имеет вид

$$P(H_i | A) = \frac{P(H_i)P(A | H_i)}{\sum_{i=1}^n P(H_i)P(A | H_i)}, \quad i = \overline{1, n}.$$

Здесь  $H_1, H_2, \dots, H_n$  – полная группа несовместных гипотез. Вероятности этих гипотез до опыта (*априорные* вероятности) равны  $P(H_1), P(H_2), \dots, P(H_n)$ . Произвели опыт, и появилось событие  $A$ . Требуется определить, как в связи с появлением события  $A$  изменились вероятности гипотез (*апостериорные* вероятности), т. е.  $P(H_i | A)$ .

По теореме умножения вероятностей  $P(AH_i) = P(A)P(H_i | A)$  или  $P(AH_i) = P(H_i)P(A | H_i)$ . Так что

$$P(A)P(H_i | A) = P(H_i)P(A | H_i).$$

Откуда  $P(H_i | A) = \frac{P(H_i)P(A | H_i)}{P(A)}$ . Из  $P(A) = \sum_{i=1}^n P(H_i) \cdot P(A | H_i)$

следует  $P(H_i | A) = \frac{P(H_i)P(A | H_i)}{\sum_{i=1}^n P(H_i)P(A | H_i)}$ , что и требовалось доказать.

### 1.3. Повторение испытаний

*Испытания называются независимыми относительно события  $A$ , если вероятность события  $A$  при каждом повторении испытания не зависит от исходов других испытаний.*

Независимые испытания могут производиться в одинаковых или различных условиях. В первом случае  $P(A)$  во всех испытаниях одинаковы. Здесь это и предполагается.

Укажем некоторые примеры:

- несколько последовательных бросаний монеты есть независимые испытания;
- несколько последовательных извлечений карты из колоды являются независимыми, если вынутая карта каждый раз возвращается в колоду;
- несколько выстрелов являются независимыми, если каждый раз перед выстрелами производится прицеливание.

Определим *вероятность  $P_{m,n}$  того, что событие  $A$  в  $n$  опытах появится ровно  $m$  раз*, если вероятность появления события  $A$  в каждом опыте постоянна и равна  $p$  (соответственно вероятность не появления  $A$  равна  $q = 1 - p$ ).

Для этого обозначим через  $B_m$  событие, состоящее в появлении  $A$  в  $n$  произведенных опытах ровно  $m$  раз. Обозначим  $A_i$  – появление события  $A$  в  $i$ -м опыте,  $\bar{A}_i$  – не появление  $A$  в этом же опыте.

Тогда  $B_m = A_1 A_2 \cdots A_m \bar{A}_{m+1} \cdots \bar{A}_n + \cdots + \bar{A}_1 \bar{A}_2 \cdots \bar{A}_{n-m} A_{n-m+1} \cdots A_n$ .

В каждое произведение  $A$  входит  $m$  раз,  $\bar{A}$  —  $(n - m)$  раз. Число всех таких комбинаций равно  $C_n^m$ , то есть числу способов, какими из  $n$  способов можно выбрать  $m$ , в которых произошло событие  $A$ . Вероятность каждой комбинации по теореме умножения независимых испытаний равна  $p^m q^{n-m}$ . Так как комбинации несовместны, то воспользуемся теоремой сложения. Будем иметь

$$P_{m,n} = P(B_m) = \underbrace{p^m q^{n-m} + \cdots + p^m q^{n-m}}_{C_n^m \text{ раз}} = C_n^m p^m q^{n-m}.$$

Таким образом, вероятность появления события  $A$  в  $n$  независимых опытах ровно  $m$  раз определяется формулой Бернулли

$$P_{m,n} = C_n^m p^m q^{n-m}.$$

В связи с тем, что вероятности  $P_{m,n}$  по форме представляют собой члены разложения бинома  $(q + p)^n$ , распределение вероятностей вида

$$P_{m,n} = C_n^m p^m q^{n-m}$$

называется *биномиальным распределением*.

#### 1.4. Случайные величины и законы их распределения

Случайной величиной называется величина, которая в результате опыта может принять одно и только одно из возможных значений, наперед неизвестное и зависящее от случайных причин, которые заранее не могут быть учтены.

Так, например:

- число  $X$  отказавших элементов в приборе, состоящем из 5 элементов — случайная величина: 0, 1, 2, 3, 4, 5 — ее возможные значения;

- число родившихся мальчиков среди 100 новорожденных в некотором родильном доме — случайная величина, 0, 1, 2, ..., 100 — ее возможные значения.

Случайные величины будем обозначать большими буквами  $X, Y, Z, \dots$ , их возможные значения соответствующими малыми  $x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_m; z_1, z_2, \dots, z_k; \dots$ .

Например, если  $X$  — число попаданий при трех выстрелах, то ее возможные значения —  $x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3$ .

Дискретной называется случайная величина, которая может принимать отдельные, изолированные значения с определенными вероятностями.

Число ее возможных значений может быть как конечным, так и бесконечным. Здесь ограничимся рассмотрением случайных величин с конечным числом возможных значений.

Приведенные примеры – примеры дискретных случайных величин.

В результате опыта дискретная случайная величина  $X$  примет одно из возможных значений: произойдет одно событие из полной группы несовместных событий. Если вероятности этих событий соответственно равны  $P(X = x_1) = p_1, P(X = x_2) = p_2, \dots, P(X = x_n) = p_n$ ,

то  $\sum_{i=1}^n p_i = 1$ .

Эта суммарная вероятность *распределена* между отдельными возможными значениями.

*Законом распределения случайной величины называется всякое соответствие, устанавливающее связь между возможными значениями  $x_i$  и соответствующими им вероятностями  $p_i$*  (говорят, что  $X$  подчиняется данному закону распределения).

Простейшей формой задания этого закона является таблица:

$X$	$x_1$	$x_2$	...	$x_n$
$p$	$p_1$	$p_2$	...	$p_n$

Она называется *рядом распределения* случайной величины  $X$ .

### Функция распределения

Рассмотрим случайную величину, имеющую бесчисленное множество возможных значений, *сплошь* заполняющих некоторый промежуток. Здесь составление указанной таблицы ряда распределения невозможно.

В этом случае для количественной характеристики распределения вероятностей пользуются не вероятностью события  $X = x$ , а вероятностью события  $X < x$ , где  $x$  – текущая переменная. Вероятность  $P(X < x)$  этого события есть функция  $x$ .

Функция  $F(x) = P(X < x)$  называется *функцией распределения случайной величины  $X$*  (или интегральной функцией распределения, или интегральным законом распределения).

Функция распределения – наиболее универсальная характеристика случайной величины и существует для всех случайных величин; является одной из форм закона распределения.

*Свойства функции распределения:*

1.  $F(x)$  – неубывающая функция (если  $x_2 > x_1$ , то  $F(x_2) \geq F(x_1)$ );
2.  $F(-\infty) = 0$ ;
3.  $F(+\infty) = 1$ .

Зная ряд распределения случайной величины, легко построить функцию распределения.

*Случайная величина называется непрерывной, если ее интегральная функция  $F(x)$  непрерывна в любой точке и дифференцируема всюду, кроме, быть может, отдельных точек.*

### **Вероятность попадания случайной величины на заданный участок $[\alpha, \beta]$**

Попадание случайной величины на заданный участок  $[\alpha, \beta]$  равносильно выполнению неравенства  $\alpha \leq X < \beta$ .

Пусть  $A$  – событие, состоящее в том, что  $X < \beta$ ;  $B$  – событие, состоящее в том, что  $X < \alpha$ ;  $C$  – событие, состоящее в том, что  $\alpha \leq X < \beta$ . Тогда  $A = B + C$ , и по теореме сложения вероятностей несовместных событий будем иметь  $P(A) = P(B) + P(C)$ , то есть

$$P(X < \beta) = P(X < \alpha) + P(\alpha \leq X < \beta).$$

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

Таким образом, вероятность попадания случайной величины на заданный участок равна приращению функции распределения на этом участке.

Справедливо утверждение: *вероятность любого отдельного значения непрерывной случайной величины равна нулю.*

Действительно, будем уменьшать участок  $(\alpha, \beta)$ , полагая  $\beta \rightarrow \alpha$ .

Тогда  $P(X = \alpha) = \lim_{\beta \rightarrow \alpha} P(\alpha \leq X < \beta) = \lim_{\beta \rightarrow \alpha} [F(\beta) - F(\alpha)] = 0$  (в силу непрерывности  $F(x)$ ).

*Если  $X$  – непрерывная случайная величина, то вероятность попадания случайной величины в интервал  $(\alpha, \beta)$  не зависит от того, является этот интервал открытым или закрытым, то есть*

$$P(\alpha < X < \beta) = P(\alpha \leq X < \beta) = P(\alpha < X \leq \beta) = P(\alpha \leq X \leq \beta)$$

Действительно,

$$\begin{aligned} P(\alpha \leq X \leq \beta) &= P(X = \alpha) + P(\alpha < X < \beta) + P(X = \beta) = \\ &= 0 + P(\alpha < X < \beta) + 0 = P(\alpha < X < \beta). \end{aligned}$$

### Плотность распределения (дифференциальная функция)

В силу предыдущего вероятность попадания непрерывной случайной величины  $X$  на участок  $[x, x + \Delta x]$  определится в виде:

$$P(x \leq X \leq x + \Delta x) = F(x + \Delta x) - F(x).$$

Отношение этой вероятности к длине участка даст среднюю вероятность, приходящуюся на единицу длины на этом участке. При  $\Delta x \rightarrow 0$  получим

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

Используется обозначение  $F'(x) = f(x)$ ;  $f(x)$  — называется *плотностью распределения*. Она характеризует плотность, с которой распределяются значения случайной величины в данной точке.

Кривая, изображающая плотность распределения случайной величины, называется *кривой распределения*.

С точностью до бесконечно малых высших порядков вероятность попадания  $X$  на элементарный участок  $[x, x + \Delta x]$  равна  $f(x)dx$ :

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

Геометрически вероятность попадания  $X$  на участок  $[x, x+dx]$  равна заштрихованной площади кривой распределения (рис. 1).

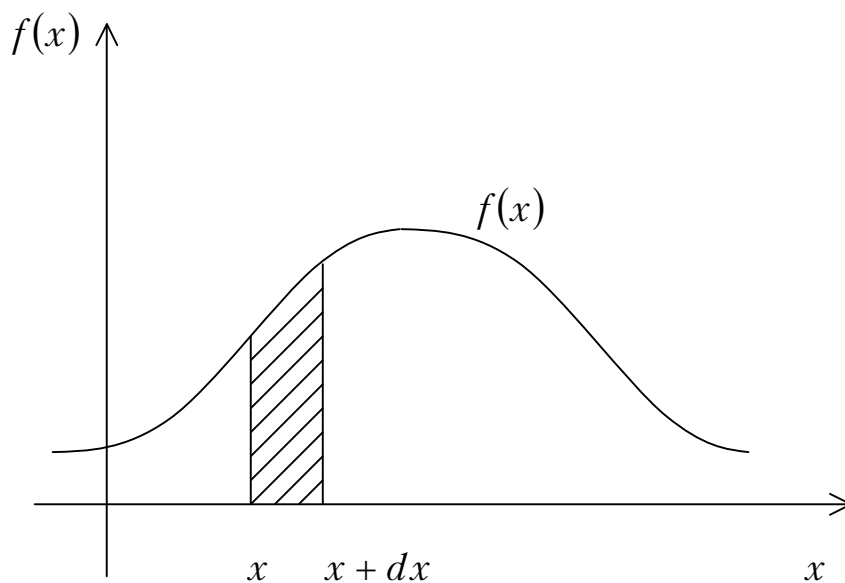


Рис. 1

Выразим функцию распределения через плотность распределения. Имеем

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

Таким образом,

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

### Основные свойства плотности распределения

1.  $f(x) \geq 0$  (так как  $F(x)$  – неубывающая функция).
2.  $\int_{-\infty}^{\infty} f(x)dx = 1$  (следует из  $F(+\infty) = 1$ ).

### 1.5. Эмпирическая функция распределения

Рассмотрим случайную величину  $X$ , закон распределения которой в точности неизвестен. Пусть этот закон требуется определить из опыта или проверить экспериментально гипотезу о том, что величина  $X$  подчинена тому или иному закону.

В каждом из опытов случайная величина  $X$  принимает определенное значение. Совокупность наблюдаемых значений представляет собой первичный статистический материал (*простая статистическая совокупность* или *простой статистический ряд*), подлежащий обработке и анализу. Обычно он оформляется в виде таблицы с указанием номера опыта и соответствующих наблюдаемых значений.

Простой статистический ряд обрабатывается различными способами, одним из которых является построение *статистической функции распределения* случайной величины или как иначе называют, *эмпирической функции распределения*.

*Эмпирической функцией распределения* называют функцию  $F^*(x)$ , определяющую для каждого значения  $X$  относительную частоту события  $X < x$ :

$$F^*(x) = \frac{n_x}{n},$$

где  $n_x$  – число опытов, при которых наблюдаемые значения  $X$  были меньше  $x$ ,  $n$  – общее число наблюдений (опытов).



Из формулы очевиден способ нахождения  $F^*(x)$ .

Различие между эмпирической функцией распределения  $F^*(x)$  и теоретической функцией распределения  $F(x)$  состоит в том, что  $F(x)$  для заданного значения  $x$  определяет вероятность события  $X < x$ , а  $F^*(x)$  – определяет относительную частоту этого события. По теореме, носящей название *теоремы Бернулли*, при  $n \rightarrow \infty$  частота события  $X < x$  стремится к  $P(X < x)$ . Это обстоятельство позволяет приближенно использовать  $F^*(x)$  вместо  $F(x)$ .

### Статистический ряд. Гистограмма

При большом числе наблюдений простая статистическая совокупность становится громоздкой и мало наглядной. В этом случае целесообразней построение статистического ряда.

Разобьем весь диапазон наблюдаемых значений  $X$  на интервалы (разряды) и подсчитаем количество значений  $n_i$ , приходящихся на каждый разряд. Это число разделим на общее число наблюдений  $n$  и найдем частоту, соответствующую данному разряду  $p_i^* = \frac{n_i}{n}$ ;  $\sum_i p_i^* = 1$ .

Построим таблицу, в которой приведем разряды в порядке возрастания  $X$  и соответствующие им частоты.

$I_i$	$x_1, x_2$	$x_2, x_3$	...	$x_i, x_{i+1}$	...	$x_k, x_{k+1}$
$p_i^*$	$p_1^*$	$p_2^*$	...	$p_i^*$	...	$p_k^*$

Такая таблица называется *статистическим рядом*. Здесь  $I_i$  – обозначение  $i$ -го разряда,  $x_i, x_{i+1}$  – его границы,  $k$  – число разрядов.

Если для каждого возможного значения  $x_i$  указывается его относительная частота, то статистический ряд называется *статистическим распределением*.

Статистический ряд часто оформляется в виде *гистограммы*. Она представляет собой ступенчатую фигуру из прямоугольников с основаниями, равными интервалам значений признака и высотами, равными частотам интервалов (рис. 2).

Для построения гистограммы нужно относительную частоту каждого разряда  $W_i = p_i^* = \frac{n_i}{n}$  разделить на длину разряда и полученное число взять в качестве высоты прямоугольника. Если длины разрядов равны  $h$ , то высоты равны  $\frac{W_i}{h}$ . Площадь гистограммы

$$S = \sum h \frac{W_i}{h} = \sum W_i = 1.$$

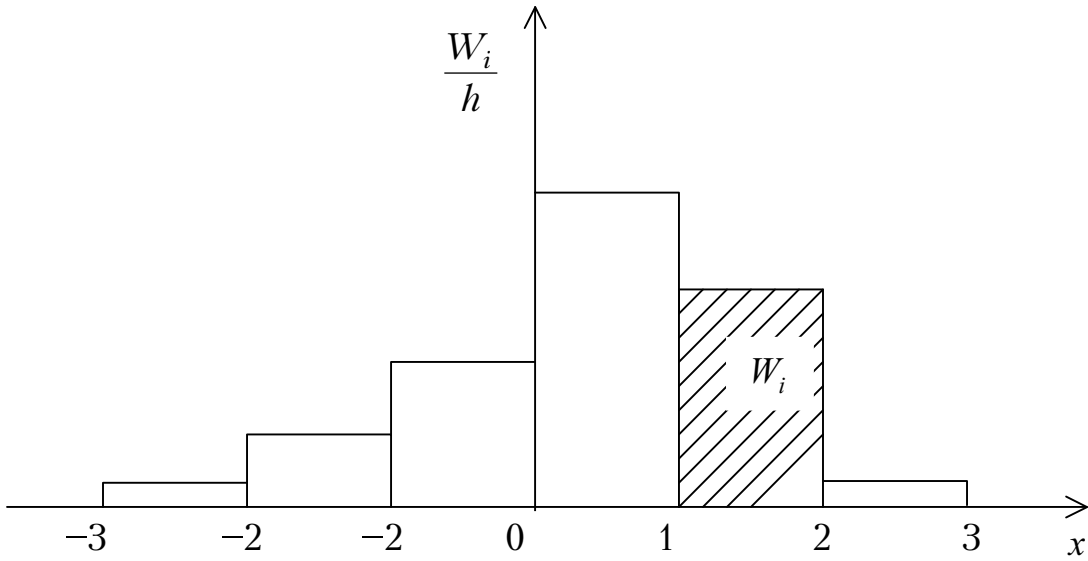


Рис.2

По статистическому ряду легко определяется эмпирическая функция распределения. Если  $x_1, x_2, \dots, x_{k+1}$  – границы разрядов ряда, то  $F^*(x_1) = 0$ ,  $F^*(x_2) = p_1^*$ ,  $F^*(x_3) = p_1^* + p_2^*, \dots$ ,  $F^*(x_k) = \sum_{i=1}^{k-1} p_i^*$ ,

$$F^*(x_{k+1}) = \sum_{i=1}^k p_i^* = 1.$$

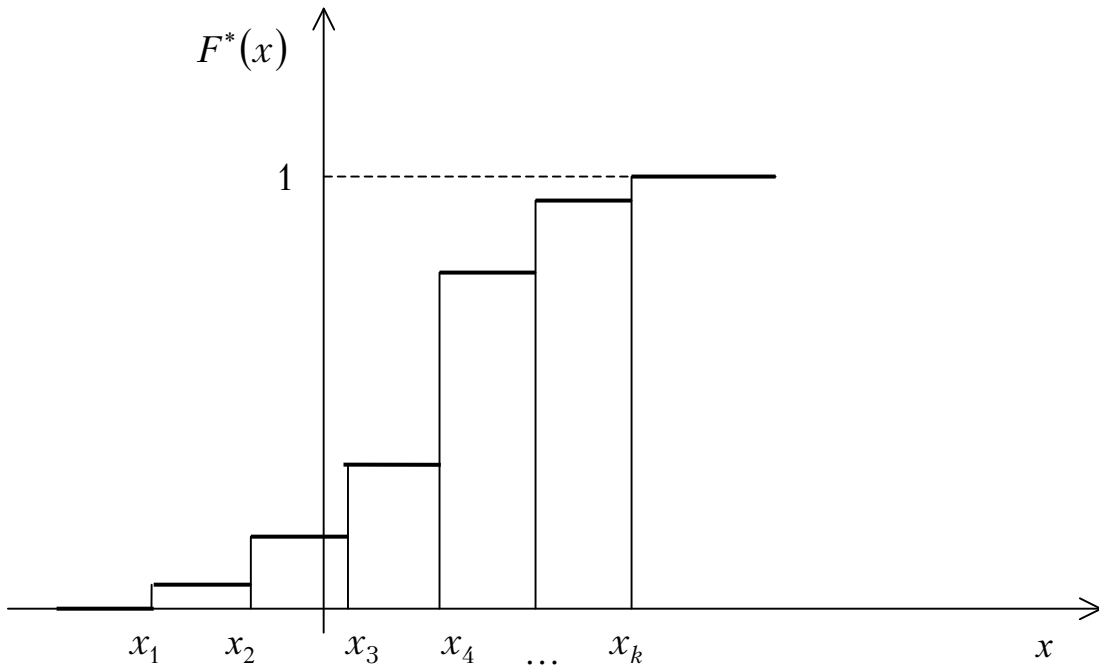


Рис. 3

## 1.6. Числовые характеристики случайных величин

При решении многих практических вопросов отсутствует возможность полного описания случайной величины. Часто и закон распределения неизвестен. Приходится ограничиваться меньшими сведениями. Иногда даже выгоднее пользоваться некоторыми числами для интегрального (суммарного) описания случайной величины. Такие числа называются *числовыми характеристиками случайной величины*.

**Математическое ожидание.** Для дискретной случайной величины – это сумма произведений всех возможных значений случайной величины на вероятности этих значений

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

Оно характеризует среднее взвешенное значение случайной величины с учетом различных значений вероятностей возможных значений.

Математическое ожидание случайной величины  $X$  связано со своеобразной зависимостью со средним арифметическим наблюдаемых значений случайной величины при большом числе опытов.

Действительно, пусть в  $N$  независимых опытах  $x_1$  появилось  $m_1$  раз,  $x_2$  –  $m_2$  раз, ...,  $x_n$  –  $m_n$  раз;  $\sum_{i=1}^n m_i = N$ .

Среднее арифметическое наблюдаемых значений

$$\begin{aligned} M^*[X] &= \frac{x_1 m_1 + x_2 m_2 + \dots + x_n m_n}{N} = x_1 \frac{m_1}{N} + x_2 \frac{m_2}{N} + \dots + x_n \frac{m_n}{N} = \\ &= x_1 p_1^* + x_2 p_2^* + \dots + x_n p_n^* . \end{aligned}$$

Здесь  $p_i^* = \frac{m_i}{N}$  – относительная частота, которая при  $N \rightarrow \infty$

стремится к вероятности  $p_i$ . Откуда следует, что  $M^*[X] = \sum_{i=1}^n x_i p_i^*$

стремится к математическому ожиданию  $M[X] = \sum_{i=1}^n x_i p_i$ .

Для непрерывной случайной величины математическим ожиданием называется

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

### Свойства математического ожидания

Если  $C$  – неслучайная величина, то:

1.  $M[C] = C$ ,
2.  $M[CX] = CM[X]$ .

Если  $X, Y, \dots, Z$  – взаимно независимые случайные величины, то:

3.  $M[X + Y + \dots + Z] = M[X] + M[Y] + \dots + M[Z]$ ,
4.  $M[X_1 \cdot X_2 \cdot \dots \cdot X_n] = M[X_1] \cdot M[X_2] \cdot \dots \cdot M[X_n]$ ,
5.  $M\left[\sum_{i=1}^n a_i X_i + b\right] = \sum_{i=1}^n a_i M[X_i] + b$ .

**Моменты.** Начальным моментом  $s$ -го порядка **дискретной** случайной величины называется

$$\alpha_s[X] = \sum_{i=1}^n x_i^s p_i,$$

для **непрерывной** случайной величины

$$\alpha_s[X] = \int_{-\infty}^{\infty} x^s f(x) dx.$$

Как видим, начальным моментом  $s$ -го порядка является математическое ожидание  $s$ -й степени случайной величины

$$M[X^s] = \alpha_s[X],$$

в частности, математическое ожидание

$$M[X] = \alpha_1[X].$$

*Центрированной* случайной величиной, соответствующей случайной величине  $X$ , называется

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

где  $m_x = M[X]$ .

Заметим,

$$M\left[\overset{\circ}{X}\right] = M[X] - M[m_x] = M[X] - m_x = 0.$$

Центрирование случайной величины равносильно переносу начала координат в среднюю точку («центр»), абсцисса которой равна математическому ожиданию.

**Центральным моментом**  $s$ -го порядка называется

$$\mu_s[X] = M\left[\overset{\circ}{X}^s\right] = M[(X - m_x)^s]:$$

– для *дискретной* случайной величины  $X$

$$\mu_s[X] = \sum_{i=1}^n (x_i - m_x)^s p_i;$$

– для *непрерывной* случайной величины  $X$

$$\mu_s[X] = \int_{-\infty}^{\infty} (X - m_x)^s f(x) dx.$$

Для любой случайной величины центральный момент первого порядка  $\mu_1[X] = 0$ .

Второй центральный момент называется **дисперсией** случайной величины

$$\mu_2[X] = D[X] = M\left[\overset{\circ}{X}^2\right]:$$

– для *дискретной* случайной величины  $X$

$$D[X] = \sum_{i=1}^n (x_i - m_x)^2 p_i;$$

– для *непрерывной* случайной величины  $X$

$$D[X] = \int_{-\infty}^{\infty} (X - m_x)^2 f(x) dx.$$

Дисперсия случайной величины есть характеристика рассеивания, разбросанности значений  $X$  около ее математического ожидания. Дисперсия имеет размерность квадрата случайной величины. Поэтому удобнее пользоваться **среднеквадратическим отклонением** случайной величины

$$\sigma[X] = \sqrt{D[X]}$$

(часто обозначают  $D[X] = D_x$ ,  $\sigma[X] = \sigma_x$ ;  $\sigma_x = \sqrt{D_x}$ ).

На практике для вычисления  $\sigma_x$  удобнее пользоваться формулой

$$\sigma_x = \sqrt{\alpha_2 - m_x^2}.$$

Справедливость формулы для дискретной случайной величины очевидна:

$$D[X] = \mu_2 = M\left[\overset{\circ}{X}^2\right] = \sum_{i=1}^n (x_i - m_x)^2 p_i = \sum_{i=1}^n x_i^2 p_i - 2m_x \sum_{i=1}^n x_i p_i + m_x^2 \sum_{i=1}^n p_i = \\ = \alpha_2 - 2m_x^2 + m_x^2 = \alpha_2 - m_x^2.$$

Математическое ожидание и среднее квадратическое отклонение характеризуют положение и степень разбросанности и являются наиболее часто применяемыми числовыми характеристиками случайной величины.

### Свойства дисперсии и среднее квадратическое отклонения

1.  $D[X] = M[X^2] - m_x^2.$

Если  $C$  – неслучайная величина, то:

2.  $D[C] = 0,$

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

Если  $X_1, X_2, \dots, X_n$  – взаимно независимые случайные величины, то:

4.  $D\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n D[X_i],$

5.  $D[C + X] = D[X],$

6.  $D[X \cdot Y] = D[X] \cdot D[Y],$

7.  $D\left[\sum_{i=1}^n a_i X_i + b\right] = \sum_{i=1}^n a_i^2 D[X_i],$

8.  $\sigma[X_1 + X_2 + \dots + X_n] = \sqrt{\sigma^2[X_1] + \sigma^2[X_2] + \dots + \sigma^2[X_n]}.$

### Математическое ожидание и дисперсия числа появлений события в серии независимых опытов

**Теорема.** Математическое ожидание  $M[X]$  числа появлений  $X$  события  $A$  в  $n$  независимых испытаниях равно произведению числа испытаний на вероятность  $p$  появления события в каждом испытании:

$$M[X] = np.$$

Действительно, число появлений  $X$  события  $A$  во всей серии опытов

$$X = X_1 + X_2 + \dots + X_n,$$

где  $X_i$  – число появлений события  $A$  в  $i$ -м опыте.

Так что  $M[X] = M[X_1] + M[X_2] + \dots + M[X_n].$

Каждая из  $X_i$  принимает два значения:

$x_1 = 1$  с вероятностью  $p$  (соответствует появлению события  $A$ ),

$x_2 = 0$  с вероятностью  $q = 1 - p$  (соответствует не появлению события  $A$ ).

Откуда

$$M[X_i] = 0 \cdot q + 1 \cdot p = p;$$

$$M[X] = np.$$

**Теорема.** *Дисперсия числа появлений события  $A$  в  $n$  независимых опытах, в каждом из которых вероятность появления  $A$  постоянна, равна произведению числа испытаний на вероятности появления и не появления события в одном испытании:*

$$D[X] = npq.$$

Действительно, число появлений события в серии опытов  $X = \sum_{i=1}^n X_i$ .

В силу независимости  $X_1, X_2, \dots, X_n$  по свойству 4 для дисперсий получим

$$D_x = \sum_{i=1}^n D_{x_i}.$$

Имеем

$$D[X_i] = \alpha_2 - M_x^2 = M[X_i^2] - m_x^2.$$

Подставляя

$$M[X_i^2] = 0 \cdot q + 1 \cdot p = p,$$

$$m_x = 0 \cdot q + 1 \cdot p = p,$$

$$m_x^2 = p^2,$$

получим

$$D[X_i] = p - p^2 = p(1 - p) = p \cdot q.$$

Откуда

$$D[X] = \sum_{i=1}^n D[X_i] = npq.$$

### **1.7. Закон больших чисел и центральная предельная теорема**

Физическое содержание закона больших чисел заключается в устойчивости среднего результата воздействия массы случайных явлений: при очень большом числе случайных явлений средний их результат практически перестает быть случайным и может быть предсказан с большой степенью определенности. Конкретные особен-

ности каждого отдельного случайного явления почти не сказываются на среднем результате массы случайных явлений; случайные отклонения от среднего, неизбежные в каждом отдельном явлении, в массе взаимно погашаются, нивелируются, выравниваются.

Для практики очень важно знать условия, при выполнении которых совокупное действие многих случайных причин приводит к результату, почти не зависящему от случая, что позволяет предвидеть ход явлений. Эти условия и указываются в теоремах, носящих общее название *закона больших чисел*.

Наиболее общим законом больших чисел является обобщенная теорема Л.П. Чебышева, простейшим (следствием из теоремы Чебышева) – теорема Бернулли.

**Обобщенная теорема Чебышева.** *Если  $X_1, X_2, \dots, X_n$  – попарно независимые случайные величины, причем дисперсии их не превышают постоянного числа  $C$ , то как бы мало ни было положительное число  $\varepsilon$ , вероятность неравенства*

$$\left| \frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n M[X_i]}{n} \right| < \varepsilon$$

*будет как угодно близка к 1, если число случайных величин  $n$  – достаточно велико.*

Другими словами, в условиях теоремы будет иметь место равенство

$$\lim_{n \rightarrow \infty} P \left( \left| \frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n M[X_i]}{n} \right| < \varepsilon \right) = 1.$$

Таким образом, теорема утверждает, что при возрастании  $n$  среднее арифметическое наблюдаемых значений величин  $X_1, X_2, \dots, X_n$  сходится по вероятности к среднему арифметическому их математических ожиданий.

**Примечание.** Говорят, что случайная величина  $X = \{X_n\}$  сходится по вероятности к  $a$ , если при увеличении  $n$  вероятность того, что  $X_n$  и  $a$  будут сколь угодно близки, неограниченно стремится к 1, то есть при достаточно большом  $n$

$$P(|X_n - a| < \varepsilon) > 1 - \delta,$$

где  $\varepsilon, \delta > 0$  – произвольные малые числа.



**Теорема Чебышева.** Если  $X_1, X_2, \dots, X_n$  – попарно независимые случайные величины, имеющие одно и то же математическое ожидание  $a$ , и если дисперсия этих величин не превышает постоянного числа  $C$ , то как бы мало ни было число  $\varepsilon > 0$ , вероятность неравенства

$$\left| \frac{\sum_{i=1}^n X_i}{n} - a \right| < \varepsilon.$$

сколь угодно близка к 1, если число  $n$  случайных величин достаточно велико, то есть

$$\lim_{n \rightarrow \infty} P \left( \left| \frac{\sum_{i=1}^n X_i}{n} - a \right| < \varepsilon \right) = 1.$$

Эта теорема – частный случай предыдущей. Здесь математические ожидания  $X_i$ ,  $i = \overline{1, n}$  предполагаются одинаковыми (в общем случае – разными).

**Теорема Бернулли.** Если в каждом из  $n$  независимых испытаний вероятность появления события  $A$  есть  $p = \text{const}$ , то вероятность того, что отклонение относительной частоты от вероятности  $p$  по абсолютной величине будет сколь угодно малым, если число испытаний достаточно велико, будет как угодно близко к 1, то есть

$$\lim_{n \rightarrow \infty} P \left( \left| \frac{M}{n} - p \right| < \varepsilon \right) = 1.$$

Теорема Бернулли объясняет, почему относительная частота при достаточно большом числе испытаний обладает свойством устойчивости и оправдывает статистическое определение вероятности.

Значение теоремы Чебышева для практики трудно переоценить по следующим причинам.

1. Обычно при измерении некоторой постоянной физической величины производят несколько измерений. Затем находят их среднее арифметическое, которое и принимают в качестве искомого значения измеряемой величины. Теорема Чебышева утверждает возможность и правильность этого способа измерения.

Действительно, результаты измерения можно рассматривать как случайные величины  $X_1, X_2, \dots, X_n$ . В теореме Чебышева случайные величины должны удовлетворять следующим условиям:

- должны быть попарно независимы,
- должны иметь одинаковые математические ожидания,
- их дисперсии должны быть ограничены одним и тем же числом.

Первое условие обычно выполняется, так как результат каждого измерения не зависит от результатов других измерений.

Второе условие выполняется, если измерения проводятся без систематических ошибок.

Третье условие также выполняется, если прибор обеспечивает определенную точность измерений.

Однако надо исходить из того, что, увеличивая число измерений, достичь сколь угодно большой точности нельзя. Если прибор дает показания лишь с точностью  $\pm \alpha$ , то и их среднее арифметическое будет измерено лишь с точностью  $\pm \alpha$ .

2. Теорема Чебышева объясняет, почему о качестве большого количества однородного материала можно судить по небольшой пробе.

3. На теореме Чебышева основан широко применяемый в статистике *выборочный метод*, когда по сравнительно небольшой выборке судят о всей совокупности исследуемых объектов (называется *генеральной совокупностью*).

**Центральная предельная теорема.** Во всех формах закона больших чисел рассматривалась сходимости по вероятности случайных величин к определенным постоянным, но нигде не оговаривались их законы распределения. Все формы центральной предельной теоремы посвящены установлению условий, при которых возникает нормальный закон распределения. Он возникает во всех случаях, когда исследуемая случайная величина представима в виде суммы большого числа независимых или слабо зависимых элементарных слагаемых, каждое из которых в отдельности мало влияет на сумму.

Нормальный закон распределения является доминирующим во многих областях. Он был впервые обоснован Лапласом и Гауссом исходя из теории ошибок измерения. В большинстве случаев ошибки измерения физических величин распределяются по нормальному закону. Это связано с тем, что ошибки измерения складываются из многочисленных элементарных ошибок, вызванных различными причинами. Одна из общих форм центральной предельной теоремы доказана А. М. Ляпуновым в 1900 году. Различные формы центральной предельной теоремы отличаются между собой условиями, накладываемыми на распределение образующих сумму случайных слагаемых. Приведем одну из самых простых форм центральной предельной теоремы, когда все слагаемые суммы распределены одинаково.

Если  $X_1, X_2, \dots, X_n$  – независимые случайные величины, имеющие один и тот же закон распределения с математическим ожиданием  $m$  и дисперсией  $\sigma^2$ , то при неограниченном увеличении  $n$  закон распределения суммы  $Y_n = \sum_{k=1}^n X_k$  неограниченно приближается к нормальному.

### 1.8. Нахождение числовых характеристик случайной величины из опыта

Заметим, что любое значение характеристики  $X$ , вычисленное на основе ограниченного числа опытов, всегда будет содержать элемент случайности. Такое приближенное, случайное значение числовой характеристики называется ее *оценкой*. При оценке одним числом оценка называется *точечной*.

Наблюденные значения случайной величины  $X_1, X_2, \dots, X_n$  можно рассматривать как  $n$  экземпляров случайной величины  $X$ , то есть  $n$  случайных величин, каждая из которых распределена по тому же закону, что и  $X$ .

Оценку параметра  $a$  будем обозначать  $\tilde{a}$ . Так как  $\tilde{a}$  вычисляется на основе  $X_1, X_2, \dots, X_n$  (случайные величины), то  $\tilde{a} = \tilde{a}(X_1, X_2, \dots, X_n)$  – случайная величина.

Чтобы оценка была доброкачественной, она должна удовлетворять ряду требований:

1. Потребуем, чтобы при  $n$ , достаточно большом,  $P(|\tilde{a} - a| < \varepsilon) > 1 - \delta$  ( $\delta$  сходилась к  $a$  по вероятности). В этом случае точечная оценка называется *состоятельной*.

2. При использовании  $\tilde{a}$  вместо  $a$  не должно быть систематической ошибки в сторону завышения или занижения, то есть надо, чтобы  $M[\tilde{a}] = a$ . Такая оценка называется *несмещенной*.

3. Необходимо, чтобы эта несмещенная оценка обладала по сравнению с другими оценками наименьшей дисперсией, то есть  $D[\tilde{a}] = \min$ . Такая оценка называется *эффективной*.

Оценки для математического ожидания и дисперсии. В качестве *оценки для математического ожидания* естественно предложить среднее арифметическое наблюдаемых значений  $\tilde{m}^*$ :

$$\tilde{m} \approx \tilde{m}^* = \frac{\sum_{i=1}^n X_i}{n}.$$

Эта оценка является состоятельной по теореме Чебышева. Она является также несмещенной, так как

$$M[\tilde{m}] = \frac{\sum_{i=1}^n m_i}{n} = m \quad (M[X] = m).$$

Дисперсия этой оценки

$$\begin{aligned} D[\tilde{m}] &= D\left[\frac{X_1}{n} + \frac{X_2}{n} + \dots + \frac{X_n}{n}\right] = D\left[\frac{X_1}{n}\right] + D\left[\frac{X_2}{n}\right] + \dots + D\left[\frac{X_n}{n}\right] = \\ &= \left(\frac{1}{n}\right)^2 D[X_1] + \left(\frac{1}{n}\right)^2 D[X_2] + \dots + \left(\frac{1}{n}\right)^2 D[X_n] = \\ &= \underbrace{\frac{1}{n^2} D \dots + \frac{1}{n^2} D}_{n \text{ раз}} = n \cdot \frac{1}{n^2} D = \frac{1}{n} D. \end{aligned}$$

Эффективность оценки зависит от вида закона распределения  $X$ . Можно доказать, что  $\tilde{m}$  является эффективной оценкой при нормальном распределении  $X$ .

Для оценки дисперсии  $D$  наиболее естественной представляется статистическая дисперсия, вычисленная на основе статистического распределения

$$D^* = \frac{\sum_{i=1}^n (X_i - m^*)^2}{n},$$

где  $m^*$  — статистическое среднее.

Однако эта оценка, будучи состоятельной, оказывается смещенной, так как  $M[D^*] = \frac{n-1}{n} D$ .

Чтобы ликвидировать это смещение достаточно ввести поправку, умножив  $D^*$  на  $\frac{n}{n-1}$ .

Такая исправленная статистическая дисперсия и берется в качестве оценки для  $D$ :

$$\tilde{D} = \frac{\sum_{i=1}^n (X_i - \tilde{m})^2}{n-1}.$$

При большом  $n$  обе оценки  $D^*$  и  $\tilde{D}$  будут различаться мало, то есть  $n \approx n-1$  и введение поправочного множителя  $\frac{n}{n-1}$  теряет смысл.

## 1.9. Некоторые виды законов распределения случайной величины

### 1.9.1. Закон биномиального распределения вероятностей

Пусть производится  $n$  независимых испытаний, в каждом из которых вероятность появления события  $A$  постоянна и равна  $p$ ;  $X$  – число появлений события  $A$  в этих испытаниях ( $X$  – дискретная случайная величина; может принимать значения:  $x_0 = 0, x_1 = 1, \dots, x_n = n$ ). Установим закон распределения  $X$ .

Вероятности появления события  $A$  ровно  $m$  раз в  $n$  независимых испытаниях по предыдущему определяются по формуле Бернулли:

$$P(X = m) = P_{m,n} = C_n^m p^m q^{n-m} = \frac{n!}{m!(n-m)!} p^m q^{n-m}.$$

Получим ряд распределения:

$X$	$n$	$n-1$	$\dots$	$m$	$\dots$	$0$
	$p^n$	$np^{n-1}q$	$\dots$	$C_n^m p^m q^{n-m}$	$\dots$	$q^n$

### 1.9.2. Закон равномерной плотности (равномерного распределения вероятностей)

Распределение вероятностей называют равномерным, если на интервале, которому принадлежат все возможные значения случайной величины  $X$ , дифференциальная функция (плотность распределения) имеет постоянное значение (рис. 4). Здесь речь идет о непрерывном законе распределения, то есть о законе распределения *непрерывной* случайной величины.

*Пример.* Шкала измерительного прибора проградуирована в некоторых единицах. Ошибку отсчета при округлении до целого деления можно рассматривать как случайную величину  $X$ , которая может принимать с постоянной плотностью вероятности любые значения между двумя соседними целыми делениями. Таким образом,  $X$  имеет равномерное распределение.

Здесь

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

$$\left( \int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^a 0 \cdot dx + \int_a^b c \cdot dx + \int_b^{\infty} 0 \cdot dx = \int_a^b c \cdot dx = 1; \quad c(b-a) = 1; \right. \\ \left. c = \frac{1}{b-a} \right).$$

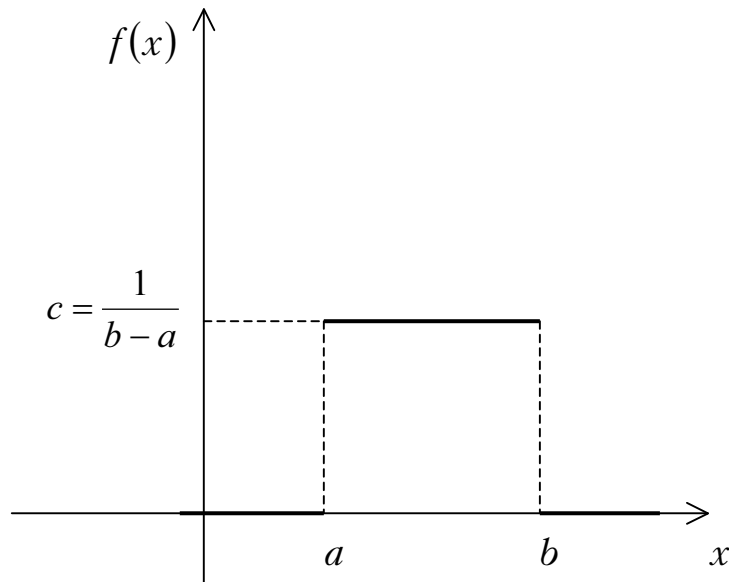


Рис. 4

### 1.9.3. Закон распределения Пуассона

Распределение Пуассона встречается в задачах, связанных с *простейшим потоком событий*.

Под потоком событий понимается последовательность событий, наступающих одно за другим в случайные моменты времени (*например, поток вызовов на телефонной станции, поток заявок в системе массового обслуживания, поток отказов при работе ЭВМ и т.д.*).

*Простейший поток событий* характеризуется свойствами:

- вероятность наступления того или иного числа событий за любой промежуток времени зависит только от длительности этого промежутка, но не зависит от начала отсчета;

- эта вероятность не зависит от того, какое число событий наступило до начала рассматриваемого промежутка времени;

- за малый промежуток времени  $\Delta t$  вероятность наступления одного события  $P_1(\Delta t)$  приближенно пропорциональна длительности этого промежутка, то есть  $P_1(\Delta t) = \lambda \cdot \Delta t + 0(\Delta t)$ ; вероятность наступ-

ления двух и более событий за этот промежуток пренебрежимо мала (события наступают поодиночке).

Пусть  $X$  – число событий простейшего потока, наступивших за фиксированный промежуток времени  $t$ . Значениями этой *случайной величины* могут быть любые целые числа  $m = x_m = 0, 1, 2, \dots$ . Соответствующие этим возможным значениям  $m$  дискретной случайной величины  $X$  вероятности того, что за фиксированный промежуток времени  $t$  наступит ровно  $m$  событий простейшего потока, обозначим

$$P_m(t) = P(X = m).$$

Для искомых вероятностей справедлива общая формула

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

где  $\lambda$  – среднее число событий, наступающих за единицу времени.

Обычно вводят  $a = \lambda t$  и закон распределения записывают в виде:

$$P(X = m) = \frac{a^m}{m!} e^{-a} \quad (m = 1, 2, \dots).$$

Этот закон распределения и называют *законом распределения Пуассона*.

Здесь ряд распределения имеет вид, приведенный в табл. 1

Т а б л и ц а 1

Число событий, наступивших за время $t$	0	1	2	3	...	$m$	...
Вероятность $P$	$e^{-a}$	$\frac{a}{1!} e^{-a}$	$\frac{a^2}{2!} e^{-a}$	$\frac{a^3}{3!} e^{-a}$	...	$\frac{a^m}{m!} e^{-a}$	...

Указанный ряд действительно является рядом распределения, так как

$$\sum_{m=0}^{\infty} P_m = \sum_{m=0}^{\infty} \frac{a^m}{m!} e^{-a} = e^{-a} \underbrace{\sum_{m=0}^{\infty} \frac{a^m}{m!}}_{e^a} = e^{-a} \cdot e^a = 1.$$

Можно показать: если случайная величина  $X$  распределена по закону Пуассона, то

$$M[X] = a; \quad D[X] = a.$$

Из  $M[X] = a$  следует, что  $a$  есть среднее число событий, наступающих за время  $t$ .

Закон распределения Пуассона может использоваться в качестве приближения для биномиального распределения, если в этом распределении количество независимых испытаний  $n$  полагается большим, а вероятность  $p$  появления события  $A$  в этом испытании полагается очень малым.

Действительно, обозначим  $np = a$ ;  $p = \frac{a}{n}$ ;  $q = 1 - p = 1 - \frac{a}{n}$ . Тогда

$$P(X = m) = P_{m,n} = C_m^n p^m q^{n-m} = C_m^n \left(\frac{a}{n}\right)^m \left(1 - \frac{a}{n}\right)^{n-m}.$$

Разлагая в степенной ряд по степеням  $\frac{1}{n}$ , будем иметь:

$$P_{m,n} = \frac{a^m}{m!} e^{-a} \left[ 1 + \frac{1}{n} \left( ma - \frac{a^2}{2} - \frac{n(m-1)}{2} \right) + \dots \right].$$

Откуда

$$C_m^n p^m q^{n-m} \approx \frac{a^m}{m!} e^{-a}.$$

*Пример.* На АТС поступают вызовы со средней плотностью  $k$  вызовов в час. Считая, что число вызовов на любом участке времени распределены по закону Пуассона, найти вероятность того, что за 2 минуты наступит ровно 3 вызова.

Среднее число вызовов за 2 минуты  $a = 2 \cdot \frac{k}{60} = \frac{k}{30}$ .

$$\text{Откуда } P(X = 3) = \frac{a^3}{m!} e^{-a} = \frac{\left(\frac{k}{30}\right)^3}{1 \cdot 2 \cdot 3} e^{-\frac{k}{30}}.$$

#### 1.9.4. Закон нормального распределения вероятностей (Гаусса)

Нормальный закон распределения (закон Гаусса) в теории вероятностей играет исключительно важную роль и занимает среди других законов распределения особое положение. Это — наиболее часто встречающийся на практике закон распределения.

Главная особенность, выделяющая его среди других законов, состоит в том, что он является *предельным* законом, к которому приближаются другие законы распределения при весьма часто встречающихся типичных условиях.



Нормальный закон распределения характеризуется плотностью вероятности вида:

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

график дифференциальной функции  $f(x)$  называют *нормальной кривой Гаусса* (рис. 5; получен исследованием  $f(x)$  методами дифференциального исчисления).

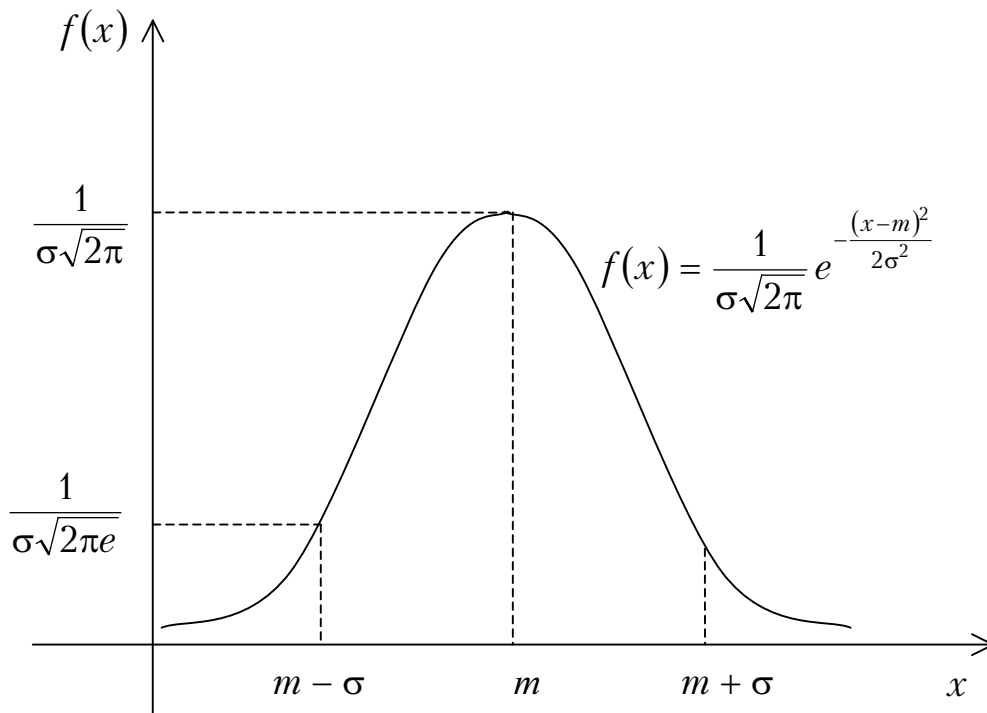


Рис. 5

Справедливо:

- $f(x)$  имеет максимум при  $x = m$ ;  $\max f(x) = \frac{1}{\sigma\sqrt{2\pi}}$ ;
- график функции симметричен относительно прямой  $x = m$  (так как  $(x - m)$  в  $f(x)$  входит в квадрате);
- $f(x) \geq 0$ ;
- $x = m \pm \sigma$ ;
- $x = m \pm \sigma$  - абсциссы точек перегиба,  $f(m \pm \sigma) = \frac{1}{\sigma\sqrt{2\pi}e}$ .

При  $m = 0$ ,  $\sigma = 1$

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

распределение называется *стандартным нормальным распределением*.

Часто вместо параметра  $\sigma$  вводят  $h = \frac{1}{\sigma\sqrt{2}}$ , который называют *мерой точности*.

Тогда

$$f(x) = \frac{h}{\sqrt{\pi}} e^{-h^2(x-m)^2}.$$

При нормальном распределении случайной величины  $X$  математическое ожидание  $M[X] = m$  (часто, особенно в задачах стрельбы, называется *центром рассеивания*);  $D[X] = \sigma^2$  ( $\sigma$  в  $f(x)$  есть среднеквадратическое отклонение  $X$ ).

Если изменить центр рассеивания  $m$ , кривая распределения будет смещаться вдоль оси абсцисс, не изменяя своей формы. Таким образом, центр рассеивания  $m$  характеризует положение распределения на оси абсцисс. Параметр  $\sigma$  характеризует не положение, а форму кривой Гаусса: при увеличении  $\sigma$  кривая «растягивается» вдоль оси абсцисс ( $\max f(x) = \frac{1}{\sigma\sqrt{2\pi}}$  обратно пропорционален  $\sigma$ , а площадь под кривой распределения всегда равна 1).

**Вероятность попадания случайной величины, подчиненной нормальному закону, на заданный участок  $[\alpha, \beta]$  определится в виде:**

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-m)^2}{2\sigma^2}} dx.$$

Сделаем замену переменной  $\frac{x-m}{\sigma} = t$ . Тогда  $x = \sigma t + m$ ,  
 $dx = \sigma dt$ .

При

$$x = \alpha \quad t = \frac{\alpha - m}{\sigma};$$

$$x = \beta \quad t = \frac{\beta - m}{\sigma}.$$

Получим

$$\begin{aligned}
 P(\alpha < X < \beta) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{\frac{\alpha-m}{\sigma}}^{\frac{\beta-m}{\sigma}} e^{-\frac{t^2}{2}} (2dt) = \\
 &= \frac{1}{\sqrt{2\pi}} \int_{\frac{\alpha-m}{\sigma}}^0 e^{-\frac{t^2}{2}} dt + \frac{1}{\sqrt{2\pi}} \int_0^{\frac{\beta-m}{\sigma}} e^{-\frac{t^2}{2}} dt = \\
 &= \frac{1}{\sqrt{2\pi}} \int_0^{\frac{\beta-m}{\sigma}} e^{-\frac{t^2}{2}} dt - \frac{1}{\sqrt{2\pi}} \int_0^{\frac{\alpha-m}{\sigma}} e^{-\frac{t^2}{2}} dt.
 \end{aligned}$$

Введя функцию Лапласа  $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$  (интеграл Пуассона  $\int e^{-x^2} dx$  в элементарных функциях не берется, поэтому для  $\Phi(x)$  составлены таблицы), будем иметь:

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - m}{\sigma}\right) - \Phi\left(\frac{\alpha - m}{\sigma}\right).$$

Аналогично интегральная функция  $F(x) = \int_{-\infty}^x f(x)dx$  представляется в виде

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

Справедливо (рис. 6):

$$P(m < X < m + \sigma) = \Phi\left(\frac{m + \sigma - m}{\sigma}\right) - \Phi\left(\frac{m - m}{\sigma}\right) = \Phi(1) - \Phi(0) \approx 0,341;$$

$$\begin{aligned}
 P(m + \sigma < X < m + 2\sigma) &= \Phi\left(\frac{m + 2\sigma - m}{\sigma}\right) - \Phi\left(\frac{m + \sigma - m}{\sigma}\right) = \\
 &= \Phi(2) - \Phi(1) \approx 0,136;
 \end{aligned}$$

$$\begin{aligned}
 P(m + 2\sigma < X < m + 3\sigma) &= \Phi\left(\frac{m + 3\sigma - m}{\sigma}\right) - \Phi\left(\frac{m + 2\sigma - m}{\sigma}\right) = \\
 &= \Phi(3) - \Phi(2) \approx 0,012.
 \end{aligned}$$

Из симметрии нормальной кривой относительно  $x = m$  следует

$$P(m - 3\sigma < X < m + 3\sigma) \approx 2(0,341 + 0,136 + 0,012) = 0,978.$$

Для оценки диапазона возможных значений нормально распределенной случайной величины получили следующее **«правило трех сигм»**: если  $X$  распределена по нормальному закону, то отклонение (по абсолютной величине) этой величины от математического ожидания не превосходит утроенного среднеквадратического отклонения, то есть

$$P(|X - m| < 3\sigma) \approx 1$$

(для приближенного определения  $\sigma$  максимально практически возможное отклонение от среднего следует разделить на три).

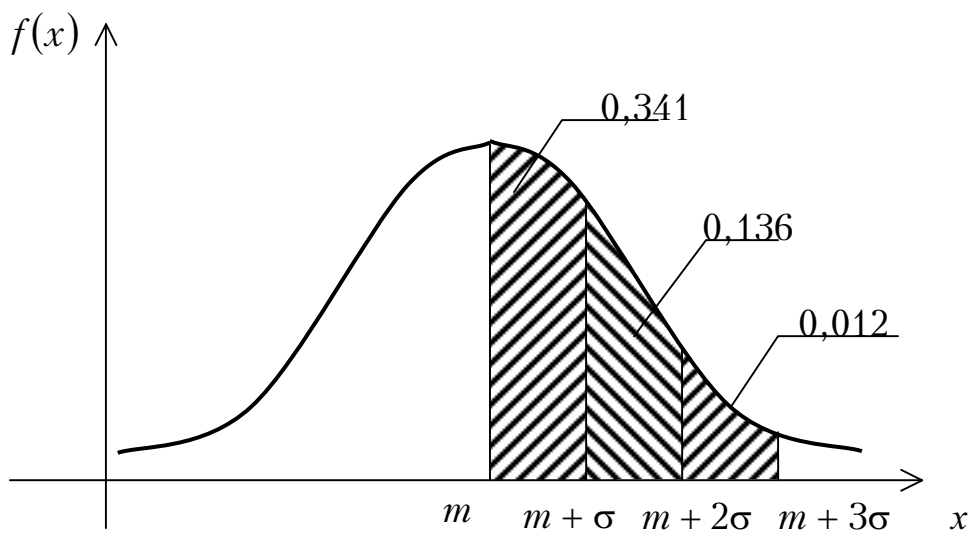


Рис. 6

### 1.10. Системы случайных величин

В практических задачах часто результат опыта описывается не одной случайной величиной, а системой случайных величин. Например, попадание снаряда в цель определяется абсциссой и ординатой на плоскости (в пространстве – тремя координатами).

*Функцией распределения двух случайных величин  $(X, Y)$*  называется вероятность совместного выполнения двух неравенств  $X < x, Y < y$ :

$$F(x, y) = P(X < x, Y < y).$$

Геометрически  $F(x, y)$  есть вероятность того, что случайная точка  $(X, Y)$  попадет в бесконечный квадрант с вершиной  $(x, y)$ , расположенный левее и ниже этой точки (рис. 7).

Вводя в рассмотрение плотность распределения случайной величины  $X$ , ее определяли как предел отношения вероятности попадания на малый участок к длине этого участка при ее неограниченном уменьшении. Аналогично *плотность распределения системы двух случайных величин  $(X, Y)$*  определяется как *предел отношения вероятности попадания случайной точки  $(X, Y)$  в малый прямоугольник  $R$  к площади этого прямоугольника, когда оба его размера стремятся к нулю.*

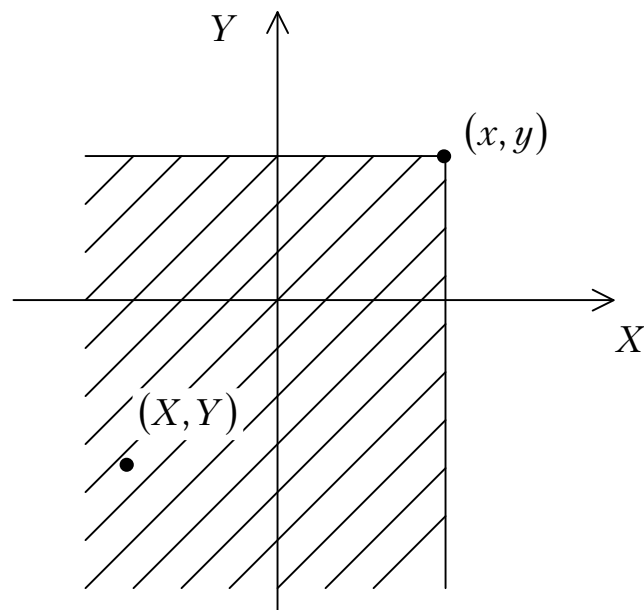


Рис. 7

Легко определить, что указанная плотность распределения

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = F''_{xy}(x, y).$$

Действительно,  $(X, Y) \subset R$  равносильно произведению событий  $\alpha \leq X \leq \beta$  и  $\gamma \leq Y \leq \delta$ . Вероятность попадания в прямоугольник  $R$  определится в виде

$$P((X, Y) \subset R) = F(\beta, \delta) - F(\alpha, \delta) - F(\beta, \gamma) + F(\beta, \gamma).$$

Полагая  $\alpha = x$ ,  $\gamma = y$ ,  $\delta = y + \Delta y$ ,  $\beta = x + \Delta x$  получим:

$$\begin{aligned} P(x \leq X \leq x + \Delta x, y \leq Y \leq y + \Delta y) &= \\ &= F(x + \Delta x, y + \Delta y) - F(x, y + \Delta y) - F(x + \Delta x, y) + F(x, y). \end{aligned}$$

Тогда, если  $F(x, y)$  не только непрерывна, но и дифференцируема, получим:

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{P(x \leq X \leq x + \Delta x, y \leq Y \leq y + \Delta y)}{\Delta x \cdot \Delta y} = \frac{\partial^2 F(x, y)}{\partial x \partial y}.$$

Укажем основные свойства функции распределения.

1. При одном из аргументов, равным  $+\infty$ , функция распределения системы превращается в функцию распределения случайной величины, соответствующей другому аргументу:

$$F(x, +\infty) = F_1(x),$$

$$F(+\infty, y) = F_2(y),$$

где  $F_1(x)$ ,  $F_2(y)$  – функции распределения случайных величин  $X$  и  $Y$  соответственно;

$$F_1'(x) = f_1(x) = \int_{-\infty}^{\infty} f(x, y) dx, \quad F_2'(y) = f_2(y) = \int_{-\infty}^{\infty} f(x, y) dy.$$

2.  $F(-\infty, +\infty) = 1$ .

3.  $F(x, -\infty) = F(-\infty, y) = F(-\infty, -\infty) = 0$ .

4.  $F(x, y)$  – неубывающая функция.

5.  $F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy$ .

6.  $f(x, y) \geq 0$ , так как плотность распределения  $f(x, y)$  есть предел отношения двух неотрицательных величин (вероятности попадания в прямоугольник и площади прямоугольника).

7.  $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$  (вероятность попадания на всю плоскость

$XOY$  – достоверное событие).

В общем случае по заданному распределению отдельных величин, входящих в систему, нельзя определить закон распределения системы. Для этого надо знать так называемый *условный закон распределения*.

**Условным законом распределения случайной величины  $X$ , входящей в систему  $(X, Y)$ , называется закон ее распределения при условии, что другая случайная величина  $Y$  приняла определенное значение** (обозначается  $F(X | Y)$ ), а соответствующая ей условная плотность распределения –  $f(x | y)$ ).

Справедливо соотношение (теорема умножения законов распределения)

$$f(x, y) = f_1(x) \cdot f(y | x)$$

$$\left( f(y | x) = \frac{f(x, y)}{f_1(x)} = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y) dy}, \quad f(x | y) = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y) dx} \right).$$

Случайная величина  $Y$  называется **независимой** от случайной величины  $X$ , если закон распределения  $Y$  не зависит от того, какое значение приняла  $X$ .

Для непрерывных случайных величин  $X$  и  $Y$  это означает, что  $f(y | x) = f_2(y)$ .

Очевидно, если  $Y$  не зависит от  $X$ , то и  $X$  не зависит от  $Y$ , то есть  $f(x | y) = f_1(x)$ .

Из предыдущего следует, что случайные величины  $X$  и  $Y$  будут **независимыми**, если закон распределения каждой из них не зависит от того, какое значение приняла другая.

Для независимых случайных величин:

$$f(x, y) = f_1(x) \cdot f_2(y).$$

Верно и обратное, если  $f(x, y) = f_1(x) \cdot f_2(y)$ , то  $X$  и  $Y$  независимы.

### 1.11. Числовые характеристики системы двух случайных величин

Для описания системы двух случайных величин, кроме математических ожиданий и дисперсий составляющих, пользуются и другими характеристиками, к числу которых относятся *корреляционный момент и коэффициент корреляции*.

Корреляционным моментом (или моментом связи)  $K_{xy}$  случайных величин  $X$  и  $Y$  называют математическое ожидание произведения центрированных случайных величин  $\overset{\circ}{X} = (X - M[X])$  и  $\overset{\circ}{Y} = (Y - M[Y])$ , то есть  $K_{xy} = M\left[\overset{\circ}{X} \overset{\circ}{Y}\right] = M[(X - M[X]) \cdot (Y - M[Y])]$ .

При *дискретных*  $X$  и  $Y$  для вычисления корреляционного момента пользуются формулой:

$$K_{xy} = \sum_i \sum_j (X_i - m_x)(Y_j - m_y) p_{ij}, \quad p_{ij} = P(X = x_i, Y = y_j).$$

Для *непрерывных* величин:

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

Корреляционный момент двух независимых случайных величин равен нулю.

Действительно, из независимости  $X$  и  $Y$  следует независимость  $(X - M[X])$  и  $(Y - M[Y])$ . Будем иметь:

$$\begin{aligned} K_{xy} &= M[(X - M[X]) \cdot (Y - M[Y])] = M[X - M[X]] \cdot M[Y - M[Y]] = \\ &= (M[X] - M[M[X]]) \cdot (M[Y] - M[M[Y]]) = 0. \end{aligned}$$

Если  $X$  мало отличается от  $M[X]$  (т.е. она почти не случайна), то корреляционный момент будет мал. Так что корреляционный момент характеризует не только зависимость величин, но и их рассеивание.

Размерность корреляционного момента равна произведению размерностей  $X$  и  $Y$ , то есть зависит от единиц их измерения. Чтобы избавиться от этого недостатка вводят безразмерную характеристику – коэффициент корреляции:

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y},$$

где  $\sigma_x, \sigma_y$  – среднеквадратические отклонения  $X$  и  $Y$  ( $D_x, D_y$  характеризуют рассеивание  $X$  и  $Y$  соответственно).

В силу предыдущего для независимых величин  $X$  и  $r_{xy} = 0$ .

Если  $K_{xy} \neq 0$ , то случайные величины  $X$  и  $Y$  называются коррелированными.

Две коррелированные величины также зависимы. Обратное не всегда верно. Для двух зависимых величин возможно как  $K_{xy} = 0$ , так и  $K_{xy} \neq 0$ .

### 1.12. Оценки для числовых характеристик системы случайных величин

Нахождение *точечных* оценок для числовых характеристик  $m_x, m_y, D_x, D_y, K_{xy}$  системы случайных величин  $(X, Y)$  производится аналогично тому, как это осуществлялось для одной случайной величины. Так, если в  $n$  независимых опытах над системой случайных величин  $(X, Y)$  получены значения  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , то – несмещенные оценки математических ожиданий  $X$  и  $Y$  есть

$$\tilde{m}_x = \frac{\sum_{i=1}^n X_i}{n}, \quad \tilde{m}_y = \frac{\sum_{i=1}^n Y_i}{n};$$