МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
**Рубцовский индустриальный институт (филиал)**
федерального государственного бюджетного образовательного
учреждения высшего образования
«Алтайский государственный технический университет им. И.И. Ползунова»
(РИИ АлтГТУ)

**О.В. ЕФРЕМЕНКОВА**

# МАТЕМАТИЧЕСКИЕ ОСНОВЫ ТЕОРИИ НАДЕЖНОСТИ

Учебное пособие
для студентов всех форм обучения направления
«Строительство»

***Рекомендовано Рубцовским индустриальным институтом (филиалом) ФГБОУ ВО «Алтайский государственный технический университет им. И.И. Ползунова» в качестве учебного пособия для студентов, обучающихся по направлению подготовки «Строительство»***

Рубцовск 2016

УДК 517.8

Ефременкова О.В. Математические основы теории надежности: Учебное пособие для студентов всех форм обучения направления «Строительство» / Рубцовский индустриальный институт. – Рубцовск, 2016. - 71с.

Данное учебное пособие предназначено для студентов всех форм обучения, направления «Строительство». В пособии представлен широкий спектр как теоретического материала, так и задач прикладной математики. Содержание задач затрагивает разные отрасли практических знаний, с которыми будущий инженер-электрик может столкнуться в своей профессиональной деятельности.

Рассмотрено и одобрено на заседании
НМС Рубцовского индустриального
института.
Протокол № 5 от 22.09.2016 г.

Рецензент: к.т.н. О.А. Михайленко

СОДЕРЖАНИЕ

# Глава 1. ПОНЯТИЕ НАДЕЖНОСТИ И ЕЕ СВОЙСТВА

Обычный детерминистический подход к расчету конструкций состоит из двух этапов:

1) Вычисляются напряжения, деформации и перемещения в конструкциях, подверженных действию внешних нагрузок. Эта задача решается методами строительной механики, теории упругости, теории пластичности и т.д.

2) Вычисленные величины сопоставляются с нормативно допустимыми значениями. При этом решается задача надежности, долговечности и экономичности конструкции.

Однако реальная система и ее условия эксплуатации отличаются от идеализированной системы и условий, рассматриваемых на стадии проектирования. Фактически напряжения, деформации и перемещения являются случайными величинами из-за случайного характера внешних воздействий, прочностных и др. внешних условий. Поэтому надежность конструкции может быть определена с привлечением методов математической и статистической теории вероятностей.

В теории вероятностей главная задача - зная состав генеральной совокупности, изучить распределения для состава случайной выборки. Это прямая задача теории вероятностей. Обратная задача - когда известен состав выборки и по нему требуется определить, какой была генеральная совокупность. Это обратная задача математической статистики. Или, точнее, в теории вероятностей мы, зная природу некоторого явления, выясняем, как будут вести себя (как распределены) те или иные изучаемые нами характеристики, которые можно наблюдать в экспериментах. В математической статистике наоборот – исходными являются экспериментальные данные (обычно это наблюдения над случайными величинами), а требуется вынести то или иное суждение о природе рассматриваемого явления.

*Надежность* – свойство объекта сохранять во времени в установленных пределах значения всех параметров, характеризующих способность выполнять требуемые функции в заданных режимах и условиях применения, технического обслуживания, ремонта и транспортирования. Или *надежность* также – устойчивость качества по отношению ко всем возможным возмущениям. Надежность – количественный показатель (промежуток времени, число рабочих циклов, число километров и т.д.).

В зависимости от назначения системы и условий ее эксплуатации надежность включает свойства: 1) безотказность; 2) долговечность; 3) ремонтопригодность; 4) сохраняемость и любые их сочетания.

*Безотказность* – вероятность безотказной работы конструкции за определенный промежуток времени.

*Долговечность* – вероятный промежуток времени безотказной работы конструкции.

*Ремонтопригодность* – вероятность того, что неисправная система может быть восстановлена за заданное время.

Содержание теории надежности – разработка методов оценки надежности систем и создание систем, обладающих заданными показателями надежности и долговечности.

Задачи расчета на надежность состоят в определении вероятности выхода конструкции из строя в заданных условиях, нахождении по заданной экономически целесообразной надежности требуемых размеров конструкции, допустимых нагрузок или оптимального срока эксплуатации, а также оценки надежности системы по имеющимся оценкам надежности составляющих ее элементов. В задачу теории надежности строительных конструкций входит также обоснование процедур нормирования расчетных характеристик. Специфика теории надежности строительных конструкций состоит в необходимости учета случайных свойств нагрузок и воздействий на сооружения, а также учета совместного действия случайных нагрузок на систему со случайными прочностными характеристиками.

Основное понятие теории надежности – *отказ* – событие, состоящее в нарушении работоспособности системы. Понятие отказа близко по смыслу к понятию предельного состояния. К предельным состояниям 1-й группы относятся: общая потеря устойчивости формы, потеря устойчивости положения, любое разрушение, переход в изменяемую систему, качественное изменение конфигурации; состояния, при которых возникает необходимость прекращения эксплуатации в результате текучести материала, сдвига в соединениях, ползучести или чрезмерного раскрытия трещин. Предельные состояния 2-й группы – недопустимые деформации конструкций в результате прогиба, поворота или осадок, характеризуемых разностью вертикальных перемещений узлов, отнесенных к расстоянию между ними, креном сооружения в целом, относительным прогибом или выгибом, кривизной элемента, относительным углом закручивания, горизонтальным или вертикальным смещением элемента или сооружения в целом, углом перекоса или поворота. К предельным состояниям 2-й группы относятся также недопустимые колебания конструкции, изменение положения, образование или раскрытие трещин.

Примеры отказов - обрушения, опрокидывания, потеря устойчивости, хрупкое разрушение, большие деформации и прогибы, механический или коррозионный износ, растрескивание и т.д.

Отказы вызваны влиянием случайных факторов, поэтому они носят случайный характер. За показатель (меру) надежности системы может быть принята вероятность $P$ безотказной работы в течение всего срока службы $T$.

Недостатки теории надежности - сложно получить опытные данные в количестве, достаточном для последующей их обработки методами теории вероятностей. Сложно длительный срок проводить испытания конструкции для получения надежных выводов о ее долговременной работе.

Безотказность (и другие составляющие свойства надежности) систем обработки информации и управления проявляется через случайные величины, например, наработку до очередного отказа и количество отказов за заданное время.

Для описания этих случайных величин используют стандартные распределения случайных величин, рассматриваемых в теории вероятностей. Для каждого распределения рассматриваются четыре основные характеристики:

- функция распределения $F(t)$;
- плотность распределения $f(t)$;
- математическое ожидание (средняя наработка до отказа) $T_0$;
- дисперсия $DT_0$.

Наработка до очередного отказа есть продолжительность или объем работы объекта.

Для систем обработки информации и управления естественно исчисление наработки в единицах времени, тогда как для других технических средств могут быть удобнее иные единицы измерения (например, наработка автомобиля - в километрах пробега).

Для невосстанавливаемых и восстанавливаемых изделий понятия наработки различаются:

- в первом случае подразумевается наработка до первого отказа (он же является и последним отказом),
- во втором случае – наработка между двумя соседними во времени отказами (после каждого отказа производится восстановление работоспособного состояния).

Вероятностные характеристики наработки до очередного отказа и являются показателями безотказности объекта.

Эти вероятностные характеристики определяются по результатам наблюдений за некоторым множеством экземпляров однотипных изделий, но используются в качестве показателя надежности каждого конкретного изделия.

Средняя наработка до отказа (между отказами) есть математическое ожидание наработки до очередного отказа (интегрирование осуществляется в пределах от 0 до ∞):

$$M[T] = \int t f(t) dt = T_0,$$

где через $t$ обозначено текущее значение наработки,

а $f(t)$ - плотность вероятности ее распределения.

Вероятностью безотказной работы $P(t)$ называют вероятность того, что изделие будет работоспособно в течение заданной наработки при заданных условиях эксплуатации.

По статистическим данным об отказах вероятность безотказной работы определяют по формуле:

$$P(t) = [N(0) - n(t)]/(N(0),$$

где N(0) – число изделий в начале наблюдений;

$n(t)$ – число изделий, отказавших за время $t$.

В начальный момент времени $P(0) = 1$. При увеличении времени вероятность $P(t)$ монотонно уменьшается и для любых технических изделий асимптотически приближается к нулю.

Вероятность отказа $F(t)$ есть вероятность того, что при заданных условиях эксплуатации в течение заданной наработки произойдет отказ.

Отказ и безотказная работа – противоположные события. Поэтому вероятность $F(t)$ отказа можно найти по формуле:

$$F(t) = 1 - P(t).$$

Частота отказов есть плотность распределения времени безотказной работы изделия:

$$f(t) = d\,F(t)/dt = d[1 - P(t)]\,/dt = -\,d\,P(t)/dt,$$

откуда очевидно, что частота отказов характеризует скорость уменьшения вероятности безотказной работы во времени.

Интенсивностью отказов $\lambda(t)$ называют условную плотность вероятности возникновения отказа изделия при условии, что к моменту $t$ отказ не возник:

$$\lambda(t) = f(t)/P(t) = -\,[1/P(t)]dP(t)/dt.$$

Функции $\lambda(t)$ и $f(t)$ измеряются в $ч^{-1}$.

Интегрируя $\lambda(t)$ в пределах от 0 до $t$, можно получить:

$$P(t) = e^{-\int \lambda(t)dt}.$$

Это выражение, называемое основным законом надежности, позволяет установить временное изменение вероятности безотказной работы при любом характере изменения интенсивности отказов во времени.

В частном случае постоянства интенсивности отказов $\lambda(t) = \lambda = \text{const}$ равенство переходит в известное в теории вероятностей экспоненциальное распределение:

$$P(t) = e^{-\lambda t}; \qquad F(t) = 1 - e^{-\lambda t}; \qquad f(t) = \lambda e^{-2t}.$$

Экспоненциальное распределение однопараметрическое и обладает уникальными свойствами:

- вероятность безотказной работы изделия $P(t) = e^{-\lambda t}$ не зависит от того, сколько времени изделие проработало до рассматриваемого интервала времени;
- средняя остаточная (предстоящая) наработка до отказа $T_0$ также не зависит от того, сколько времени проработало изделие ранее.

Эти закономерности являются проявлением свойства, называемого отсутствием последствия.

В этом случае показатели надежности изделия зависят только от состояния изделия в начале рассматриваемого интервала времени, но не зависят от наработки до этого интервала времени.

Существуют физические предпосылки, объясняющие это свойство. Одно из объяснений. Любое изделие работает в условиях определенной нагрузки (электрической, механической и пр.) и имеет ограниченную «прочность», поэтому существует некоторая предельная нагрузка, которую изделие способно выдержать без отказа. Если же нагрузка превосходит предельное значение, то наступает внезапный отказ. Пиковые значения нагрузок возникают случайным образом. В теории случайных процессов доказывается, что при определенных условиях время до первого пересечения случайным процессом некоторого порогового уровня имеет как раз экспоненциальное распределение.

В приведенном объяснении важным является то, что отказ возникает не вследствие постепенного изменения внутреннего состояния изделия, а вследствие внешнего воздействия, значение которого превышает допустимое.

Отсюда следует, что при экспоненциальном распределении наработки до отказа профилактические работы, включающие в себя замену элементов или их периодический ремонт, теряют всякий смысл, так как не могут повлиять на причину отказа.

Естественный путь повышения надежности изделия состоит либо в его конструктивном улучшении, либо в снижении действующих нагрузок.

Поток отказов при постоянстве интенсивности отказов $\lambda(t)$=const называется простейшим потоком отказов, и именно он реализуется для большинства систем обработки информации и управления в течение периода нормальной эксплуатации от окончания приработки до начала старения и износа.

Подставив выражение плотности вероятности $f(t) = \lambda e^{-\lambda t}$ экспоненциального распределения, получим:

$$T_0 = 1 / \lambda,$$

т. е. при простейшем потоке отказов средняя наработка $T_0$ обратно пропорциональна интенсивности отказов $\lambda$ .

Можно показать, что за время средней наработки, $t = T_0$ , вероятность безотказной работы изделия составляет 1/e.

## Глава 2. ТЕОРИЯ ВЕРОЯТНОСТИ

События $A, B, C,...$ несовместные, если появление одного из них исключает появление других событий в одном и том же испытании.

События образуют полную группу событий, если в результате испытаний появится хотя бы одно из них.

Испытание. Бросание монеты: $A$ – герб, $B$ – цифра;

Испытание. Выстрел: $A$ - попадание, $\overline{A}$ – промах (противоположное событие). $A$ и $\overline{A}$ - образуют полную группу событий при одном выстреле.

Испытание. Покупают 23 лотерейных билета: событие $A$ – выигрыш выпал на первый билет, отсутствие выигрыша на втором; событие $B$ – выигрыш не на первом, а на втором билете; событие $C$ – выигрыш обоих билетов; событие $D$ – выигрыш не выпал ни на один билет.

Данные события образуют полную группу событий, если проверяются два билета. А при проверке трех билетов что образует полную группу событий?

Испытание - игральная кость: событие $A$ – появление числа 1, событие $B \Rightarrow 2$, событие $C \Rightarrow 3$, $D \Rightarrow 4$, $E \Rightarrow 5$, $F \Rightarrow 6$. События равновероятны?

Вероятность – мера объективной возможности появления события

Классическое определение вероятности:

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

где $m$ – число исходов испытания, которое благоприятствует появлению события $A$, $n$ – общее число исходов испытания.

Свойства вероятности:

1) Вероятность достоверного события равна 1: $P(A) = \frac{n}{n} = 1.$

2) Вероятность невозможного события равна 0: $P(A) = \frac{0}{n} = 0.$

3) Вероятность случайного события есть положительное число, заключенное между 0 и 1: $P(A) = \frac{m}{n}$, $0 \le m \le n$, $\Rightarrow 0 \le P(A) \le 1.$

«Судебная машина» Лапласа. Парадокс кавалера де Грие.
Аксиоматика Колмогорова

Теоретико-множественный подход к определению вероятности

$\Omega$ - некоторое пространство элементарных событий, в котором с исходами испытаний связывают точки $\Omega_i$ пространства. Каждому $\Omega_i$ ставят в соответствие $p_i$.

***Пример 2.1***. Для игральной кости: $i = 1, 2, 3, 4, 5, 6$. Событие $A$ – подмножество пространства элементарных событий - появление четного числа (событие): $A = \{2, 4, 6\}$.

Недостатки классического определения вероятности

1. Неприменимость при бесконечном числе исходов $(n \to \infty)$.

Выход находят путем введения геометрической вероятности: как отношение мер длин, площадей. Примеры: рулетка, попадание точки в отрезок длиной $l$ на отрезке длиной $L$: $P(A) = \frac{l}{L}$.

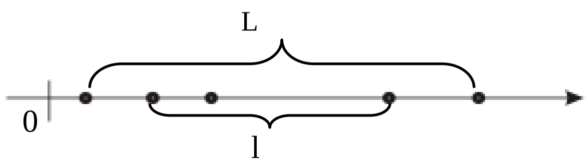



Рис. 2.1

2. Априори трудно представить результат испытаний в виде совокупности элементарных событий, еще труднее указать основание, позволяющее считать элементарные события равновозможными.

Тогда вводят статистическую (апостериорную) вероятность

$$p^*(A) = \frac{m}{n},$$

где $m$ – число исходов, в которых событие появилось,
$n$ – общее число исходов

$$P^*(A) \xrightarrow[n \to \infty]{p} P(A).$$

Алгебра событий

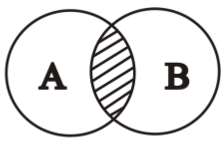
1. Суммой двух событий $A$ и $B$ называют событие $(A+B)$, состояще в появлении или $A$, или $B$, или обоих (если они совместны).

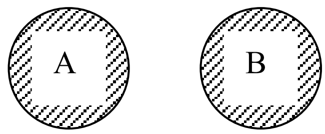
***Пример 2.2.*** Для двух выстрелов сумма событий: $A$ – попадание при первом выстреле, $B$ – попадание при втором выстреле, $AB$ – попадание при обоих выстрелах, т.е. $A+B=A+B+A\cdot B$.

Суммой нескольких событий называют событие, заключающееся в появлении хотя бы одного из них. $C=A+B+C+AB+AC+BC+ABC$.

При несовместных событиях: $A+B=A+B,\ C=A+B+C+AB+AC+BC+ABC$.

Диаграммы Эйлера-Венна




$P(A+B)=P(A)+P(B)-P(AB)$ $\qquad$ $P(A+B)=P(A)+P(B)$

***Пример 2.3.*** В урне 30 шаров: 10 красных, 5 синих, 15 белых. Найти вероятность появления цветного шара при вынимании из урны одного шара. С появлением цветного шара свяжем или событие $A$ - вынимание красного шара, или событие $B$ – вынимание синего шара.

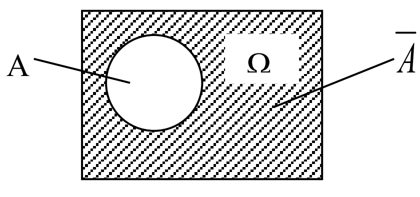
$$P(A+B)=P(A)+P(B)\Rightarrow P(A)=\frac{10}{30},\ P(B)=\frac{5}{30}\Rightarrow P(A+B)=\frac{1}{3}+\frac{1}{6}=\frac{1}{2}.$$

Сумма вероятностей событий $A_1, A_2, ..., A_n$, образующих полную группу, равна 1 (условие полноты группы событий).

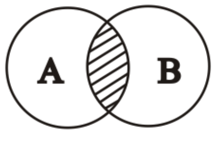
$$P(A_1)+P(A_2)+\ldots+P(A_n)=1.$$

2. Противоположными называют два единственно возможных события, образующих полную группу.

$$P(A)+P(\overline{A})=1,\quad P(\overline{A})=1-P(A).$$




3. Произведением двух событий $A$ и $B$ называют событие $A\cdot B$, состоящее




в совместном появлении этих событий.

***Пример 2.4.*** $A$ – деталь годна, $B$ – деталь окрашена$\Rightarrow A\cdot B$ – деталь годная и окрашенная.

Произведением нескольких событий называют событие, состоящее в совместном появлении всех этих событий.

***Пример 2.5.*** Событие $A$ – появление герба при первом броске монеты, $B$ - появление герба при втором броске, $C$ - появление герба при третьем броске, $A \cdot B \cdot C$ – появление герба при всех трех бросках.

4. Условной вероятностью $P(A|B)$ (или $P_B(A)$) называют вероятность события $A$, вычисленную в предположении, что событие $B$ наступило.

***Пример 2.6.*** В урне 3 белых шаров и 3 черных шара. Из урны дважды вынимают по одному шару, не возвращая обратно. Найти вероятность появления белого шара при втором испытании (событие $B$), если при первом испытании (событие $A$) был извлечен черный шар. $P(B\,|\,A) = 3/5,\ P(B) = 1/2 \Rightarrow P(B\,|\,A) \neq P(B)$.

События $A$ и $B$ в этом примере зависимые.

$P(B)$ - безусловная вероятность, а $P(B\,|\,A)$ - условная вероятность.

5. Формула вероятности произведения двух зависимых событий.

$$P(AB) = P(A)\cdot P(B\,|\,A) = P(B)\cdot P(A\,|\,B) \Rightarrow\ P(B\,|\,A) = \frac{P(AB)}{P(A)} \text{ и } P(A\,|\,B) = \frac{P(AB)}{P(B)}.$$

6. Формула вероятности произведения двух независимых событий:

$P(AB) = P(A)\cdot P(B) \Rightarrow$ Условие независимости: $P(AB) = P(A)\cdot P(B)$.

7. Формула вероятности произведения нескольких зависимых событий:

$$P(A_1, A_2, A_3, \ldots, A_n) = P(A_1)\cdot P(A_2\,|\,A_1)\cdot P(A_3\,|\,A_1A_2)\cdot\ldots\cdot P(A_n\,|\,A_1A_2\ldots A_{n-1}).$$

***Пример 2.7.*** В урне 5 белых шаров, 4 черных и 3 синих шара. Наудачу извлекают по одному шару, не возвращая обратно. Найти вероятность того, что при первом испытании появится белый шар (событие А), при втором – черный (событие В), при третьем–синий (событие *С*).

$P(A) = 5/12,\ P(B\,|\,A) = 4/11,\ P(C\,|\,AB) = 3/10 \Rightarrow P(A\cdot B\cdot C) = (5/12)(4/11)(3/10) = 1/22.$

*Замечания:*

1. Порядок, в котором расположены события, может быть выбран любым, т.е. безразлично, какие события можно считать первыми, вторыми и т.д.

2. Независимость событий взаимна: если $A$ не зависит от $B$, то и $B$ не зависит от $A$.

3. Формула вероятности произведения нескольких независимых событий:

$P(A_1, A_2, A_3 \ldots A_n) = P(A_1)\cdot P(A_2)\cdot\ldots\cdot P(A_n)$, $P(ABC) = P(A)\cdot P(B)\cdot P(C)$.

***Пример 2.8***. Вероятности появления каждого из трех событий $A_1, A_2, A_3$ равны: $P(A_1) = p_1, P(A_2) = p_2, P(A_3) = p_3$. Найти вероятность появления только одного из этих событий.

$$B_1 = A_1\overline{A_2}\,\overline{A_3},\ B_2 = \overline{A_1}A_2\overline{A_3},\ B_3 = \overline{A_1}\,\overline{A_2}A_3,\ P(B_1 + B_2 + B_3) =$$
$$= P(B_1) + P(B_2) + P(B_3) = p_1(1-p_2)(1-p_3) + (1-p_1)P_2(1-p_3) + (1-p_1)(1-p_2)p_3.$$

***Пример 2.9.*** Вероятности попадания в цель при стрельбе из трех орудий равны: $p_1 = 0{,}8;\ p_2 = 0{,}7;\ p_3 = 0{,}9.$ Найти вероятность хотя бы одного попадания (событие *A*) при одном залпе из всех орудий. Событие $A_1$ – попадание первым орудием, $A_2$ – попадание вторым орудием, $A_3$ – попадание третьим орудием. $P(A_1) = 0{,}8;\ P(A_2) = 0{,}7;\ P(A_3) = 0{,}9$. Прямое событие

$$A = A_1\overline{A_2}\,\overline{A_3} + \overline{A_1}A_2\overline{A_3} + \overline{A_1}\,\overline{A_2}A_3 + A_1A_2\overline{A_3} + A_1\overline{A_2}A_3 + \overline{A_1}A_2A_3 + A_1A_2A_3,$$

Обратное событие $\overline{A}$ (не попал ни разу) определится следующей алгеброй $\overline{A} = \overline{A_1}\,\overline{A_2}\,\overline{A_3}$ и вероятностями

$$P(\overline{A}) = (1 - p_1)(1 - p_2)(1 - p_3) =$$
$$= (1 - 0,8)(1 - 0,7)(1 - 0,9) = 0,006;\ P(A) = 1 - P(\overline{A}) = 1 - 0,006 = 0,994.$$

***Пример 2.10.*** Вероятность того, что при одном выстреле стрелок попадает в цель, равна 0,4. Сколько выстрелов должен произвести стрелок, чтобы с вероятностью не менее 0,9 попал в цель хотя бы один раз. Пусть событие $A$ – при $n$ – выстрелах стрелок попадает в цель хотя бы один раз.

$$P(A) \ge 0,9;\ n - ? \Rightarrow \begin{cases} P(\overline{A}) = 1 - (1 - P)^n, \\ P(A) \ge 0,9; \end{cases} \Rightarrow$$
$$\Rightarrow 1 - (0,6)^n \ge 0,9; \Rightarrow$$
$$\Rightarrow (0,6)^n \le 0,1 \Rightarrow$$
$$\Rightarrow n \ge 4,5;\ n \ge 5.$$

Вывод: стрелок должен произвести не менее пяти выстрелов.

## Формула полной вероятности

Пусть событие $A$ может наступить при условии появления одного из несовместных событий $H_1, H_2, ..., H_n$, называемых гипотезами и образующих полную группу событий. При этом вероятности $P(H_i)$ и $p(A \mid H_i)$ будем считать известными.

Тогда справедлива формула полной вероятности $P(A) = \sum\limits_{i=1}^{n} P(H_i) \cdot P(A \mid H_i)$.

Формула полной вероятности – «удобная схема» (форма) расчета вероятности событий.

***Пример 2.11.*** Имеются 2 набора деталей. Вероятность того, что деталь первого набора стандартна, равна 0,8, а второго – 0,9. Найти вероятность того, что взятая наудачу деталь из наудачу взятого набора – стандартна. Обозначим через $A$ событие – извлеченная деталь стандартна. В качестве гипотез удобно принять события $H_1$ - взята деталь из первого набора, $H_2$ - взята деталь из второго набора. Эти события несовместны (берут деталь один раз), образуют полную группу событий (деталь берут) и в данном случае равновероятны (набор выбирается наудачу): $P(H_1) = \frac{1}{2},\ P(H_2) = \frac{1}{2},\ P(H_1) + P(H_2) = 1.$ Выбор того или иного набора – условие. В различных наборах вероятность извлечения стандартной детали различна: $P(A \mid H_1) = 0,8;\ P(A \mid H_2) = 0,9.$ Тогда рассмотрим сумму двух событий, каждое из которых, в свою очередь, состоит из произведения, т.е. формулу полной вероятности $P(A) = P(H_1) \cdot P(A \mid H_1) + P(H_2) \cdot P(A \mid H_2) = \frac{1}{2} \cdot 0,8 + \frac{1}{2} \cdot 0,9 = 0,85.$

***Пример 2.12.*** В первом наборе 20 деталей и из них 18 - стандартны. Во втором наборе 10 деталей и из них 9 - стандартны. Из второго набора наудачу взята деталь и переложена в первый. Найти вероятность того, что деталь, наудачу извлеченная из первого набора, будет стандартна. Введем обозначения для событий: $A$ – из первого набора извлечена стандартная деталь, $H_1$ – из второго набора извлечена стандартная деталь, $H_2$ – из второго набора извлечена нестандартная деталь. Можно ли $H_1$ и $H_2$ - считать гипотезами?

$$P(H_1)=\frac{9}{10},\ P(H_2)=\frac{1}{10},\ P(A\,|\,H_1)=\frac{19}{21},\ P(A\,|\,H_2)=\frac{18}{21}\Rightarrow$$

$$P(A)=P(H_1)\cdot P(A\,|\,H_1)+P(H_2)\cdot P(A\,|\,H_2)=\frac{9}{10}\cdot\frac{19}{21}+\frac{1}{10}\cdot\frac{18}{21}=0,9.$$

## Формула вероятности гипотез

До сих пор рассматривалась - априорная, безусловная вероятность гипотез $P(H_i)$. Опытные данные (реальные события) могут уточнить априорную характеристику объекта анализа, определить $P(H_i\,|\,A)$ - условную (апостериорную) вероятность гипотезы. По формуле произведения вероятности событий можем записать:

$$P(AH_i)=P(A)P(H_i\,|\,A)=P(H_i)\cdot P(A\,|\,H_i).$$

Откуда выразим условную апостериорную (условие - событие $A$ произошло) вероятность

$$P(H_i\,|\,A)=\frac{P(H_i)\cdot P(A\,|\,H_i)}{P(A)}.$$

Знаменатель раскроем по формуле полной вероятности и получим формулу вероятности гипотез (формулу Байеса, лежащую в основе известного «байесовского подхода» в уточнении гипотез):

$$P(H_i\,|\,A)=\frac{P(H_i)\cdot P(A\,|\,H_i)}{\sum_i P(H_i)\cdot P(A\,|\,H_i)}.\ \text{Например, }\ P(H_1\,|\,A)=\frac{P(H_1)\cdot P(A\,|\,H_1)}{\sum_i P(H_i)\cdot P(A\,|\,H_i)}.$$

***Пример 2.13.*** Детали, изготовленные цехом, попадают для проверки на стандартность к одному из контролеров. Вероятность того, что деталь попадет к первому контролеру, равна 0,6, а ко второму – 0,4 (их загрузка или производительность). Вероятность того, что годная деталь будет признана стандартной первым контролером, – 0,94, а вторым – 0,98 (второй имеет лучшую квалификацию). Годная деталь при проверке была признана стандартной. Найти вероятность того, что годную деталь проверил первый контролер.

Введем обозначения: $A$ – деталь признана стандартной (событие произошло); гипотезы: $H_1$ – деталь проверил первый контролер, $H_2$ – деталь проверил второй контролер: $P(H_1)=0,6;\ P(H_2)=0,4.$ Условные вероятности

$P(A/H_1)=0,94;\ P(A/H_2)=0,98.$ $P(H_1/A)=\dfrac{0,6\cdot 0,94}{0,6\cdot 0,94+0,4\cdot 0,98}=0,59.$ Сравним априорную и апостериорную вероятности гипотез

$$P(H_1)=0,6;\ P(H_1\,|\,A)=0,59\Rightarrow P(H_1)\neq P(H_1\,|\,A).$$

## Формула Бернулли

Часто практике соответствует схема независимых испытаний Бернулли: проводятся испытания, в которых вероятность появления события $A$ («успеха») одна и та же, а исходы независимы друг от друга.

Задача: определить вероятность того, что при $n$ испытаниях событие $A$ произойдет $k$ раз (не произойдет $n-k$ раз). При этом не требуется, чтобы событие $A$ повторялось ровно $k$ раз в определенной последовательности. Например, при $n=4$ один «успех» может быть реализован следующим образом: $\Rightarrow AAA\bar{A},\ AA\bar{A}A,\ A\bar{A}AA,\ \bar{A}AAA$.

Для общего случая вероятность $k$ «успехов» из $n$ испытаний равна

$$P_n(k)=C_n^k\cdot p^k\cdot q^{n-k},$$

где $C_n^k$ - число сочетаний из $n$ по $k$, $C_n^k=\dfrac{n!}{k!(n-k)!};\ q=1-p.$

***Пример 2.14.*** Вероятность того, что операционные расходы фирмы в течение 1 месяца не превысят установленный бюджет, равна 0,75. Найти вероятность того, что в ближайшие 6 месяцев операционные расходы в течение 4 месяцев из них не превысят норму. $p=0,75;\ q=1-p=1-0,75=0,25;$

$$P_6(4)=C_6^4\cdot p^4\cdot q^2=\frac{6!}{4!(6-4)!}\cdot 0,75^4\cdot 0,25^2=0,30\,.$$

*Замечание.* Формула Бернулли требует больших вычислений при больших значениях $n$.

## Локальная формула Лапласа

Используется в схеме испытаний Бернулли для более простого асимптотического определения вероятности $k$ успехов из $n$ – испытаний, если $n\to\infty$:

$$P_n(k)\simeq\frac{1}{\sqrt{npq}}\cdot\varphi(x),$$

где $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}},\ x=\dfrac{k-np}{\sqrt{npq}}-$ нормированная и центрированная случайная (стандартная) величина, функция $\varphi(x)$ табулирована.

***Пример 2.15.*** Найти вероятность того, что событие $A$ появится ровно 80 раз в 400 испытаниях, если вероятность события в каждом испытании равна 0,2.

$$p = 0,2;\ n = 400;\ k = 80;\ P_{400}(80) - ? \Rightarrow x = \frac{80 - 400 \cdot 0,2}{\sqrt{400 \cdot 0,2 \cdot 08}} = 0,$$

$$P_{400}(80) = \frac{1}{\sqrt{400 \cdot 0,2 \cdot 0,8}} \cdot 0,3989 = 0,04986\,.$$

## Интегральная формула Лапласа

Имеем вновь схему испытаний Бернулли, но требуется вычислить вероятность того, что событие $A$ появится в $n$ испытаниях от $k_1$ до $k_2$ раз (не менее $k_1$ и не более $k_2$ раз). Эта задача распространена в прикладных вопросах теории вероятностей:

$$P_n(k_1,k_2) = \frac{1}{\sqrt{2\pi}} \cdot \int\limits_{x'}^{x''} e^{-\frac{z^2}{2}} dz,\ x' = \frac{k_1 - np}{\sqrt{npq}},\ x'' = \frac{k_2 - np}{\sqrt{npq}},$$

$$P_n(k_1,k_2) = \frac{1}{\sqrt{2\pi}} \cdot \int\limits_{x'}^{x''} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \cdot \int\limits_{x'}^{0} e^{-\frac{z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \cdot \int\limits_{0}^{x''} e^{-\frac{z^2}{2}} dz =$$

$$= \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{x''} e^{-\frac{z^2}{2}} dz - \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{x'} e^{-\frac{z^2}{2}} dz = \Phi(x'') - \Phi(x'),$$

где $\Phi(x)$ - табулированная функция.

***Пример 2.16.*** В страховой компании 10 000 клиентов. Страховой взнос – 2 000 руб. Вероятность страхового случая $p = 0,005.$ Страховая выплата – 200 000 руб. Определить размер прибыли стразовой компании с вероятностью $p = 0,9.$

Прибыль $R = (20 - 0,2k)$ млн. руб., где $k$ - число страховых случаев. Найдем такое $N$, чтобы $P_{10000}(N < k < 10\,000) \le 0,1$.

$$x' = \frac{N - np}{\sqrt{npq}} = \frac{N - 50}{7,05},\ x'' = \frac{10000 - 9950}{\sqrt{npq}} = 1411,34\,.$$

Из таблицы найдем $\Phi(x) = 0,5$ при $|x| > 5$.
$\Rightarrow 0,5 - \Phi(\frac{N-50}{7,05}) \le 0,1 \Rightarrow \Phi(\frac{N-50}{7,05}) \ge 0,4$ из таблиц $\frac{N-50}{7,05} \ge 1,28 \Rightarrow N \ge 60$.
Окончательно $R = 20 - 0,2 \cdot 60 = 8$ млн. руб.

Случайные одномерные величины

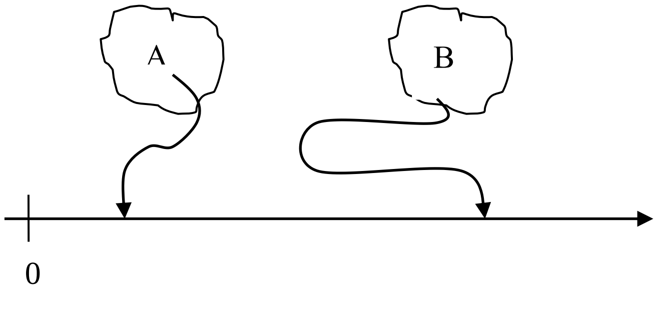



Рис. 2.2

$$\left.\begin{matrix} A \\ B \\ C \end{matrix}\right\} \text{ случайные события,} \Rightarrow \left.\begin{matrix} X \\ Y \\ Z \end{matrix}\right\} \text{ случайные величины (с.в.)} \Rightarrow$$

$$\left.\begin{matrix} x \\ y \\ z \end{matrix}\right\} \text{ значения (в том числе и отрицательные) с.в. на числовой оси.}$$

Случайной называют величину, которая в результате испытания примет одно и только одно возможное значение, априори неизвестное и не зависящее от случайных причин, которые заранее не могут быть учтены.

С.в. может быть непрерывной (аналоговой) или дискретной (прерывистой).

***Пример непрерывной с.в.:*** расстояние, которое пролетает снаряд из орудия, зависит от большого количества факторов (ветра, $t^0$, угла прицела, изменений количества и состояния пороха и т.д.), точности измерений (шаги, метры, микроны и т.д.). Имеет место и «эллипс» рассеяния (влево-вправо).

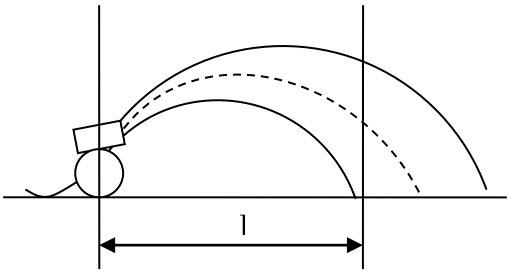



Рис. 2.3

Непрерывная с.в. имеет бесконечное число значений и несчетна. Вероятность отдельного конкретного значения случайной величины равна 0 (но это событие возможно). Можно говорить о вероятности диапазона значений непрерывной с.в.

Дискретная с.в. может быть конечной или бесконечной, но она счетна (ей можно поставить в соответствие натуральный ряд чисел).

Примеры дискретной случайной величины:

1) Число родившихся мальчиков из 100 новорожденных случайно: 0, 1, 2, 3, ... 100. Однако оно всегда дискретно и устойчиво больше, чем число родившихся девочек. С возрастом соотношение выравнивается, затем становится меньше.

2) Вероятность числа «успехов» в схеме Бернулли.

Характеристики с.в.:

1) Графики, таблицы.

2) Аналитические функции (интегральная и дифференциальная функция распределения), функционалы от них (числовые характеристики), квантили.

## Распределения вероятностей дискретной случайной величины

$$X \Rightarrow x_i \Rightarrow P(x_i) = p_i,$$

где $i = \overline{1,n}$; $\sum\limits_{i=1}^{n} p_i = 1$ - условие полноты группы событий (нормировки).

Ряд распределения, если наложено условие, $x_{i+1} \geq x_i$ => вариационный ряд.

| | | | | |
|---|---|---|---|---|
| | | | .. | |
| $p$ | | | .. | |

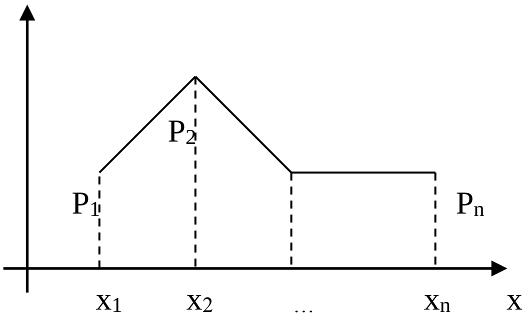



Рис. 2.4

График: многоугольник или полигон распределения.

***Пример 2.17.*** В денежной лотерее выпущены 100 билетов. При этом могут быть 1 выигрыш по 50 руб., 10 выигрышей по 1 руб. Найти закон распределения. Пусть $x$ – величина возможного выигрыша при покупке 1 билета.

$x_1$=50 руб.$\Rightarrow$ $p_1$=1/100;

$x_1$=10 руб.$\Rightarrow$ $p_1$=10/100;

$x_1$=0 руб.$\Rightarrow$ $p_1$=89/100.

| | | | |
|---|---|---|---|
| $x$ | 0 | 1 | 50 |
| $p$ | 0,89 | 0,1 | 0,01 |

Известно более 100 аналитических распределений.

1) Биноминальное распределение (схема испытаний Бернулли)

$$P_n(k) = C_n^k p^k q^{n-k}.$$

***Пример 2.18.*** Монета брошена 2 раза. Определить закон распределения числа выпадения герба. Пусть $x$ – появление «герба». Тогда $p = 1/2$; $g = 1/2$.

$P_2(0) = C_2^0 q^2 = 1 \cdot (1/2)^2 = 0,25$;

$P_2(1) = C_2^1 pq = 2 \cdot (1/2)^2 = 0,5$;

$P_2(2) = C_2^2 p^2 = 1 \cdot (1/2)^2 = 0,25;$

$\sum_{i=0}^{2} P_i = 1.$

2) Распределение Пуассона

Схема испытаний Бернулли$\Rightarrow$формула Бернулли. При дополнительном условии $n \to \infty \Rightarrow$ локальная теорема Лапласа. При дополнительных условиях $p \to 0$, $np = \lambda$. вероятность $k$ «успехов» из $n$ испытаний определится асимптотической формулой Пуассона:

$$P_n(k) = \frac{\lambda^k}{K!} \cdot e^{-\lambda}.$$

Формула Пуассона широко используется в теории массового обслуживания и в теории надежности, где $\lambda$ - имеет смысл интенсивности отказов.

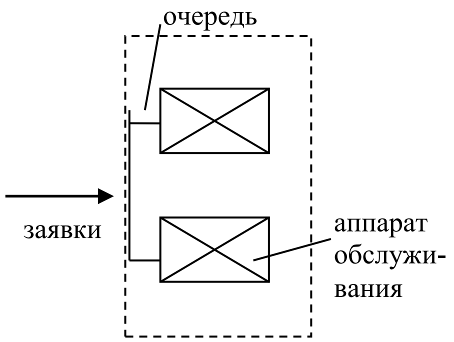



Рис. 2.5

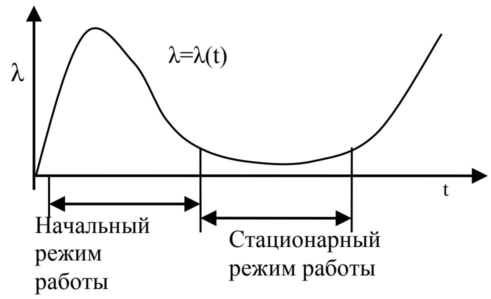



Рис. 2.6

***Пример 2.19.*** Поставщик отправил дистрибьютеру 5000 товаров. Вероятность того, что единица товара выйдет из строя, 0,0002. Найти вероятность того, что у дистрибьютера выйдут из строя 3 единицы товара.

$$n = 5000;\ p = 0,0002. \Rightarrow \left.\begin{matrix} n \to \infty \\ p \to 0 \end{matrix}\right\} \lambda = np = 1\ \ P_{5000}(3) = \frac{1^3}{3!} \cdot e^{-1} = 0,06.$$

## Основные формулы комбинаторики, используемые для определения вероятностей алгебры событий

Комбинаторика изучает количество комбинаций, подчиненных определенным условиям для элементов конечного множества.

1. *Перестановками* называют комбинации, составленные из одних и тех же элементов множества, отличающихся только порядком расположения $P_n = n!$.

***Пример 2.20.*** Сколько трехзначных чисел можно составить из цифр 1, 2, 3, если каждая цифра входит в число один раз. $P_3 = n!\big|_{n=3} = 3! = 6.$

2. *Размещениями* $A_m^n$ называются комбинации, составленные из $n$ – различных элементов по $m$ – элементам, которые отличаются либо составом элементов, либо их порядком.

$$A_n^m = n(n-1)(n-2)\ldots(n-(m-1)) = n(n-1)(n-2)\ldots(n-m+1)\,.$$

***Пример 2.21.*** Сколько сигналов можно составить из 6 букв: *A, B, C, D, E, F* по 2 элемента?

*AB, AC, AD, AE, AF*
*BA, BC, BD, BE, BF*
*CA, CB, CD, CE, CF*
*DA, DB, DC, DE, DF*
*EA, EB, EC, ED, EF*
*FA, FB, FC, FD, FE* $\Rightarrow A_6^2 = 6 \cdot 5 = 30.$

3. *Сочетаниями* называют комбинации, составленные из n- различных элементов по m – элементам, которые отличаются хотя бы одним элементом. $C_n^m = \dfrac{n!}{m!(n-m)!}$.

***Пример 2.22.*** Сколькими способами можно выбрать 2 детали из ящика, содержащего 10 деталей?

$$C_{10}^{\,2} = \frac{10!}{2!\cdot 8!} = 45\,.$$

$A_n^m = P_m \cdot C_n^m$ - связь числа размещений, перестановок и сочетаний.

**Функция (интегральный закон) распределения с.в.**

$F(x)$ – является универсальной характеристикой и для непрерывных и для дискретных одномерных с.в. и описывает вероятность события $P(X < x)$:

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

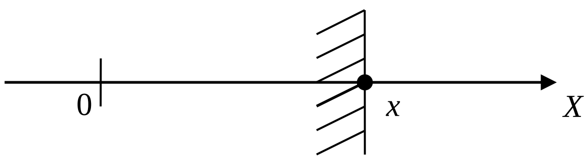



Рис. 2.7

***Пример 2.23.*** Пусть случайная величина X задана таблицей распределения (т.е. X – дискретная случайная величина).

| | | | |
|---|---|---|---|
| $x$ | 1 | 4 | 8 |
| $p$ | 0,3 | 0,1 | 0,6 |

$$P(X<1)=0;\ P(X<4)=0{,}3;\ P(X<8)=0{,}4;\ P(X<9)=1\,;\ P(X<10)=1. \Rightarrow \sum_i P_i = 1\,.$$

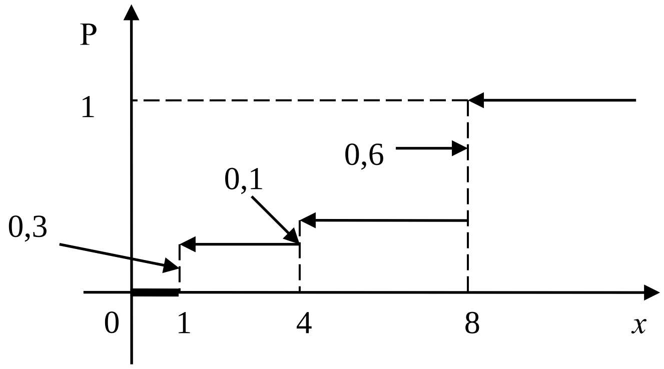
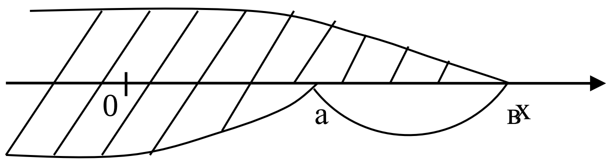



Рис. 2.8

## Свойства функции распределения

1. $0 \le F(x) \le 1$;
2. $F(+\infty) = P(x < +\infty) = 1$;
3. $F(-\infty) = P(x < -\infty) = 0$;
4. $P(X \ge x) = 1 - F(x)$ - вероятность противоположного события;
5. $P(a \le x \le b) = F(b) - F(a)$- вероятность попадания в интервал значений.




6. $x_2 < x_1 \Rightarrow F(x_2) \ge F(x_1)$

## Плотность вероятности (дифференциальный закон) распределения непрерывной с.в.

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

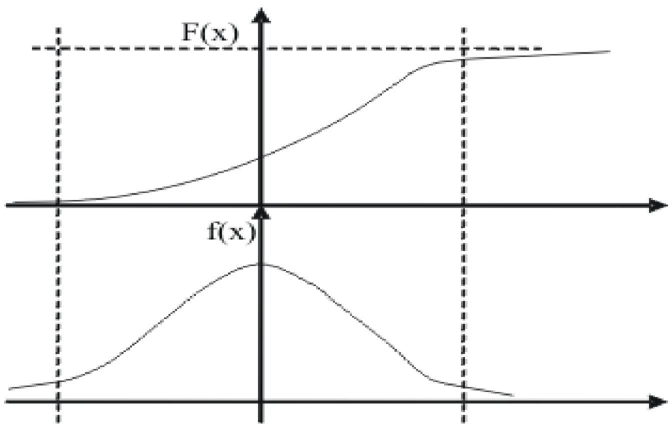



Рис. 2.9

Свойства плотности вероятности

1. $f(x) \geq 0$;

2. $\int_{-\infty}^{+\infty} f(x)dx = 1$ - условие нормировки (единичная площадь под кривой распределения, полнота группы событий);

3. $F(x) = \int_{-\infty}^{x} f(t)dt$ - выражение функции распределения через плотность;

4. $P(a \leq x \leq b) = \int_{a}^{b} f(x)dx = F(b) - F(a)$ - вероятность попадания в интервал значений.

Функции $F(x)$, $f(x)$, ряды распределения $p_i$ исчерпывающим образом описывают одномерную с.в., однако они довольно сложны, а информация, содержащаяся в них, зачастую избыточна, поэтому широко используют числовые характеристики.

## Числовые характеристики распределения с.в.

Математическое (безусловное) ожидание с.в.

$$M[x] = m_X = \begin{cases} \sum_i x_i p_i \\ \int_{-\infty}^{+\infty} x f(x)dx \end{cases},$$

где $p_i$ - вероятность, а $f(x)dx = dp$ - элемент вероятности (вероятность попадания в $dx$),

$M[.]$ - оператор математического ожидания.

$m_X$ - центр группирования, «среднее значение», центр тяжести плоской фигуры, ограниченной кривой $f(x)$.

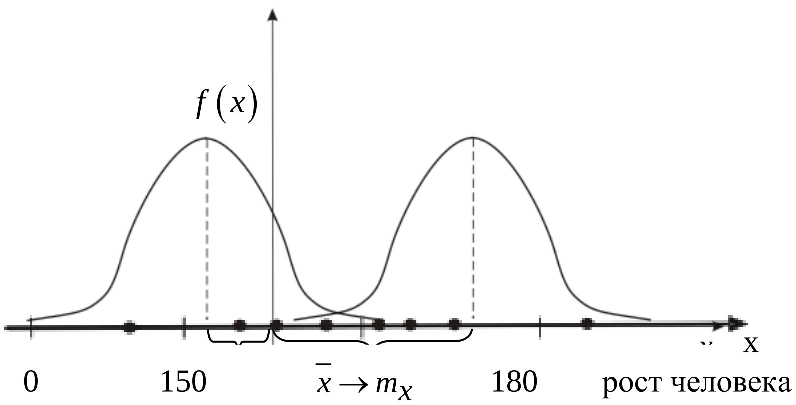



Рис. 2.10

Свойства математического ожидания

1. $M[x] = c \cdot 1 = c;$
2. $M[cx] = cm_x;$
3. $M[x \pm y] = m_x \pm m_y;$
4. $M[ax + b] = am_x + b;$
5. $M[xy] = m_x m_y,$ если $x$ и $y$ – независимые с.в.

$\alpha_k[x] = \begin{cases} \sum\limits_i x_i \cdot P_i \\ \int\limits_{-\infty}^{\infty} x^k f(x)dx \end{cases}$ - начальный момент $k$ – того порядка,

$\alpha_0 = 1,$

$\alpha_1[x] = m_x.$

1. Дисперсия (безусловная) с.в.

$$D_x = D[x] = \begin{cases} \sum\limits_i (x_i - m_x)^2 \cdot P_i \\ \int\limits_{-\infty}^{\infty} (x - m_x)^2 \cdot f(x)dx \end{cases}$$

$x_i - m_x = \overset{0}{x}$ - центрированная с.в.;

$D_x$ - мера разброса (рассеяния) вокруг матожидания (в квадратных единицах);

$D[.]$ - оператор дисперсии.

Свойства дисперсии

1. $D[c] = (c - c)^2 \cdot 1 = 0$.
2. $D[cx] = c^2 D_x$.
3. $D[x \underline{+} y] = D_x + D_y$.
4. $D[ax + b] = a^2 D_x$.

$\sigma_x = \sqrt{D_x}$ - среднее квадратическое отклонение (в линейных единицах).

$k_в = \dfrac{\sigma_x}{m_x}$ - коэффициент вариации (изменчивости) (в относительных единицах).

*Замечание.* Коэффициент вариации применяют при анализе риска инвестирования.

Удобная формула расчета дисперсии:

$$D[X] = M[(x - m_x)^2] =$$
$$= M[x^2 - 2xm_x + m_x^2] = M[x^2] - 2m_x M[x] + m_x^2 = M[x^2] - m_x^2.$$

$$\mu_k[x] = \begin{cases} \sum\limits_i (\overset{0}{x}_i)^k p_i \\ \int\limits_{-\infty}^{\infty} (\overset{0}{x})^k f(x)dx \end{cases}$$ - центральный момент $k$ – того порядка;

$\mu_2[x] = D_x$;

$k_A = \dfrac{\mu_3[x]}{\sigma_x^{\ 3}}$ - коэффициент асимметрии;

$E_x$ - эксцесс (коэффициент островершинности).

$\begin{matrix} k_A = 0 \\ E_x = 0 \end{matrix} \Rightarrow$ нормальное (Гауссово) распределение выступает в качестве меры (эталона):

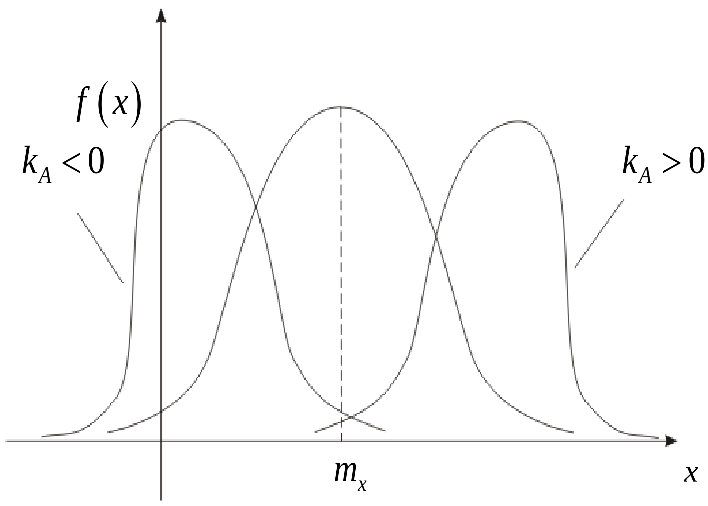



Рис. 2.11

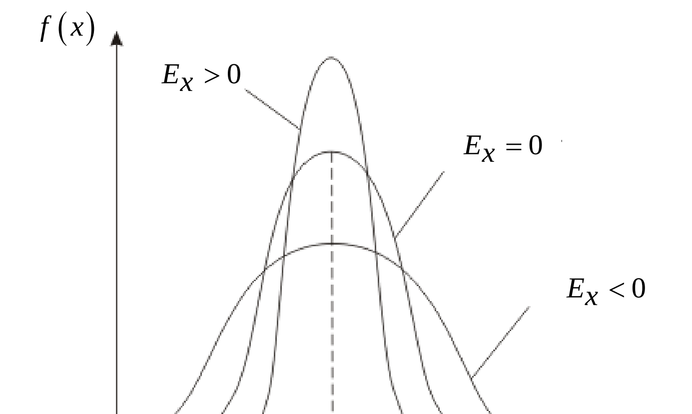



Рис. 2.12

5. Квантили распределения

$x_p$ - квантиль уровня (вероятности) $p$ определяется как корень уравнения $F\left(x_p\right) = p$ или, соответственно, $P(X < x_p) = p$.

Квантили:

$x_{1/2} = \mu e$; (медиана) квартили; децили; процентили.

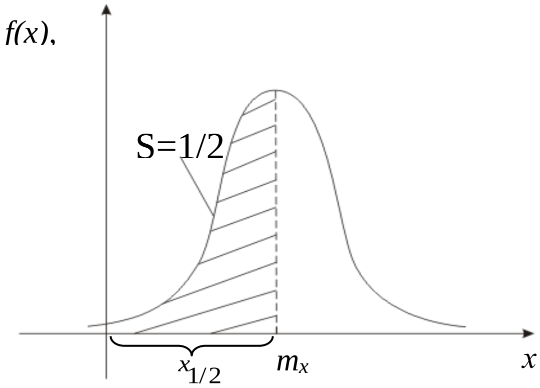



Рис. 2.13

Квартили

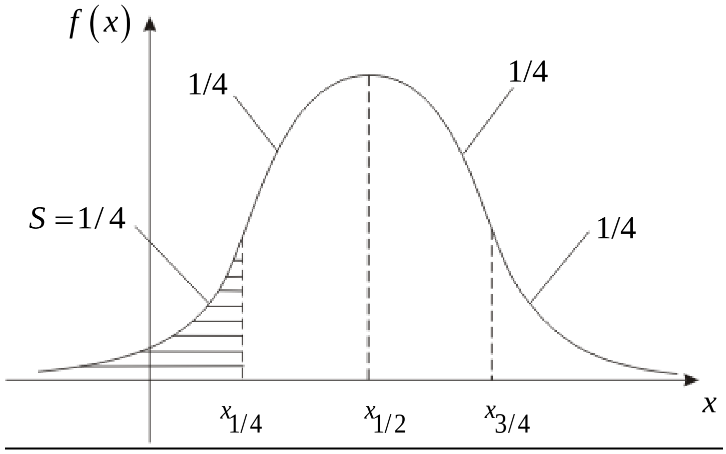



Рис. 2.14

Квартили - $x_{1/4}, x_{2/4}, x_{3/4}$.

Децили - $x_{1/10}, x_{2/10}, \ldots, x_{9/10}$.

Процентили - $x_{1/100}, x_{2/100}, \ldots, x_{99/100}$.

Квантили широко используются в робастной статистике: 178; 187; 1655, 175; 178 $\Rightarrow$ $m_X = 474{,}6$. Наблюдение 1655 – выброс (промах).

175; 178; 178; 187; 1655 $\Rightarrow \mu e = 178$.

Мода распределения - $\mu o = \arg\max f(x)$.

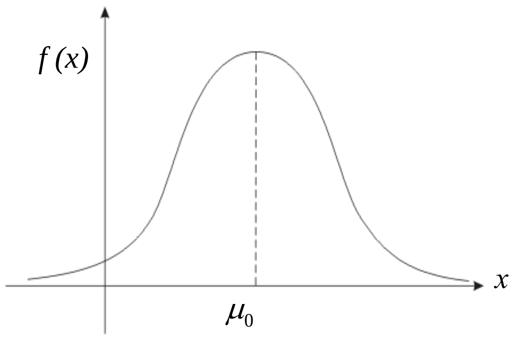



Рис. 2.15. Одномодальное (унимодальное) распределение

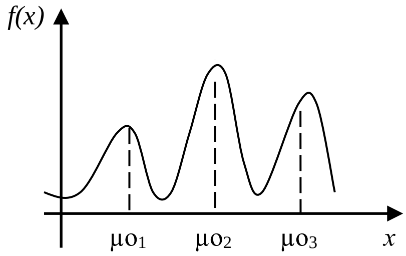



Рис. 2.16. Полимодальное распределение

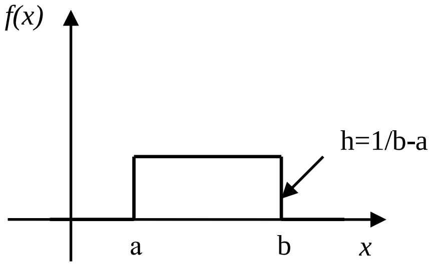



Рис. 2.17. Амодальное распределение

Пример использования моды: размеры одежды, перепись населения.

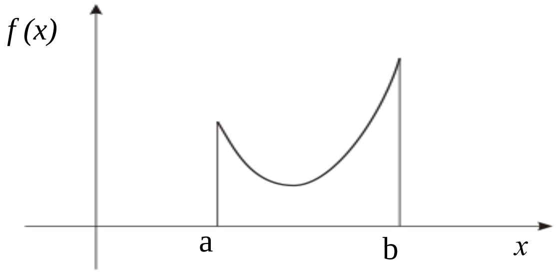



Рис. 2.18. Антимодальное распределение

## Нормальное (Гауссово) распределение

$$f(x)=\frac{1}{\sigma_x\sqrt{2\pi}}e^{\frac{-(x-m_x)^2}{2\sigma_x^2}}, \qquad F(x)=\frac{1}{\sigma_x\sqrt{2\pi}}\cdot\int_{-\infty}^{x}e^{\frac{-(t-m_t)^2}{2\sigma_t^2}}dt.$$

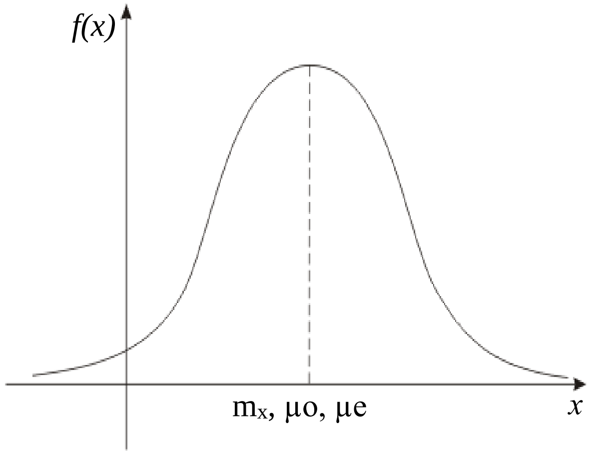



Рис. 2.19

$f(x)=f(x,m_x,\sigma_x)$ - двухпараметрическое распределение,

$x \sim N(m_x,\sigma_x)$ - обозначение с.в., имеющей нормальное распределение

$u \sim N(0,1)$ - $\left.\begin{matrix} m_x=0 \\ \sigma_x=1 \end{matrix}\right\}$ - обозначение с.в., имеющей нормированное (стандартное) нормальное распределение

Операции нормирования и центрирования:

$$\frac{x-m_x}{\sigma_x}\Rightarrow M\left[\frac{x-m_x}{\sigma_x}\right]=\frac{1}{\sigma_x}\{M[x]-m_x\}=0;$$

$$D\left[\frac{x-m_x}{\sigma_x}\right]=\frac{1}{\sigma_x^2}\cdot\sigma_x^2=1.$$

Правило трех сигм

$P\left(x\in[m_x-\sigma_x, m_x+\sigma_x]\right)=0,68$.

$P\left(x\in[m_x-2\sigma_x, m_x+2\sigma_x]\right)=0,95$.

$P\left(x\in[m_x-3\sigma_x, m_x+3\sigma_x]\right)=0,99$.

$m_x\pm 3\sigma_x$ - зона практического рассеивания

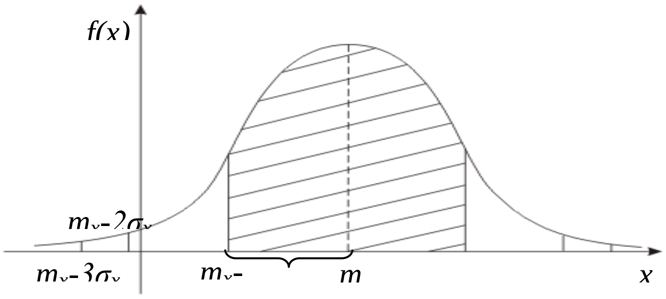


Рис. 2.20

## Функция Лапласа

$u = \dfrac{x - m_X}{\sigma_X}$, $f(u) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$, $F(u) = \dfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{u} e^{-\frac{t^2}{2}} dt$ - стандартное нормальное распределение с.в. от $-\infty$ до $u$. Проще рассматривать случай, когда $u \ge 0$.

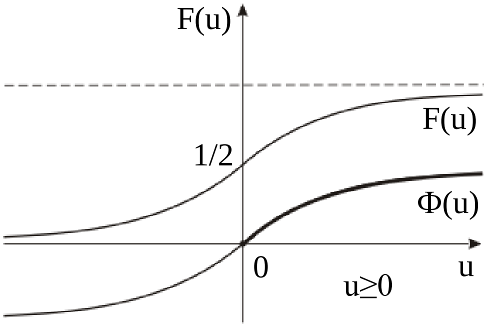



Рис. 2.21

$$\Phi(u) = F(u) - 1/2 = \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{u} e^{-\frac{t^2}{2}} dt .$$

- табулированная функция Лапласа, широко используемая для определения вероятности попадания в диапазон значений:

$$P(a \le x \le b) = \Phi(\frac{b - m_X}{\sigma_X}) - \Phi(\frac{a - m_X}{\sigma_X}) .$$

## Роль нормального распределения. Теоретики и практики («мифы»)

*Центральная теорема:* сумма конечного числа нормальных с.в. есть нормальная с. в.

*Центральная предельная теорема:* сумма бесконечного числа с.в. с любыми законами распределения, но с примерно одинаковыми дисперсиями имеет нормальное распределение.

Многие экономические показатели имеют нормальный или близкий к нормальному закон распределения: доход населения, прибыль фирмы в отрасли и др.

***Пример 2.24.*** В результате длительных наблюдений определено, что дивиденды $X'$ и $Y'$ по акциям фирм $A$ и $B$ являются нормальными с. в.:

$X'\sim N(5;5)$; $Y'\sim N(15;15)$. Стоимость каждой акции равна 100$. Инвестор хочет приобрести акции на 1000 $.

а) Какие законы распределения имеют доходы X, Y от вложения всей суммы в акции фирмы A или B? б) Каков закон распределения имеет доход Z от покупки акций в пропорции 2/3? в) Построить графики функций случайных величин X, Y, Z. г) Какова вероятность того, что полученный доход Z от вложения будет лежать в пределах от 110$ до 150$?

а) $X=10X'\Rightarrow$ X ~ $N\,(50;\sqrt{10^2\cdot 25}=50)$ или $Y=10Y'\Rightarrow$ Y ~ $N\,(150, \sqrt{10^2\cdot 125}=150)$;

б) $Z=4\cdot X'+6\cdot Y'$ $\Rightarrow$ Z ~ $N$ ($m_z=4\cdot 5+6\cdot 15=110$, $\sigma_z=\sqrt{16\cdot 25+36\cdot 225}=92,2$);

в) При построении графиков целесообразно пользоваться правилом трех сигм и обязательно соблюдать условие нормировки (площадь под кривой распределения одна и та же - равна единице).

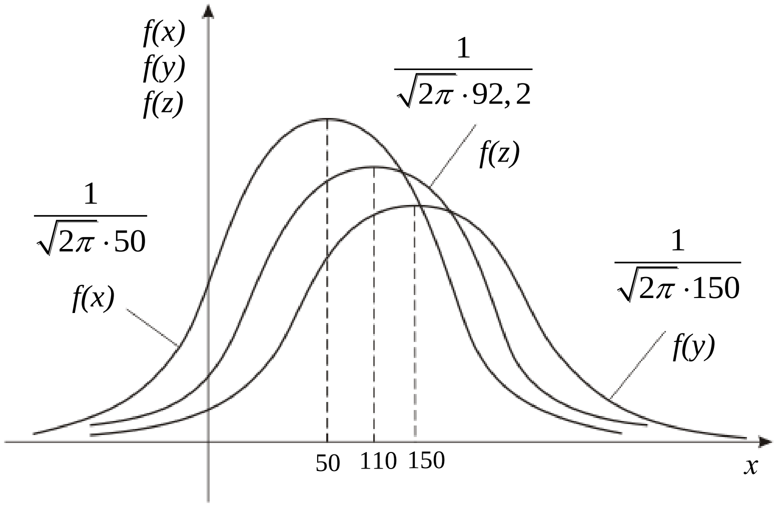



Рис. 2.22

г) P(110≤Z≤150)=Ф(0,43)-Ф(0)=0,16 (используются таблицы функции Лапласа).

**Равномерное (равновероятное, прямоугольное) распределение**

$$f(x)=\begin{cases}\dfrac{1}{b-a}, a\le x\le b,\\ 0, \qquad b<x<a.\end{cases}$$

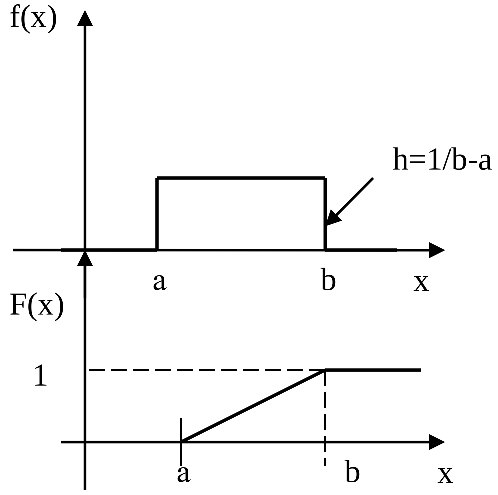



Рис. 2.23

$$F(x) = \int_{-\infty}^{\infty} \frac{1}{b-a} dt = \int_{-\infty}^{a} 0 dt + \int_{a}^{x} \frac{1}{b-a} dt = \frac{x-a}{b-a},$$

$$m_x = \frac{a+b}{2}, \qquad D_x = \frac{(b-a)^2}{12},$$

$$m_x = \mu e, \qquad k_A = 0.$$

## Показательное (экспоненциальное) распределение

$$f(x) = \begin{cases} 0, x < 0, \\ \lambda e^{-\lambda x}, x \ge 0. \end{cases}$$

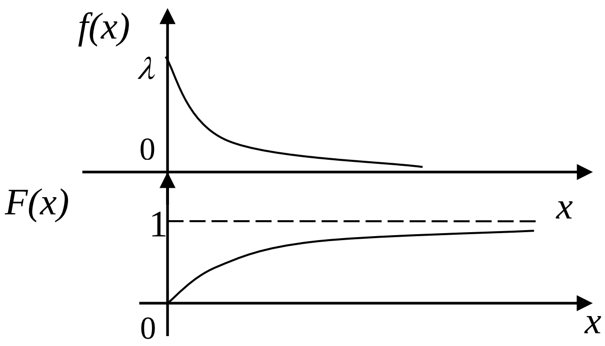



Рис. 2.24

$$F(x) = \int_{-\infty}^{x} \lambda e^{-\lambda x} dt = \int_{-\infty}^{0} 0 dt + \lambda \int_{0}^{x} e^{-\lambda t} dt = 1 - e^{-\lambda x},$$

$$m_x = \int_{0}^{\infty} \lambda e^{-\lambda x} \cdot x dx = \frac{1}{\lambda},$$

$$D_x = M[(x - \frac{1}{\lambda})^2] = \int_{0}^{\infty} (x - \frac{1}{\lambda})^2 f(x) dx = \int_{0}^{\infty} \lambda e^{-\lambda x} (x - \frac{1}{\lambda})^2 dx = \frac{1}{\lambda^2}.$$

$\sigma_x = \frac{1}{\lambda} = m_x$ - характеристическое свойство показательного распределения.

Приложение - функция надежности. Пусть объект анализа начинает функционировать в момент времени $t=0$ и по истечении $t$ выходит из строя. Обозначим через $T$ время безотказной работы. Тогда вероятность отказа за время t равна:

$$P(T>t)=1-F(t),\ F(t)=P(T<t).$$

Для оценки вероятности безотказной работы («накапливающиеся отказы») часто используется экспоненциальное распределение. При этом параметр распределения λ=λ(t) является функцией времени. Технические системы в демографии – смертность.

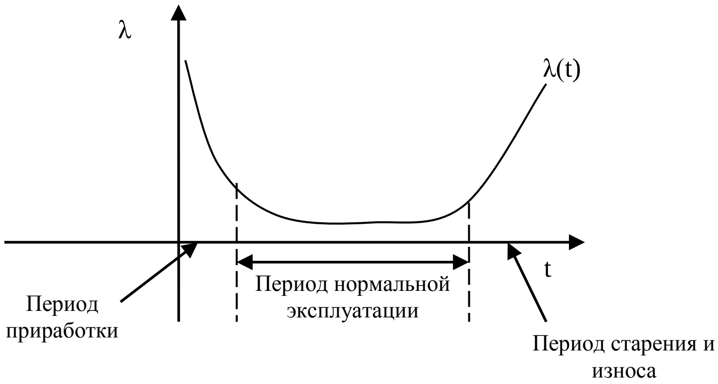



Рис. 2.25. «Внезапные отказы» – гамма-распределение

## Распределение $\chi^2$

Сумма квадратов $n$ нормальных с.в. $u_i \sim N(0,1)$ $\boxed{\chi^2 = \sum_{i=1}^{n} u_i^2}$ является с.в. и имеет табулированное распределение $\chi^2$ с $\nu = n - m$ числом степеней свободы. Здесь $m$ – число наложенных связей: обязательно - условие нормировки и связи, связанные с расчетом тех или иных центральных или начальных моментов. Аналитическое выражение не приводится (сложное), распределение табулировано. $M[\chi^2]=\nu;\ D[\chi^2]=2\nu;\quad \nu\to\infty \Rightarrow \chi^2 \sim N(\quad,\quad)$.

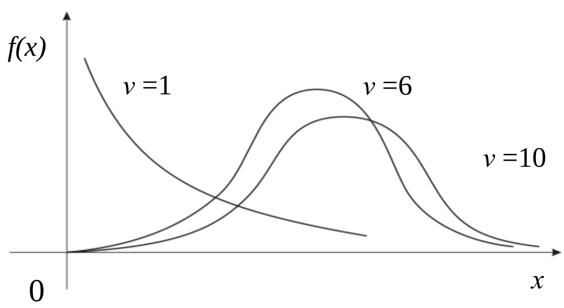



Рис. 2.26

## t - распределение (Стьюдента)

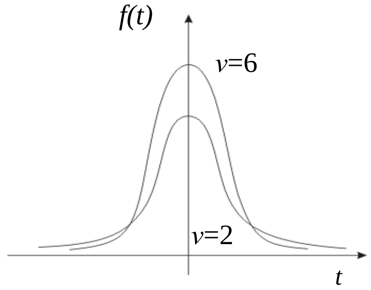



Рис. 2.27

Если $u \sim N(0,1)$, а $V \sim \chi^2$ с $\nu$ степенями свободы, то с.в. $t = u / \sqrt{\dfrac{\chi^2}{\nu}}$ имеет распределение Стьюдента с $\nu$ степенями свободы.

$M[T] = 0$; $D[T] = \dfrac{\nu}{\nu - 2}$. $\nu \to \infty \Rightarrow t \sim N(\ \ ,\ \ )$.

Основное распределение малых выборок (до 15-17 наблюдений). Распределение табулировано.

## F - распределение (Фишера)

Если $W$ и $V$ – независимые случайные величины, распределенные по закону $\chi^2$ со степенями свободы $m$ и $n$ соответственно, то с.в.

$f = \dfrac{V/m}{W/n}$ имеет распределение Фишера (табулировано).

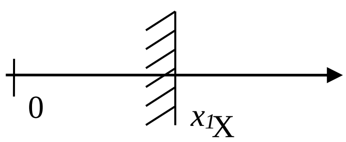
В статистике широко используют прием введения некоторой случайной величины, распределение которой не зависит от числовых характеристик исходного анализируемого распределения. Такие величины имеют распределение Стьюдента, Фишера и $\chi^2$.

## Двумерные с.в. («проклятие размерности»)

Уже было для событий и одномерных с.в.:

$$A \Rightarrow P(A) \Rightarrow x \Rightarrow \begin{cases} p(x_i) = p_i \\ f(x) \end{cases}; \qquad B \Rightarrow P(B) \Rightarrow y \Rightarrow \begin{cases} p(Y = y_j) = p_j \\ f(y) \end{cases}$$

$$P(A \cdot B) = P(A) \cdot P(B \mid A) \Rightarrow P(B \mid A) = \frac{P(A \cdot B)}{P(A)};$$




$F(x) = P(X < x)$; $\qquad f(x) = \dfrac{dF(x)}{dx}$; $F(x) = \int\limits_{-\infty}^{\infty} f(t)dt.$

$$(X,Y) \Rightarrow \begin{cases} P(X < x; Y < y) = F(x,y) \\ P(X = x_i; Y = y_j) = p_{ij} \end{cases}$$ - двумерные функция и ряд распределения.

| | $x_1$ | $x_2$ | $\dots$ | $x_n$ |
|---|---|---|---|---|
| $Y_1$ | $p_{11}$ | $p_{21}$ | $\dots$ | $p_{n1}$ |
| $Y_2$ | $p_{12}$ | $p_{22}$ | $\dots$ | $p_{n2}$ |
| $\dots$ | $\dots$ | $\dots$ | $\dots$ | $\dots$ |
| $Y_n$ | $p_{1n}$ | $p_{2n}$ | $\dots$ | $p_{nn}$ |

$\sum\limits_i \sum\limits_j P_{ij} = 1$ - условие нормировки.

$$\sum_{i=1}^{2} \sum_{j=1}^{3} P_{ij} = \sum_{i=1}^{2} (P_{i1} + P_{i2} + P_{i3}) = P_{11} + P_{21} + P_{12} + P_{22} + P_{13} + P_{23}$$

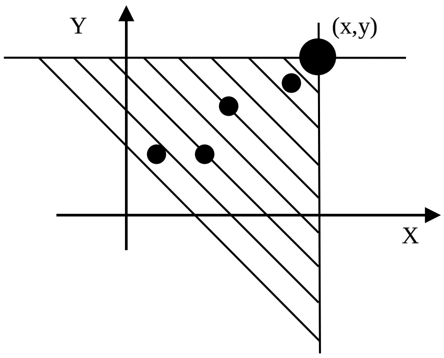



Рис. 2.28

$f(x,y) = \dfrac{\partial^2 F(x,y)}{\partial x \partial y}$; $F(x,y) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f(u,v)\,du\,dv$; $f(x,y)$ – поверхность, $f(x,y) \ge 0$.

$F(x,y)$ – неубывающая функция двух переменных, поверхность.

$\int\limits_{-\infty}^{\infty} f(x,y)dx = f_2(y)$; $\int\limits_{-\infty}^{\infty} f(x,y)dy = f_1(y)$ - получение одномерных плотностей.

$\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f(x,y)dxdy = 1$ - условие нормировки (единичный объем).

$f(x,y) = f_1(x) \cdot f(y \mid x)$;   $f(x,y) = f_1(x) \cdot f_2(y)$ - для независимых с.в.

$$f(y \mid x) = \frac{f(x,y)}{f_1(x)} = \frac{f(x,y)}{\int\limits_{-\infty}^{\infty} f(x,y)dy}$$ -условные вероятности двумерной с.в.

$$f(x \mid y) = \frac{f(x,y)}{f_2(y)} = \frac{f(x,y)}{\int\limits_{-\infty}^{\infty} f(x,y)dx}$$ (сечения поверхности $f(x,y)$).

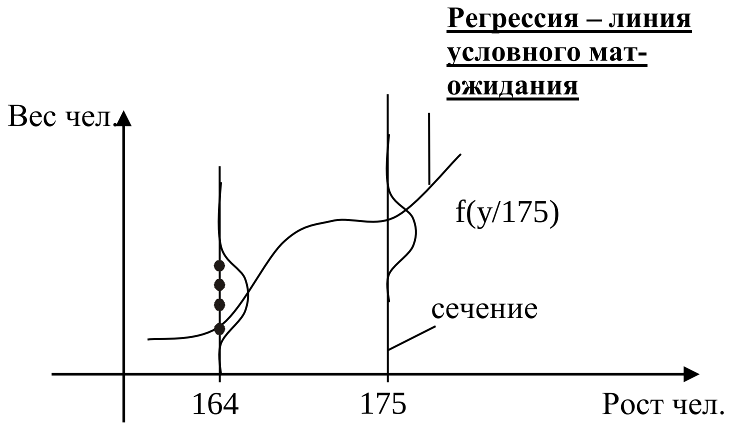



Рис. 2.29

Двумерный закон распределения является исчерпывающей характеристикой, но сложен, поэтому используют числовые характеристики:

$$\alpha_{k,l}[X,Y]=\begin{cases}\sum\limits_i\sum\limits_j x_i^k y_j^l P_{ij}\\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} x^k y^l f(x,y)dxdy\end{cases}$$ - начальный смешанный момент порядка $k+l$.

$$\mu_{k,l}[X,Y]=\begin{cases}\sum\limits_i\sum\limits_j (\overset{0}{x_i})^k (\overset{0}{y_j})^l P_{ij}\\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (\overset{0}{x})^k (\overset{0}{y})^l f(x,y)dxdy\end{cases}$$ - центральный смешанный момент порядка $k+l$.

## Корреляционный (взаимокорреляционный, кросскорреляционный) момент (ковариация)

$$\mu_{11}[X\cdot Y]=M[X\cdot Y]-m_x\cdot m_y=K_{xy}=\sigma_{xy}.$$

$r_{xy}=\dfrac{K_{xy}}{\sigma_x\sigma_y}$; $-1\le r_{xy}\le 1$ - коэффициент корреляции характеризует степень тесноты линейной связи.

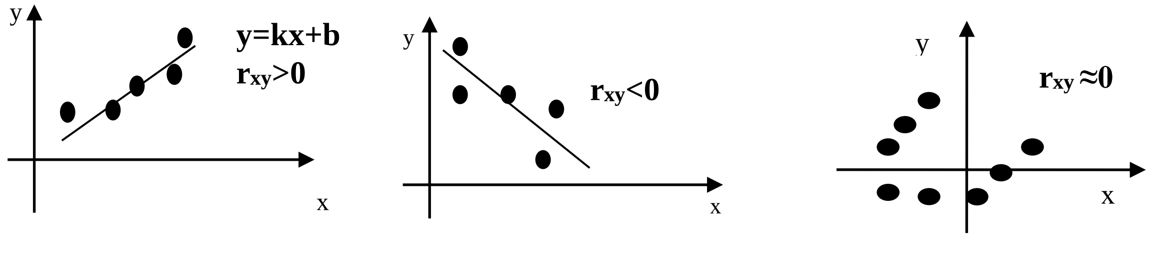



Рис. 2.30

Свойства:

1) $r_{xy} = r_{yx}$;

2) $r_{xx} = 1$;

3) если $x$ и $y$ независимы, то $r_{xy} = 0$; но если $r_{xy} = 0$, то это не означает, что $x$ и $y$ независимы;

4) $r_{xy} \le \sigma_x \cdot \sigma_y$;

5) если $\left| r_{xy} \right| = 1 \Rightarrow Y = kx + b$ - имеем детерминированную (функциональную) зависимость.

*Замечание:* Из некоррелированности нормально распределенной двумерной случайной величины следует независимость переменных:.

Если $r_{xy} = 0,$ то

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1 - r^2{}_{xy}}} \cdot \exp\left(-\frac{1}{2\sqrt{1 - r^2{}_{xy}}}\right) \times$$

$$\times\left[\frac{(x - m_x)^2}{\sigma_x{}^2} + \frac{(y - m_y)^2}{\sigma_y{}^2} - 2r_{xy} \cdot \frac{(x - m_x)(y - m_x)}{\sigma_x\sigma_y}\right] = f_1(x) \cdot f_2(y).$$

При $r_{xy} \approx 0$ возможен другой (нелинейный) вид регрессии (стохастической связи), например, $Y = ax^2 + bx + c, Y = Ae^{-\alpha x}, \ \ldots.$

## Условные математические ожидания (регрессии)

$$M[Y \mid x] = \begin{cases} \sum\limits_j y_j P(y_j \mid x) \\ \int\limits_{-\infty}^{\infty} yf(y \mid x)dy \end{cases}; \ P(y_j \mid x) = \frac{P(y_j, x)}{P(x)} = \frac{P(y_j, x)}{\sum\limits_j P(x, y_j)};$$

$$M[X \mid y] = \begin{cases} \sum\limits_i x_i p(x_i \mid y) \\ \int\limits_{-\infty}^{\infty} xf(x \mid y)dx \end{cases}.$$

***Пример 2.25.*** Дискретная двумерная с. в. задана таблицей распределения: Найти условное математическое ожидание с.в. $Y$ при значении с.в. $X = 1$:т.е. $M[Y \mid x_1 = 1]$.

| | $x_1 = 1$ | $x_2 = 3$ | $x_3 = 4$ | $x_4 = 8$ |
|---|---|---|---|---|
| $y_1 = 3$ | 0,15 | 0,06 | 0,25 | 0,04 |
| $y_2 = 6$ | 0,3 | 0,1 | 0,03 | 0,07 |

$$P[x_1] = \sum_{j=1}^{2} P(x_1, y_j) = 0.45; \ P(y_1 \mid x_1) = \frac{P(x_1, y_1)}{P(x_1)} = \frac{1}{3}; \ P(y_2 \mid x_1) = \frac{P(x_1, y_2)}{P(x_1)} = \frac{2}{3};$$

$$M[Y \mid x_1 = 1] = \sum_{j=1}^{2} y_j P(Y_j \mid x_1) = 5.$$

Пример регрессии X на Y. Условное математическое ожидание веса человека в функции от роста.

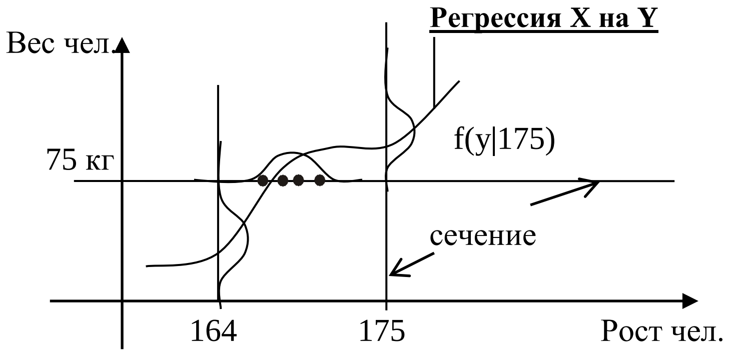



Рис. 2.31

Пример регрессии Y на X. Расчет среднего роста человека при массе (весе) 75 кг. Задавая разные веса и рассчитывая средние статистические значения роста человека при данных весах, получить стохастическую зависимость роста от веса.

Генеральная совокупность – одномоментно измеряют рост всего населения Земли (кто и как измеряет?). «Практические измерения – выборки – статистика».

## Глава 3. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Задачи математической статистики:

1) Указать способы сбора, группировки статистических методов, полученных в результате наблюдений или специально поставленных экспериментов.

2) Разработка методов анализа статистических характеристик в зависимости от цели исследования: оценка вероятности распределения, функции распределения, параметров функции распределения, зависимости случайной величины от одной или нескольких случайных величин.

Большинство решений по методам обработки статистических данных принимают до начала исследований, реже – в ходе исследования - «последовательный анализ» (примеры – приемочный контроль качества, статистическая физика).

### Основные понятия статистики

Генеральная совокупность - объемом $N$. Выборка – объемом $n$ (количество наблюдений для «разумного» результата). Репрезентативность (представительность) выборки. Построение эмпирической (статистической) функция распределения.

Вероятность $p_i = \dfrac{m}{n}$ - неслучайная величина $\Rightarrow$

$\Rightarrow F(x_i) = P(X < x_i)$ - неслучайная функция.

В статистике частота $p_i^* = \dfrac{m}{n}$ - с.в.

Сходимость частоты по вероятности к вероятности:

$$p_i^* \xrightarrow[n\to\infty]{p} p_i \;\Rightarrow \lim_{n\to\infty} P(\left|p_i^* - p_i\right| < \varepsilon) = 1, \quad \forall \varepsilon > 0.$$

$F^*(x) = p^*(X < x) = \dfrac{n_x}{n}$ - случайная функция,

где $n$ - объем выборки,

$n_x$ - число значений наблюдаемой выборки $< x$.

$$F^*(x) \xrightarrow[n\to\infty]{p} F(x) \quad \Rightarrow \lim_{n\to\infty} p(\left|F^*(x) - F(x)\right| < \varepsilon) = 1, \forall \varepsilon > 0.$$

## Свойства эмпирической функции распределения

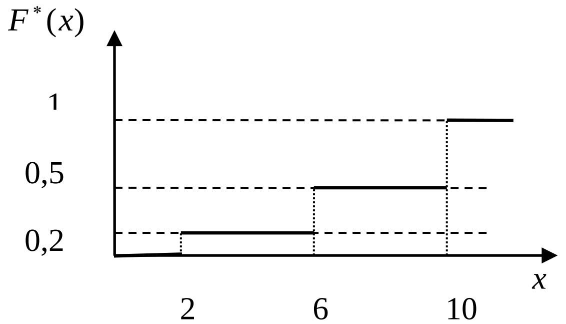
1) $0 \le F^*(x) \le 1$;
2) $F^*(x)$- неубывающая функция;
3) если $x_1$- наименьшая варианта, $\Rightarrow F^*(x) = 0, \quad x \le x_1$;
4)если $x_n$ - наибольшая варианта, $\Rightarrow F^*(x) = 1, \quad x > x_n$.

***Пример 3.1.*** Построим эмпирическую функцию распределения для данного вариационного ряда:

| $x_i$ | 2 | 6 | 10 |
|---|---|---|---|
| $n_i$ | 12 | 18 | 30 |
| $p_i^*$ | 0,2 | 0,3 | 0,5 |

$n = \sum\limits_{i=1}^{3} n_i = 60; \ \sum\limits_{i=1}^{3} p_i^* = 1,$

$$F^*(x) = \begin{cases} 0; & x \le 2 \\ 0,2; & 2 < x \le 6 \\ 0,5; & 6 < x \le 10 \\ 1; & x > 10 \end{cases};$$




***Пример3.2.*** Построение полигона для следующего вариационного ряда:

| $x_i$ | 1,5 | 3,5 | 5,5 | 7,5 |
|---|---|---|---|---|
| $p_i^*$ | 0,1 | 0,2 | 0,4 | 0,3 |
| | | | | |

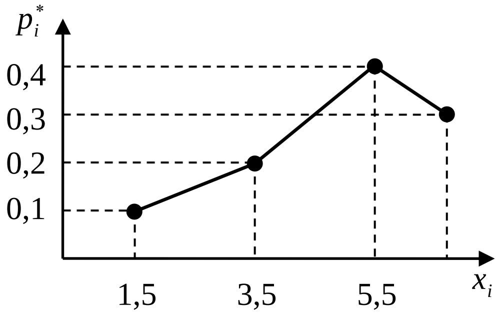



Рис. 3.1. Полигон

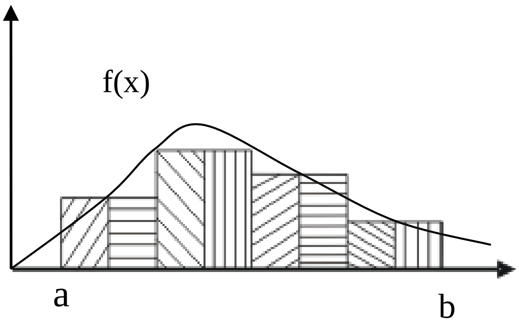



Рис. 3.2. Гистограмма

***Пример 3.3.*** Построим гистограмму для выборки. Разбиваем интервал $(b-a)$, в котором находятся все наблюдаемые значения, на несколько интервалов с шагом $h=\frac{b-a}{k}$. В общем случае $h$ может быть различным. Подсчитываем $n_i$ - число значений выборки, попавших в $i$-й интервал (в случае попадания на границу интервала относят по $1/2$ влево и вправо). Подсчитываем частоты с.в. в $i$-м интервале: $p_i^*=\frac{n_i}{n}$. Строим прямоугольники, у которых высоты равны $l_i=\frac{p_i^*}{h}$; а площади равны $S_i=h\cdot l_i=p_i^*$.

Проверяем гипотезу о виде плотности распределения вероятностей (критерий Пирсона $\chi^2$).

Для проверки гипотезы о виде интегральной функции распределения – используют критерии Романовского, Ястремского, Колмогорова.

Первым и необходимым этапом построения эмпирической функции распределения, гистограммы и расчета выборочных числовых характеристик распределения является обеспечение однородности выборки (устранение аномальных наблюдений).

Под аномальными наблюдениями понимают те отдельные наблюдения, которые не отвечают потенциальным возможностям исследуемой экономической (социально-экономической) системы и которые, оставаясь в выборке, оказывают существенное влияние на статистические характеристики выборки.

Возможные причины: ошибки при агрегировании и дезагрегировании экономических показателей, при передаче, при записи информации и т.д.

За показатель ошибочности отдельного наблюдения принимают значения его отклонения от остальных наблюдений: методы Ирвина, проверки разностей средних уровней, Фостера – Стьюдента, простой скользящей средней и экспоненциального сглаживания.

Методы обеспечения однородности обычно используют $t$ распределение.

Наибольшее сомнение обычно вызывают несколько крайних значений вариационного ряда (левая и правая варианты).

## Статистические оценки параметров распределения

Оценка математического ожидания (выборочной средней):

Неслучайное математическое ожидание: $m_x = \sum_{i=1}^{N} x_i p_i$ - генеральная средняя.

С.в. - оценка матожидания $m_x^* = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$ - выборочная средняя.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{k} x_i m_i = \sum_{i=1}^{k} \frac{1}{n} x_i m_i = \sum_{i=1}^{k} x_i P_i^*,$$

где $P_i^* = \frac{m}{n}$ - частота отдельных значений.

Так как $P_i^* \xrightarrow[n\to\infty]{p} P_i \Rightarrow \overline{X} \xrightarrow[n\to\infty]{p} m_x$.

## Оценка выборочной дисперсии

$$D_x^* = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2 = S^2 = \frac{1}{n}\sum_{i=1}^{k}\left(x_i - \overline{X}\right)^2 m_i = \sum_{i=1}^{k}\frac{1}{n}\left(x_i - \overline{X}\right)^2 m_i = \sum_{i=1}^{k}\left(x_i - \overline{X}\right)^2 P_i^*.$$

$\sigma_x^* = S$ - стандарт (выборочное стандартное – среднеквадратическое отклонение).

Удобная формула оценки дисперсии через оценку второго начального момента

$$S^2 = (\alpha_2)^* - (\overline{X})^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i\right)^2 - (\overline{X})^2.$$

Аналогично рассчитываются и другие выборочные оценки распределения.

## Точечные оценки точности оценок (статистик) генеральных числовых характеристик $\theta_x$

$$\{x_1, x_2, \dots, x_N\} \Rightarrow \theta_x.$$

$\{x_{11}, x_{21}, \dots, x_{n1}\} \Rightarrow \theta_{x1}^*;$

$\{x_{12}, x_{22}, \dots, x_{n2}\} \Rightarrow \theta_{x2}^*;$ - оценки статистических характеристик

…………………………

$\{x_{1k}, x_{2k}, \dots, x_{nk}\} \Rightarrow \theta_{xk}^*.$

$\theta_{xi}^*$ - с.в., характеризуемая законами распределения и числовыми характеристиками распределения (обычно математическим ожиданием и дисперсией).

Можно говорить о распределении оценки матожидания, о матожидании оценки матожидания, о дисперсии оценки матожидания и т.д.

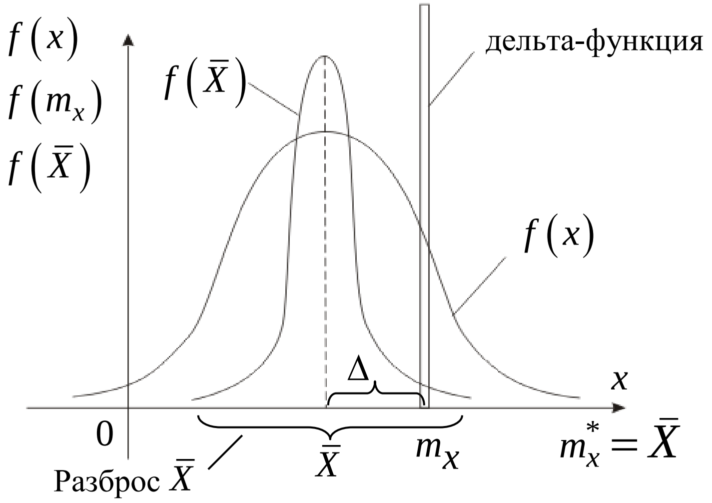



Рис. 3.3

1. Оценка называется состоятельной, если она сходится по вероятности к оцениваемой характеристике $\lim\limits_{n\to\infty} P(\left|\theta_i^* - \theta\right| < \varepsilon) = 1, \quad \forall \varepsilon > 0$.

2. Оценка называется несмещенной, если $M[\theta_i^*] = \theta$ .

$M[\theta_i^*] - \theta = \Delta$ - смещение, систематическая погрешность (от смещенности).

Асимптотически несмещенная оценка $\lim\limits_{n\to\infty} M[\left|\theta_i^* - \theta\right|] = 0.$

3. Оценка называется эффективной, если при используемом методе ее расчета выполняется условие $D[\theta_i^*] = \min$ .

***Пример 3.4***. Оценка $\overline{X}$ является несмещенной, а ее дисперсия уменьшается при усреднении в $n$ раз:

$$M\left[\overline{X}\right] = M\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n} n m_X = m_X.$$

$$D[\overline{X}] = D[\frac{1}{n}\sum_{i=1}^{n} x_i] = \frac{1}{n^2} n D_X = \frac{D_X}{n}.$$

Если $x \sim N(m_X, \sigma_X) \Rightarrow \overline{X}$ - эффективная оценка.

В прикладной статистике и в эконометрике наибольшее внимание уделяют обеспечению эффективности и несмещенности оценок.

***Пример 3.5***. Оценка дисперсии $D_X^* = S^2 = M\left[\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2\right]$ является смещенной:

Доказано, что $S^2 = M\left[\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2\right] = \frac{n-1}{n} D_X$ - т.е. данный алгоритм дает смещенную оценку дисперсии: $\Delta = \frac{D_X}{n}$. Исправленная (несмещенная) оценка дисперсии

$$\frac{nD_X^*}{n-1} = S_{испр.}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2 .$$

На практике исправленной оценкой дисперсии пользуются при $n < 30$.

### Интервальная оценка точности (надежность) генеральных математического ожидания и дисперсии

Доверительный интервал – интервал значений, в котором с заданной доверительной вероятностью (обычно назначают $\beta = 0{,}9;\quad 0{,}95;\quad 0{,}99$) находится истинное значение оцениваемой статистической характеристики $\theta_X$: $P[|\theta - \theta^*| < \delta] = \beta \Rightarrow P(\theta^* - \delta < \theta < \theta^* + \delta) = \beta$ .

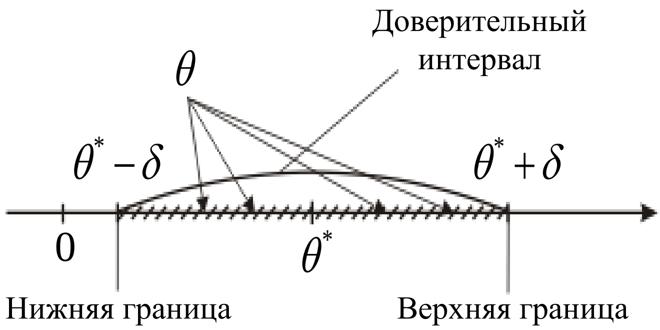



Рис. 3.4

Радиус доверительного интервала равен:

$$\delta = \frac{t \cdot \sigma_\theta}{\sqrt{n}} \approx \frac{t \cdot \sigma_{\theta^*}}{\sqrt{n}},$$

где $t$ - аргумент, соответствующий значению функции Лапласа, равной $\beta/2$:

$\Phi(t) = \beta/2$;

$\sigma_{\theta^*}$ - среднеквадратическое отклонение $\theta_X$ (его оценка).

Доверительный интервал случаен (зависит от конкретных выборок): случайно его положение на числовой оси и случайна его длина.

При $n \to \infty \Rightarrow \delta \to 0$, а при $\beta \to 1 \Rightarrow \delta \to \infty$.

Доверительный интервал для оценки математического ожидания

$$P(\overline{X} - \delta < m_X < \overline{X} + \delta) = \Phi(\frac{\overline{X} + \delta - m_{\overline{X}}}{\sigma_X}) - \Phi(\frac{\overline{X} - \delta - m_{\overline{X}}}{\sigma_X}) =$$

$$= \Phi\left(\frac{\delta}{\sigma_{\overline{X}}}\right) - \Phi\left(-\frac{\delta}{\sigma_{\overline{X}}}\right) = 2\Phi\left(\frac{\delta}{\sigma_X}\right) = \beta \Rightarrow t = \frac{\delta}{\sigma_{\overline{X}}} = \frac{\delta \cdot \sqrt{n}}{\sigma_X} \Rightarrow \quad \delta = \frac{t \cdot \sigma_X}{\sqrt{n}}.$$

Здесь $t$ рассматривается как аргумент табулированной функции распределения Лапласа (нормальной!), при котором она равна значению $\frac{\beta}{2}$: $\Phi(t)=\frac{\beta}{2}$.

Значение генерального среднеквадратического отклонения $\sigma_X$ редко известно, поэтому обычно в формуле используют оценку среднеквадратического отклонения, т.е. $\sigma_X^* = S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2}$.

***Пример 3.6.*** $X \sim N\{\bar{x};3\}$ Найти доверительный интервал для оценки неизвестного $m_x$, при выборочном среднем $\overline{X}$, если объем выборки $n$=36, а $\beta = 0,95$.

$$2\Phi(t) = 0,95 \Rightarrow \Phi(t) = 0,475 \Rightarrow t = 1,96.$$

$$\delta = \frac{t\sigma_X}{\sqrt{n}} = \frac{1,96 \cdot 3}{6} = 0.98 \Rightarrow \left(\overline{X} - 0,98; \overline{X} + 0.98\right).$$

*Замечание.* Практически важной может быть задача определения объема выборки, которая обеспечит заданный радиус доверительного интервала: $n \geq \frac{t^2\sigma_X^2}{\delta^2}$.

Более точные результаты при малых объемах выборки $n$ и неизвестном $\sigma_X$ дает использование распределения Стьюдента: для переменной $t = \frac{\overline{X} - m_X}{S/\sqrt{n}}$, имеющей распределение Стьюдента с $k = n-1$ степенями свободы, отклонение ($X \sim N(m_X;\sigma_X)$). Тогда доверительный интервал при неизвестном среднеквадратическом отклонении определяется следующим образом:

$$P\left(\overline{X} - \frac{t_\beta \cdot S}{\sqrt{n}} < m_X < \overline{X} + \frac{t_\beta \cdot S}{\sqrt{n}}\right) = \beta,$$

где $t_\beta$ – аргумент табулированного распределения Стьюдента.

## Доверительные интервалы оценки среднеквадратического отклонения

Пусть вновь $x \sim N(m_X,\sigma_X)$ и $\sigma_X$ - неизвестно, а $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2}$. Тогда

$$P(|\sigma_X - \delta| < \delta) = \beta \Rightarrow S - \delta < \sigma_X < S + \delta \ \Rightarrow S(1-q) < \sigma_X < S(1-q),$$

где $q = \frac{\delta}{S}$.

Доказано, что $q$ имеет табулированное распределение $\chi^2$, не зависящее от параметров $m_X$ и $\sigma_X$ исходного распределения, но зависящее от объема выборки и доверительной вероятности. Вычислив по выборке $S$, находим по таблице $q$, определяем границы доверительного интервала.

## Методы оценки параметров известного распределения

Пусть $X$ - с.в.; $\{x_1, x_2, ..., x_n\}$ - выборка; $f(x,\vec{Q})$ - известный вид распределения с.в., $\vec{Q} = \{Q_1, Q_2, ..., Q_l\}$ - неизвестные параметры. Например,

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{(x-m_x)^2}{2\sigma_x}} \Rightarrow \vec{Q} = \{m_X, \sigma_X\};$$

$$f(x) = \frac{1}{b-a} \Rightarrow \vec{Q} = \{b, a\}.$$

Существуют два основных метода оценки параметров (параметризации):

1. Метод моментов (Пирсона).

Метод исходит из того, что оценки начального и центрального моментов распределения являются состоятельными оценками соответствующих начальных и центральных моментов. Это дает право приравнять соответствующие эмпирические и теоретические моменты для точечной оценки параметров, например:

число $\alpha_1^*[x] = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$ = параметру $\alpha_1[x] = \begin{cases} \int_{-\infty}^{\infty} xf(x)dx = \varphi_1(\vec{Q}) \\ \sum_{i=1}^{n} x_i p_i \end{cases}$;

число $\mu_2^*[x] = S^2$ = параметру $\mu_2[x] = \begin{cases} \int_{-\infty}^{\infty} \overset{0}{x}{}^2 f(x)dx = \varphi_2(\vec{Q}) \\ \sum_i (x_i - m_X)^2 p_i \end{cases}$.

При этом надо получить не менее, чем $L$, уравнений для определения $L$ параметров распределения. Выбор начальных или центральных моментов должен производиться исходя из простоты вычислений эмпирических моментов и простоты аналитического выражения теоретических моментов.

***Пример 3.7***. $f(x) = \lambda e^{-\lambda x}; Q = \lambda.$

$$\alpha_1^*[x] = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{X} = \alpha_1[x] = \int_{-\infty}^{\infty} x\lambda e^{-\lambda x}dx = \int_0^{\infty} x\lambda e^{-\lambda x}dx = \frac{1}{\lambda}.$$

$$\overline{X} = \frac{1}{\lambda} \Rightarrow \lambda = \frac{1}{\overline{X}}.$$

***Пример 3.8.*** $f(x) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-\frac{(X-m_x)^2}{2\sigma_x}}$; $\vec{Q} = \{m_X, \sigma_X\}$.

$$\begin{cases} \overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \alpha_1[x] = m_x \\ S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2 = \mu_2[x] = \sigma_x{}^2. \end{cases}$$

Достоинство метода моментов – простота.

2. Метод максимального правдоподобия (Фишера).

На основе выборки составляется функция правдоподобия:

$$L\left\{x_1, x_2, ..., x_l, \vec{Q}\right\} = p\left(x_1, \vec{Q}\right) \cdot p\left(x_2, \vec{Q}\right) \cdot ... \cdot p\left(x_n, \vec{Q}\right),$$

где $p\left(x_1, \vec{Q}\right)$ - есть вероятность того, что первое наблюдение равно $x_1$, при векторе параметров $\vec{Q}$,

$p\left(x_2, \vec{Q}\right)$- вероятность того, что второе наблюдение равно $x_2$, при векторе параметров $\vec{Q}$, и т.д.

В качестве точечной оценки вектора параметров $\vec{Q}$ принимают такое его значение $\vec{Q} = \left\{Q_1, Q_2, ..., Q_l\right\}$, при котором функция правдоподобия достигает максимума.

Чаще для удобства вычислений берут вместо функции правдоподобия ее логарифм: $\ln L\left\{x_1, x_2, ..., x_n, \vec{Q}\right\} = \sum_{i=1}^{n} \ln p(x_i, \vec{Q})$ - логарифмическая функция правдоподобия.

И функция правдоподобия, и логарифмическая функция правдоподобия достигают максимума при одном и том же значении вектора параметров $\vec{Q}$.

Последовательность действий при реализации метода максимального правдоподобия:

1. $L\{...\}$;
2. $\ln L\{...\}$;
3. $\frac{d \ln L\{...\}}{d\vec{Q}} = 0 \Rightarrow \vec{Q}^*$ - уравнение правдоподобия; $\vec{Q}^*$ - корень уравнения;
4. $\frac{d^2 \ln L\{...\}}{d^2\vec{Q}} = \begin{cases} < 0 \Rightarrow \max \\ > 0 \Rightarrow \min \end{cases}$.

Достоинства метода максимального правдоподобия:

1. оценки состоятельны;
2. оценки распределены асимптотически нормальны (при $n \to \infty$ закон распределения оценок стремится к нормальному);
3. оценки имеют наименьшую дисперсию по сравнению с другими асимптотически нормальными оценками (эффективность является важнейшим свойством оценок в экономике);
4. метод наиболее полно использует данные выборки об оцениваемом параметре, поэтому он особенно полезен в случае малых выборок.

Недостатки метода максимального правдоподобия:

1. сложность вычислений;
2. возможна смещенность оценок параметров распределения.

## Глава 4. СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

Статистическое моделирование - численный метод решения математических задач, при котором искомые величины представляют вероятностными характеристиками какого-либо случайного явления, это явление моделируется, после чего нужные характеристики приближённо определяют путём статистической обработки «наблюдений» модели. Например, требуется рассчитать потоки тепла в нагреваемой тонкой металлической пластине, на краях которой поддерживается нулевая температура. Распределение тепла описывается тем же уравнением, что и расплывание пятна краски в слое жидкости. Поэтому моделируют плоские частицы «краски» по пластине, следя за их положениями в моменты $kt$, $k$ = 0, 1, 2,... Приближённо принимают, что за малый интервал $t$ частица перемещается на шаг $h$ равновероятно во всех направлениях. Каждый раз направление выбирается случайным образом, независимо от всего предыдущего. Соотношение между $t$ и $h$ определяется коэффициентом теплопроводности. Движение начинается в источнике тепла и кончается при первом достижении края (наблюдается налипание «краски» на край). Поток $Q$ ($C$) тепла через участок $C$ границы измеряется количеством налипшей краски. При общем количестве $N$ частиц согласно закону больших чисел такая оценка даёт случайную относительную ошибку порядка $\dfrac{1}{\sqrt{N}}$ (и систематическую ошибку порядка $h$ из-за дискретности выбранной модели).

Искомую величину представляют математическим ожиданием числовой функции $f$ от случайного исхода $\omega$ явления: $Ef\left(\omega\right)=\int f\left(\omega\right)dP$, т. е. интегралом по вероятностной мере $P$. На оценку $Ef\left(\omega\right)\approx\left[f\left(\omega_1\right)+f\left(\omega_N\right)\right]/N$, где $\omega_1, ..., \omega_N$ - смоделированные исходы, можно смотреть как на квадратурную формулу для указанного интеграла со случайными узлами ωk и случайной погрешностью RN, обычно принимают $\left|R_N\right|\le 2\sqrt{Df/N}$, считая большую погрешность пренебрежимо маловероятной; дисперсия $Df$ может быть оценена в ходе наблюдений.

В разобранном выше примере $f$ ($\omega$)= 1, когда траектория кончается на $C$; иначе $f$ ($\omega$) = 0. Дисперсия

$$Df=\left[1-Q\left(C\right)\right]Q\left(C\right)\le\frac{1}{4}.$$

Интеграл берётся по пространству ломаных со звеньями постоянной длины; он может быть выражен через кратные интегралы.

Проведение каждого «эксперимента» распадается на две части: «розыгрыш» случайного исхода ω и последующее вычисление функции $f\ (\omega)$. Когда пространство всех исходов и вероятностная мера $P$ слишком сложны, розыгрыш проводится последовательно в несколько этапов. Случайный выбор на каждом этапе проводится с помощью случайных чисел, например генерируемых каким-либо физическим датчиком; употребительна также их арифметическая имитация - псевдослучайные числа. Аналогичные процедуры случайного выбора используются в математической статистике и теории игр.

Статистическое моделирование широко применяется для решения на ЭВМ интегральных уравнений, например при исследовании *больших систем*. Они удобны своей универсальностью, как правило, не требуют большого объёма памяти. Недостаток - большие случайные погрешности, слишком медленно убывающие при увеличении числа экспериментов. Поэтому разработаны приёмы преобразования моделей, позволяющие понижать разброс наблюдаемых величин и объём модельного эксперимента.

## Метод Монте-Карло

Как мы уже отмечали, статистическое моделирование — базовый метод моделирования, заключающийся в том, что модель испытывается множеством случайных сигналов с заданной плотностью вероятности. Целью является статистическое определение выходных результатов. В основе статистического моделирования лежит *метод Монте-Карло*. Напомним, что имитацию используют тогда, когда другие методы применить невозможно.

Рассмотрим метод Монте-Карло на примере вычисления интеграла, значение которого аналитическим способом найти не удается.

Задача 1. Найти значение интеграла:

$$y = \int_{x_1}^{x_2} f\left(x\right) dx \,/.$$

На рис. 4.1 представлен график функции $f(x)$. Вычислить значение интеграла этой функции — значит найти площадь под этим графиком.

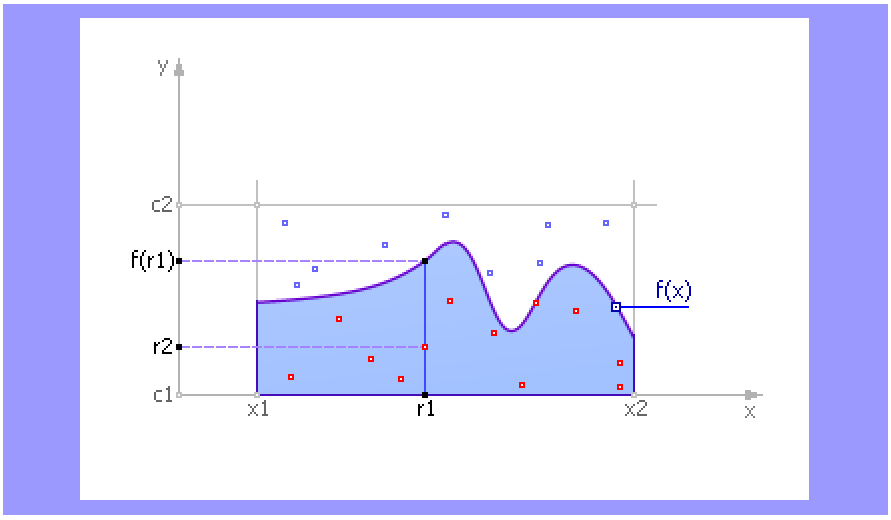



Рис. 4.1. Определение значения интеграла методом Монте-Карло

Ограничиваем кривую сверху, справа и слева. Случайным образом распределяем точки в прямоугольнике поиска. Обозначим через $N_1$ количество точек, принятых для испытаний (то есть попавших в прямоугольник) и через $N_2$ — количество точек под кривой, то есть попавших в закрашенную площадь под функцией. Тогда естественно предположить, что количество точек, попавших под кривую, по отношению к общему числу точек пропорционально площади под кривой (величине интеграла) по отношению к площади испытуемого прямоугольника. Математически это можно выразить так:

$$\frac{N_2}{N_1} = \frac{y}{(x_2 - x_1)(c_2 - c_1)}.$$

Рассуждения эти, конечно, статистические и тем более верны, чем большее число испытуемых точек мы возьмем. Фрагмент алгоритма метода Монте-Карло выглядит так, как показано на рис. 4.2.

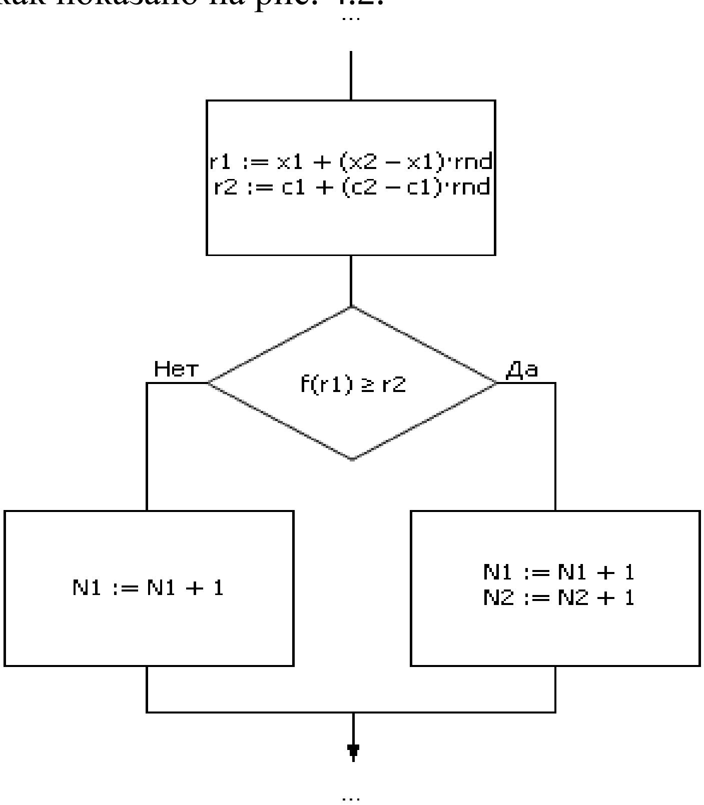



Рис. 4.2. Фрагмент алгоритма реализации метода Монте-Карло

Значения $r_1$ и $r_2$ на рис. 4.2 являются равномерно распределенными случайными числами из интервалов ($x_1$; $x_2$) и ($c_1$; $c_2$) соответственно. Метод Монте-Карло чрезвычайно эффективен, прост, но необходим «хороший» генератор случайных чисел. Вторая проблема применения метода заключается в определении объема выборки, то есть количества точек, необходимых для обеспечения решения с заданной точностью. Эксперименты показывают: чтобы увеличить точность в 10 раз, объем выборки нужно увеличить в 100 раз; то есть точность примерно пропорциональна корню квадратному из объема выборки:

$$\text{точность} \cong \sqrt{\text{объем выборки}}.$$

## Схема использования метода Монте-Карло при исследовании систем со случайными параметрами

Построив модель системы со случайными параметрами, на ее вход подают входные сигналы от генератора случайных чисел (ГСЧ), как показано на рис. 4.3. ГСЧ устроен так, что он выдает *равномерно распределенные* случайные числа $r_{\text{рр}}$ из интервала [0; 1]. Так как одни события могут быть более вероятными, другие — менее вероятными, то равномерно распределенные случайные числа от генератора подают на преобразователь закона случайных чисел (ПЗСЧ), который преобразует их в *заданный* пользователем закон распределения вероятности, например, в нормальный или экспоненциальный закон. Эти преобразованные случайные числа $x$ подают на вход модели. Модель отрабатывает входной сигнал $x$ по некоторому закону $y = \varphi(x)$ и получает выходной сигнал $y$, который также является случайным.

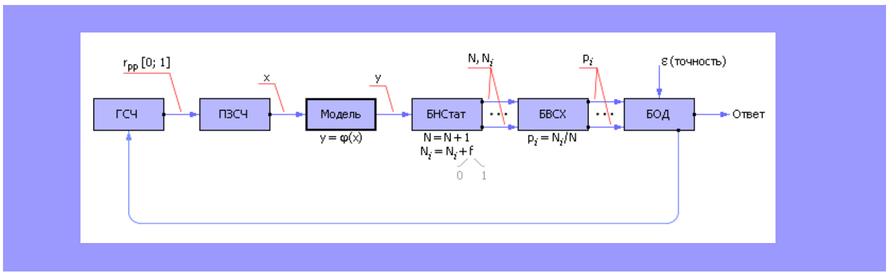



Рис. 4.3. Общая схема метода статистического моделирования

В блоке накопления статистики (БНСтат) установлены фильтры и счетчики. Фильтр (некоторое логическое условие) определяет по значению $y$, реализовалось ли в конкретном опыте некоторое событие (выполнилось условие, $f = 1$) или нет (условие не выполнилось, $f = 0$). Если событие реализовалось, то счетчик события увеличивается на единицу. Если событие не реализовалось, то значение счетчика не меняется. Если требуется следить за несколькими разными типами событий, то для статистического моделирования понадобится несколько фильтров и счетчиков $N_i$. Всегда ведется счетчик количества экспериментов — $N$.

Далее отношение $N_i$ к $N$, рассчитываемое в блоке вычисления статистических характеристик (БВСХ) по методу Монте-Карло, дает оценку вероятности $p_i$ появления события $i$, то есть указывает на частоту его выпадения в серии из $N$ опытов. Это позволяет сделать выводы о статистических свойствах моделируемого объекта.

Например, событие А совершилось в результате проведенных 200 экспериментов 50 раз. Это означает, согласно методу Монте-Карло, что вероятность совершения события равна: $p_{\text{А}} = 50/200 = 0.25$. Вероятность того, что событие не совершится, равна, соответственно, $1 – 0.25 = 0.75$.

Обратите внимание: когда говорят о вероятности, полученной экспериментально, то ее называют частостью; слово вероятность употребляют, когда хотят подчеркнуть, что речь идет о теоретическом понятии.

При большом количестве опытов $N$ частота появления события, полученная экспериментальным путем, стремится к значению теоретической вероятности появления события.

В блоке оценки достоверности (БОД) анализируют степень достоверности статистических экспериментальных данных, снятых с модели (принимая во внимание точность результата $\varepsilon$, заданную пользователем) и определяют необходимое для этого количество статистических испытаний. Если колебания значений частоты появления событий относительно теоретической вероятности меньше заданной точности, то экспериментальную частоту принимают в качестве ответа, иначе генерацию случайных входных воздействий продолжают, и процесс моделирования повторяется. При малом числе испытаний результат может оказаться недостоверным. Но чем более испытаний, тем точнее ответ, согласно центральной предельной теореме.

Заметим, что оценивание ведут по худшей из частот. Это обеспечивает достоверный результат сразу по всем снимаемым характеристикам модели.

***Пример 4.1***. Решим простую задачу. Какова вероятность выпадения монеты орлом кверху при падении ее с высоты случайным образом?

Начнем подбрасывать монетку и фиксировать результаты каждого броска (табл. 4.1).

Таблица 4.1

Результаты испытаний бросания монеты

| Количество опытов *N* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Значение счетчика выпадения орла *No* | 0 | 0 | 1 | 1 | 2 | 3 | 4 | … | … | … | … | … | … | … |
| Значение счетчика выпадения решки *Np* | 1 | 2 | 2 | 3 | 3 | 3 | 3 | … | … | … | … | … | … | … |
| Частость выпадения орла *Po =No/N* | 0 | 0 | 0.33 | 0.25 | 0.4 | 0.5 | 0.57 | … | … | … | … | … | … | … |
| Частость выпадения решки *Pp =Np/N* | 1 | 1 | 0.66 | 0.75 | 0.6 | 0.5 | 0.43 | … | … | … | … | … | … | … |

Будем подсчитывать частость выпадения орла как отношение количества случаев выпадения орла к общему числу наблюдений. Посмотрите в табл. 4.1. Случаи для $N = 1$, $N = 2$, $N = 3$ — сначала значения частости нельзя назвать достоверными. Попробуем построить график зависимости $P_о$ от $N$ — и посмотрим, как меняется частость выпадения орла в зависимости от количества проведенных опытов. Разумеется, при различных экспериментах будут получаться разные таблицы и, следовательно, разные графики. На рис. 4.4 показан один из вариантов.

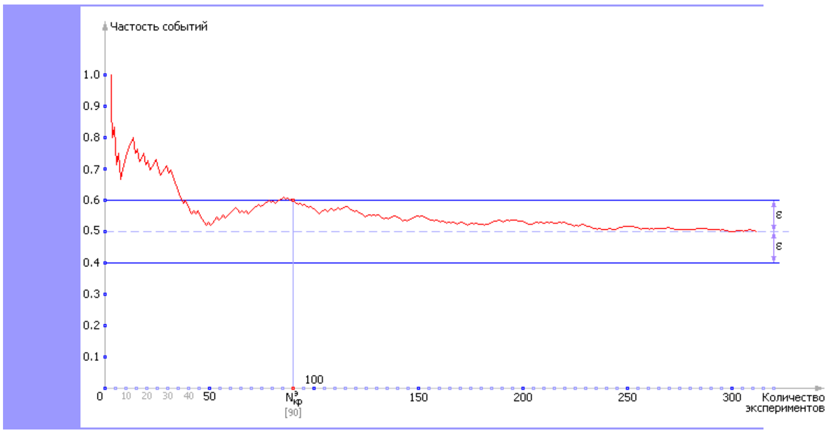



Рис. 4.4. Экспериментальная зависимость частости появления случайного события от количества наблюдений и ее стремление к теоретической вероятности

Сделаем некоторые выводы.

1. Видно, что при малых значениях $N$, например $N = 1$, $N = 2$, $N = 3$, ответу вообще доверять нельзя. Например, $P_о = 0$ при $N = 1$, то есть вероятность выпадения орла при одном броске равна нулю! Хотя всем хорошо известно, что это не так. То есть пока мы получили очень грубый ответ. Однако посмотрите на график: в процессе накопления информации ответ медленно, но верно приближается к правильному (он выделен пунктирной линией). К счастью, в данном конкретном случае правильный ответ нам известен: в идеале, вероятность выпадения орла равна 0.5 (в других, более сложных задачах, ответ нам, конечно, будет неизвестен). Допустим, что ответ нам надо знать с точностью $\varepsilon = 0.1$. Проведем две параллельные линии, отстоящие от правильного ответа 0.5 на расстоянии 0.1 (см. рис. 4.4). Ширина образовавшегося коридора будет равна 0.2. Как только кривая $P_о\,(N)$ войдет в этот коридор так, что уже никогда его не покинет, можно остановиться и посмотреть, для какого значения $N$ это произошло. Это и есть экспериментально вычисленное критическое значение необходимого количества опытов $N_{кр}^{э}$ для определения ответа с точностью $\varepsilon = 0.1$; $\varepsilon$-окрестность в наших рассуждениях играет роль своеобразной трубки точности. Заметьте, что ответы $P_о\,(91)$, $P_о\,(92)$ и так далее уже не меняют сильно своих значений (см. рис. 4.4); по крайней мере, у них не изменяется первая цифра после запятой, которой мы обязаны доверять по условиям задачи.

2. Причиной такого поведения кривой является действие центральной предельной теоремы. Пока здесь мы сформулируем ее в самом простом варианте «Сумма случайных величин есть величина неслучайная». Мы использовали среднюю величину $P_о$, которая несет в себе информацию о сумме опытов, и поэтому постепенно эта величина становится все более достоверной.

3. Если проделать еще раз этот опыт сначала, то, конечно, его результатом будет другой вид случайной кривой. И ответ будет другим, хотя примерно таким же. Проведем целую серию таких экспериментов (рис. 4.5). Такая серия называется ансамблем реализаций. Какому же ответу в итоге

следует верить? Ведь они, хоть и являются близкими, все же разнятся. На практике поступают по-разному. Первый вариант — вычислить среднее значение ответов за несколько реализаций (табл. 4.2).

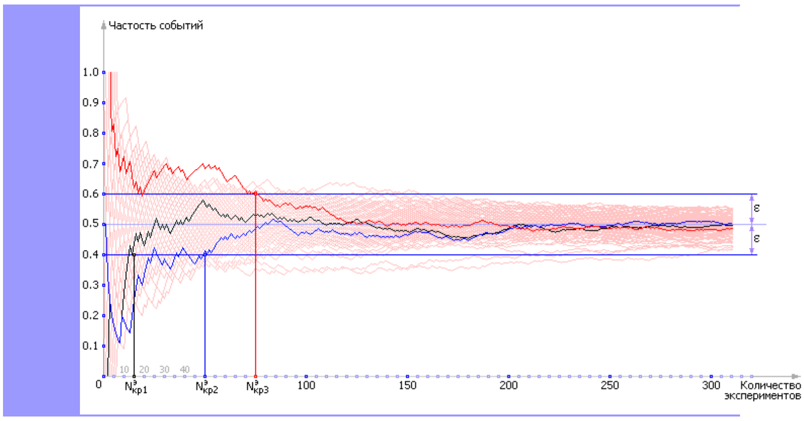



Рис. 4.5. Экспериментально снятый ансамбль случайных зависимостей частости появления случайного события от количества наблюдений

Мы поставили несколько экспериментов и определяли каждый раз, сколько необходимо было сделать опытов, то есть $N_{кр}^{э}$. Было проделано 10 экспериментов, результаты которых были сведены в табл. 4.2. По результатам 10 экспериментов было вычислено среднее значение $N_{кр}^{э}$.

Таблица 4.2

Экспериментальные данные необходимого количества бросков монеты для достижения точности $\varepsilon$ = 0.1 при вычислении вероятности выпадения орла

| Опыт | $N_{кр}^{э}$ |
|---|---|
| 1 | 288 |
| 2 | 95 |
| 3 | 50 |
| 4 | 29 |
| 5 | 113 |
| 6 | 210 |
| 7 | 30 |
| 8 | 42 |
| 9 | 39 |
| 10 | 48 |
| Среднее $N_{кр.}^{э}$ | 94 |

Таким образом, проведя 10 реализаций разной длины, мы определили, что достаточно в среднем было сделать 1 реализацию длиной в 94 броска монеты.

Еще один важный факт. Внимательно рассмотрите график на рис. 4.5. На нем нарисовано 100 реализаций — 100 красных линий. Отметьте на нем абсциссу $N = 94$ вертикальной чертой. Есть какой-то процент красных линий, которые не успели пересечь $\varepsilon$-окрестность, то есть ($P^{эксп} - \varepsilon \leq P^{теор} \leq P^{эксп} + \varepsilon$), и войти в коридор точности до момента $N = 94$. Обратите внимание, таких линий 5. Это значит, что 95 из 100, то есть 95%, линий достоверно вошли в обозначенный интервал.

Таким образом, проведя 100 реализаций, мы добились примерно 95%-ного доверия к полученной экспериментально величине вероятности выпадения орла, определив ее с точностью 0,1. Для сравнения полученного результата вычислим теоретическое значение $N_{кр}^{т}$ теоретически. Однако для этого придется ввести понятие доверительной вероятности $Q_F$, которая показывает, насколько мы готовы верить ответу. Например, при $Q_F = 0.95$ мы готовы верить ответу в 95% случаев из 100. Формула теоретического расчета числа экспериментов имеет вид: $N_{кр}^{т} = k(Q_F) \cdot p \cdot (1 - p)/\varepsilon^2$, где $k(Q_F)$ — коэффициент Лапласа, $p$ — вероятность выпадения орла, $\varepsilon$ — точность (доверительный интервал). В табл. 4.3 показаны значения теоретической величины количества необходимых опытов при разных $Q_F$ (для точности $\varepsilon = 0.1$ и вероятности $p = 0.5$).

Таблица 4.3

Теоретический расчет необходимого количества бросков монеты для достижения точности ε = 0.1 при вычислении вероятности выпадения орла

| Доверительная вероятность *QF* | Коэффициент Лапласа $k(QF)$ | Требуемое число опытов $N$крт $= k(QF)\ p \cdot (1 - p)/\varepsilon 2$ |
|---|---|---|
| 0.90 | 2.72 | 68 |
| 0.95 | 3.84 | 96 |
| 0.99 | 6.66 | 167 |

Как видите, полученная нами оценка длины реализации, равная 94 опытам, очень близка к теоретической, равной 96. Некоторое несовпадение объясняется тем, что, видимо, 10 реализаций недостаточно для точного вычисления $N_{кр}^{э}$. Если вы решите, что вам нужен результат, которому следует доверять больше, то измените значение доверительной вероятности. Например, теория говорит нам, что если опытов будет 167, то всего 1-2 линии из ансамбля не войдут в предложенную трубку точности. Но имейте в виду, количество экспериментов с ростом точности и достоверности растет очень быстро. Второй вариант, используемый на практике, — провести одну реализацию и увеличить полученное для нее $N_{кр}^{э}$ в 2 раза. Это считают хорошей гарантией точности ответа (рис. 4.6).

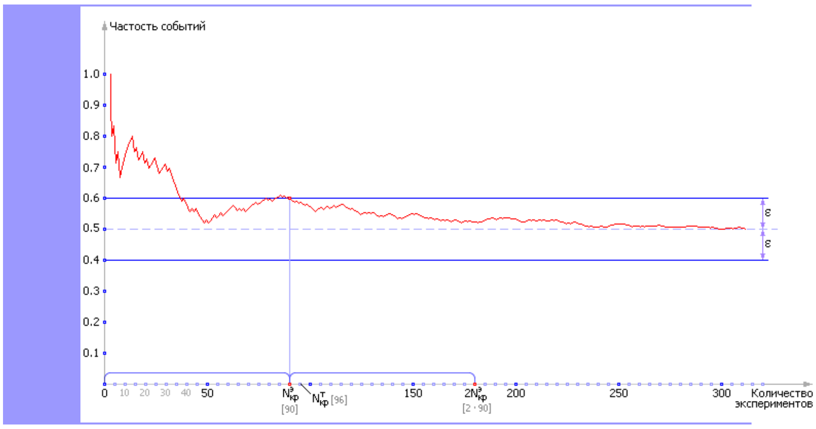



Рис. 4.6. Иллюстрация экспериментального определения $N_{кр}^{э}$
по правилу «умножь на два»

Если присмотреться к ансамблю случайных реализаций, то можно обнаружить, что сходимость частости к значению теоретической вероятности происходит по кривой, соответствующей обратной квадратичной зависимости от числа экспериментов (рис. 4.7).

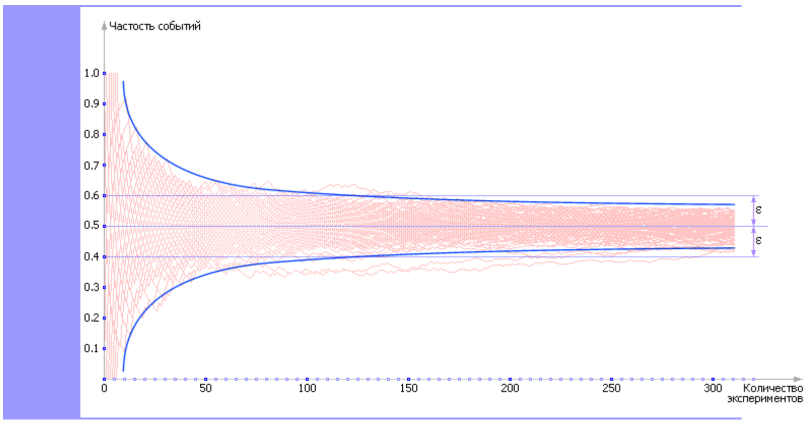



Рис. 4.7. Иллюстрация скорости схождения экспериментально получаемой
частости к теоретической вероятности

Это действительно так получается и теоретически. Если изменять задаваемую точность $\varepsilon$ и исследовать количество экспериментов, требуемых для обеспечения каждой из них, то получится табл. 4.4.

Таблица 4.4

Теоретическая зависимость количества экспериментов, необходимых
для обеспечения заданной точности,при $Q_F$ = 0.95

| Точность $\varepsilon$ | Критическое число экспериментов $N_{кр}^{т}$ |
|---|---|
| 0,1 | 96 |
| 0,01 | 9600 |
| 0,001 | 960000 |

Построим по табл. 4.4 график зависимости $N_{кр}^{т}(\varepsilon)$ (рис. 4.8).

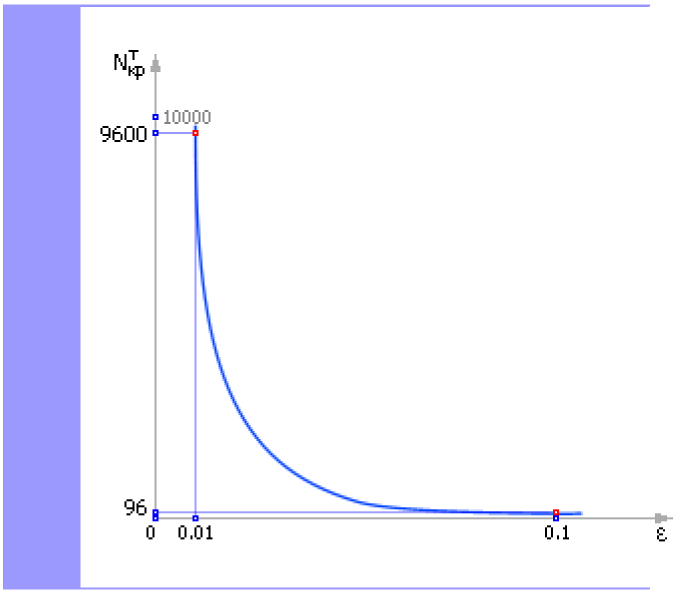



Рис. 4.8. Зависимость числа экспериментов, требуемых для достижения заданной точности ε при фиксированном $Q_F$ = 0,95

Итак, рассмотренные графики подтверждают приведенную выше оценку:

$$\text{точность} \cong \sqrt{\text{объем выборки}}\,.$$

Заметим, что оценок точности может быть несколько.

***Пример 4.2.*** Нахождение площади фигуры методом Монте-Карло. Определите методом Монте-Карло площадь пятиугольника с координатами углов (0, 0), (0, 10), (5, 20), (10, 10), (7, 0).

Нарисуем в двухмерных координатах заданный пятиугольник, вписав его в прямоугольник, чья площадь, как нетрудно догадаться, составляет (10 – 0) · (20 – 0) = 200 (рис. 4.9).

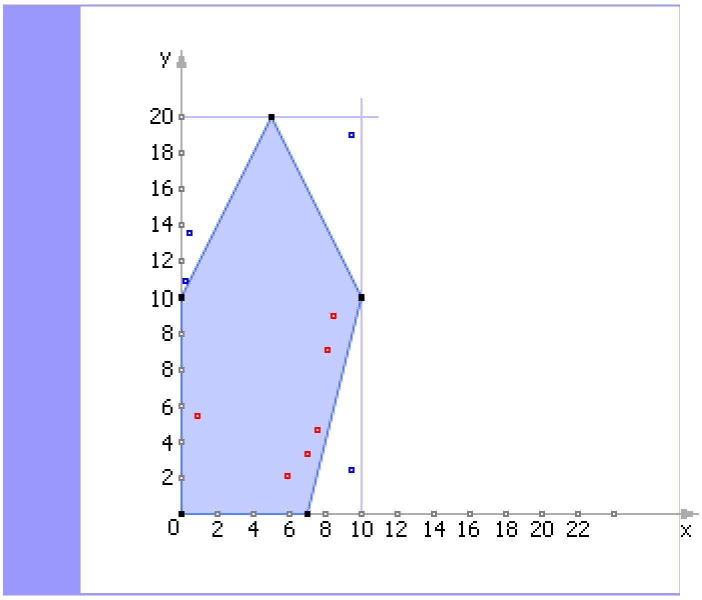



Рис. 4.9. Иллюстрация к решению задачи о площади фигуры методом Монте-Карло

Используем таблицу случайных чисел для генерации пар чисел $R, G$, равномерно распределенных в интервале от 0 до 1. Число $R$ будет имитировать координату $X$ $(0 \le X \le 10)$, следовательно, $X = 10 \cdot R$. Число $G$ будет имитировать координату $Y$ $(0 \le Y \le 20)$, следовательно, $Y = 20 \cdot G$. Сгенерируем по 10 чисел $R$ и $G$ и отобразим 10 точек $(X; Y)$ на рис. 4.9 и в табл. 4.5.

Таблица 4.5

Решение задачи методом Монте-Карло

| Номер точки | R | G | X | Y | Точка (X; Y) попала в прямоугольник? | Точка (X; Y) попала в пятиугольник? |
|---|---|---|---|---|---|---|
| 1 | 0.8109 | 0.3557 | 8.109 | 7.114 | Да | Да |
| 2 | 0.0333 | 0.5370 | 0.333 | 10.740 | Да | Нет |
| 3 | 0.1958 | 0.2748 | 1.958 | 5.496 | Да | Да |
| 4 | 0.6982 | 0.1652 | 6.982 | 3.304 | Да | Да |
| 5 | 0.9499 | 0.1090 | 9.499 | 2.180 | Да | Нет |
| 6 | 0.7644 | 0.2194 | 7.644 | 4.388 | Да | Да |
| 7 | 0.8395 | 0.4510 | 8.395 | 9.020 | Да | Да |
| 8 | 0.0415 | 0.6855 | 0.415 | 13.710 | Да | Нет |
| 9 | 0.5997 | 0.1140 | 5.997 | 2.280 | Да | Да |
| 10 | 0.9595 | 0.9595 | 9.595 | 19.190 | Да | Нет |
| Всего: | | | | | 10 | 6 |

Статистическая гипотеза заключается в том, что количество точек, попавших в контур фигуры, пропорционально площади фигуры: $6:10 = S:200$. То есть, по формуле метода Монте-Карло, получаем, что площадь $S$ пятиугольника равна: $200 \cdot 6/10 = 120$.

Проследим, как менялась величина $S$ от опыта к опыту (табл. 4.6).

Таблица 4.6

Оценка точности ответа

| Количество испытаний N | Оценка вероятности попадания случайной точки в испытуемую область | Оценка площади S методом Монте-Карло |
|---|---|---|
| 1 | 1/1 = 1.00 | 200 |
| 2 | 1/2 = 0.50 | 100 |
| 3 | 2/3 = 0.67 | 133 |
| 4 | 3/4 = 0.75 | 150 |
| 5 | 3/5 = 0.60 | 120 |
| 6 | 4/6 = 0.67 | 133 |
| 7 | 5/7 = 0.71 | 143 |
| 8 | 5/8 = 0.63 | 125 |
| 9 | 6/9 = 0.67 | 133 |
| 10 | 6/10 = 0.60 | 120 |

Поскольку в ответе все еще меняется значение второго разряда, то возможная неточность составляет пока больше 10%. Точность расчета может быть увеличена с ростом числа испытаний (рис. 4.10).

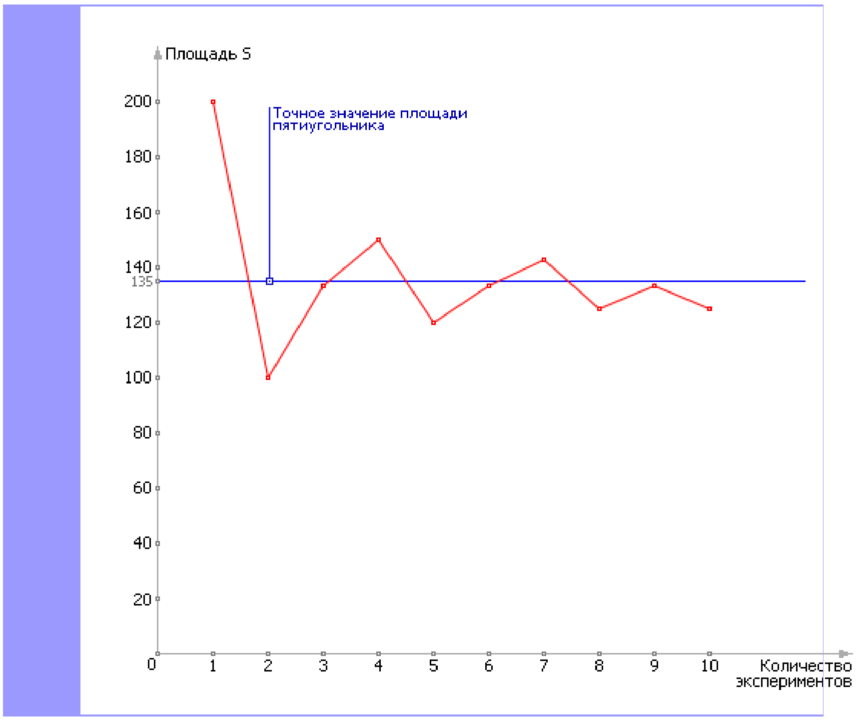



Рис. 4.10. Иллюстрация процесса сходимости определяемого экспериментально ответа к теоретическому результату

## ЛАБОРАТОРНАЯ РАБОТА «СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ»

### 1. Цель работы

Целью работы является ознакомление студентов с возможностями метода статистического моделирования – метода Монте-Карло.

Метод Монте-Карло в лабораторной работе используется для решения двух задач:

- нахождения определенного интеграла функции (цель – получение общих представлений о методе);
- моделирования работы робототехнического комплекса.

Работа состоит из двух соответствующих частей, посвященных рассматриваемым задачам.

### 2. Теоретические сведения

Метод Монте-Карло является методом статистического моделирования. Его применение эффективно там, где сложно или невозможно построение аналитической модели. Например, в системах массового обслуживания, не являющихся марковскими системами, в задачах надежности, управления, экономики и т.п., вообще для сложных систем, которые состоят из большого числа взаимодействующих элементов.

Идея метода заключается в следующем. Вместо того, чтобы описывать процесс с помощью аналитического аппарата (дифференциальных или алгебраических уравнений), производится «розыгрыш» случайного явления с помощью специально организованной процедуры, включающей в себя случайность и дающей случайный результат реализации процесса. Множество реализаций можно использовать как некий искусственно полученный статистический материал, который может быть обработан обычными методами математической статистики и получены интересующие нас статистические характеристики. При моделировании случайных явлений методом Монте-Карло мы пользуемся самой случайностью как аппаратом исследования.

В математике метод Монте-Карло применяется для вычисления интегралов, особенно многомерных, для решения систем алгебраических уравнений высокого порядка и т.п.

Метод Монте-Карло имеет простую структуру вычислительного алгоритма. Как правило, составляется программа для осуществления одного случайного испытания. Затем это испытание повторяется N раз, причем каждый опыт не зависит от всех остальных, и результаты всех опытов усредняются. Поэтому метод Монте-Карло называют также методом статистического моделирования.

Погрешность вычисления метода, как правило, пропорциональна $\sqrt{D/N}$, где $D$ – некоторая постоянная, $N$ – число испытаний. Отсюда видно, что для того, чтобы уменьшить погрешность в 10 раз, нужно увеличить $N$ в 100 раз.

Приведем важное для метода Монте-Карло соотношение:

$$P\left\{\left|\frac{1}{N}\sum\xi_i - m\right| \prec \frac{3b}{\sqrt{N}}\right\} \approx 0{,}997, \qquad (1)$$

где $\xi_i$ - случайная величина;

$m$ – неизвестная искомая величина (статистическую оценку которой необходимо получить);

$b$ – среднеквадратичное отклонение случайной величины $\xi_i$ (одинаково для всех $\xi_i$).

Соотношение (1) дает нам метод расчета $m$ и оценку погрешности. В самом деле, найдем $N$ значений случайной величины $\xi_i$. Из (1) видно, что среднее арифметическое этих значений будет приближенно равно $m$. С большой вероятностью погрешность такого приближения не превосходит величины $\frac{3b}{\sqrt{N}}$. Очевидно, эта погрешность стремится к нулю с ростом $N$.

В рассматриваемой работе требуется определить среднее время работы РТК. При этом известны интенсивность потока отказов и схема поточной линии.

Известно, что поток отказов подчиняется экспоненциальному (показательному) закону распределения с плотностью распределения

$$f(t) = \lambda e^{-\lambda t},$$

где $\lambda$ - интенсивность потока отказов (число отказов в сутки);

$t$ –последовательности значений продолжительности интервалов между отказами.

Требуется смоделировать случайную величину, распределенную в соответствии с экспоненциальным законом.

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 выбрать случайное число $\xi$ и решить уравнение $F(x) = \xi$ относительно *x*.

Для случая экспоненциального распределения выразим $t_i$ через $\xi_i$

$$t_i = x_i = -\frac{1}{\lambda}\ln \xi_i. \tag{2}$$

Значения $\xi_i$ определяем по генератору случайных чисел в *MS EXCEL* (или другой программе) или по таблице случайных чисел.

Рассмотрим применение метода Монте-Карло для статистической оценки некоторого определенного интеграла

$$I = \int\limits_{\Omega} \iint \ldots \int f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n,$$

где $\Omega$ - область в *n*-мерном пространстве.

В лабораторной работе для простоты вычислений рассматривается определенный интеграл от функции одной переменной

$$I = \int\limits_a^b f(x) dx. \tag{3}$$

Требуется найти его статистическую оценку $\bar{I}$. Задача сводится к оценке отношения площади криволинейной трапеции, соответствующей некоторому определенному интегралу, к площади квадрата, в который этот интеграл может быть вписан, т.е. имеющий координаты: (*a, 0*), (*b, 0*), (*a, b-a*), (*b, b-a*).

Идея метода заключается в следующем. Выберем пару случайных чисел *x:* $x \in [a, b]$ и *y:* $y \in [0, b-a]$ – их можно рассматривать как координаты случайной точки в указанном квадрате. Затем выберем следующую пару чисел и т.д. Когда число выбранных таким образом точек станет достаточно большим, они более-менее равномерно покроют данный квадрат. При этом множество точек $N$, попавших под кривую $y = f(x)$, будет пропорционально площади криволинейной трапеции, а множество всех точек $M$ – площади квадрата.

Тогда статистическая оценка искомого интеграла найдется по формуле

$$\bar{I} = S\frac{N}{M}, \tag{4}$$

где $S$ – площадь квадрата со стороной $b$-$a$.

Погрешность (абсолютная) $\Delta$ может быть найдена из разности:

$$\Delta = \left|I - \bar{I}\right|. \tag{5}$$

## 3. Задание и методика выполнения работы

Задание по лабораторной работе состоит из двух частей. В первой части для общего ознакомления с методом статистического моделирования рассматривается задача нахождения определенного интеграла. Во второй части рассматривается практическая задача определения среднего времени безотказной работы робототехнического комплекса (РТК).

### 3.1.Вычисление определенного интеграла методом Монте-Карло

*Задание*

Выбрать вариант интеграла функции $f(x)$ по указанию преподавателя из таблицы 1

Таблица 1

Вариант выбора определенного интеграла

| Вариант | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Интеграл | $\int_0^1 \sin x dx$ | $\int_0^1 \cos x dx$ | $\int_0^2 x^{-1} dx$ | $\int_0^2 x^{1/2} dx$ | $\int_0^2 e^x dx$ |

Требуется:

1) вычислить определенный интеграл функции методом Монте-Карло;

2) определить зависимость точности метода в зависимости от мощности множества $M$.

*Методика выполнения работы*

1) Сгенерируем множество $M$ точек на плоскости внутри квадрата с координатами ($a$, $0$), ($b$, $0$), ($a$, $b$-$a$), ($b$, $b$-$a$). Каждая точка $m_i \in M$ имеет координаты $(x_i, y_i)$. Генерация множества $M$ производится путем использования *Microsoft Excel* при числе параметров 2 , $M$ = 50 и заданных границах (рис. 1).

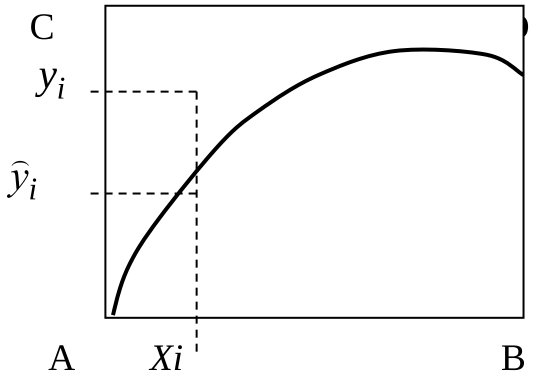



Рис. 1. Схема оценки определенного интеграла

2) Для каждой абсциссы точки $x_i$ определяем статистическую оценку функции $f(x\ )$, которую обозначим через $\widehat{y}_i$.

3) Вычислим для каждой точки $m_i$ разность $y_i - \widehat{y}_i$.

4) Если $y_i - \widehat{y}_i \le 0$ - точка принадлежит криволинейной трапеции, т.е. находится под кривой $f(x\ )$. Найдем множество N элементов $n_i$, для которых выполняется условие: $y_i - \widehat{y}_i \le 0$.

5) Статистическая оценка искомого интеграла по методу Монте-Карло найдется по формуле (4)

6) Определим погрешность метода Монте-Карло по формуле (5).

7) Для определения зависимости $\Delta = f(M)$ зададимся рядом множеств $M_j$: 50;100; 150 и определим в каждом случае $\Delta$.

8) Построим график зависимости $\Delta = f(M)$ в *Microsoft Excel.*

### 3.2. Определение среднего времени безотказной работы РТК

*Задание*

В таблице 2 задана интенсивность потока отказов для оборудования РТК.

Таблица 2

Наименование оборудования РТК и интенсивность потока отказов

| № на схеме | Наименование | Интенсивность потока отказов $\lambda$ (число отказов в сутки) |
|---|---|---|
| 1 | Токарный с ЧПУ | 0,2 |
| 2 | Токарный с ЧПУ | 0,2 |
| 3 | Фрезерный | 0,3 |
| 4 | Фрезерный | 0,32 |
| 5 | Шлифовальный | 0,35 |

Задана схема организации поточной линии, изображенная на рис. 2.

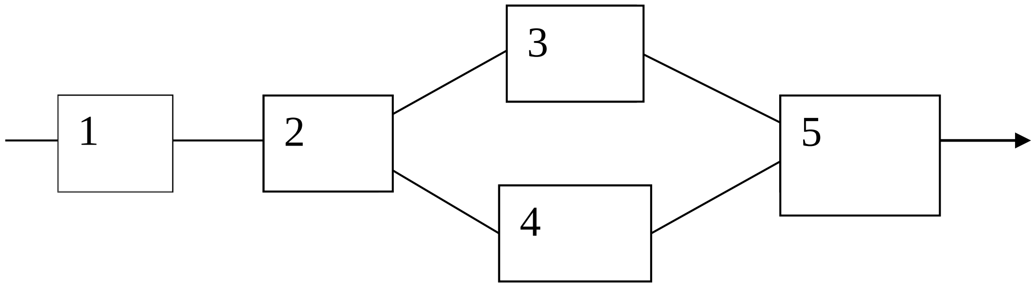



Рис. 2. Схема организации поточной линии

При отказе хотя бы одного вида оборудования линия останавливается.

Требуется найти среднее время безотказной работы РТК.

*Методика выполнения работы*

Для каждого *j*-го вида оборудования (*j*=1,5) рассчитываются по формулам 1, 2 и 3:

1) Выбираем *N* случайных чисел в интервале [0,1]: $\xi_1, \xi_2, ..., \xi_N$ .

Рекомендуется принять *N*=50.

2) Находим по формуле (2) при заданной интенсивности потока $\lambda$ (таблица 2): $t_1,\ t_2,\ ...,\ t_N$ .

3) Находим среднее время безотказной работы j-го вида оборудования

$$\bar{t}_j = \frac{1}{N}\sum_{i=1}^{N} t_i .$$

4) Вычисляем среднее время безотказной работы РТК. Для схемы, изображенной на рис. 2, оно определяется по формуле

$$\bar{t} = \min\left(\bar{t}_1, \bar{t}_2, max(\bar{t}_3, \bar{t}_4), \bar{t}_5\right). \qquad (6)$$

## 4. Содержание отчета

В отчете необходимо представить:

*по первой части работы*:

1) краткие теоретические сведения;

2) таблицы со сгенерированными случайными числами мощностью множества М = 50; 100 и 150;

3) диаграмму зависимости модуля абсолютной ошибки вычислений от М;

4)выводы по первой части работы;

*по второй части работы*:

1)таблицу со сгенерированными случайными числами $\xi_1, \xi_2, ..., \xi_N$ при *N* = 50;

2) таблицу расчета величин $t_1, t_{2,...,} t_N$ ;

3) расчет среднего времени безотказной работы РТК;

4) выводы по второй части работы.

## 5. Контрольные вопросы

1) Области применения метода Монте-Карло (статистического моделирования).

2) От каких условий зависит точность метода Монте-Карло?

3) Как изменится расчетная формула (6) при последовательном расположении оборудования РТК?

4) Можно ли находить с помощью метода Монте-Карло площади фигур в многомерной области?

## СПИСОК ЛИТЕРАТУРЫ


1. Бережная Е.В., Бережной В.И. Математические методы моделирования экономических систем: Учеб. пособие / Е.В. Бережная, В.И. Бережной –М.: Финансы и статистика, 2002. – 368 с.
2. Введение в математическое моделирование: Учеб. пособие / под ред. П.В. Трусова. – М.: Университетская книга, Логос, 2007. – 440 с.
3. Вентцель Е.С. Исследование операций. Задачи, принципы, методология. Учеб. пособие для студентов втузов.- 2-е изд., стер. – М.: Высш. шк., 2001. – 208 с.
4. Гмурман В.Е. Теория вероятностей и математическая статистика: Учеб. пособие для вузов / В.Е. Гмурман. – 9-е изд., стер. – М.: Высш. шк., 2003. – 479 с.
5. Соболь И.М. Метод Монте-Карло / И.М. Соболь – 4-е изд. – М.: Наука. Главная редакция физико-математической литературы, 1985. – 80 с.
6. Соболь И.М. Численные методы Монте-Карло / И.М. Соболь – М.: Наука. Главная редакция физико-математической литературы, 1973. – 312 с.
7. Щипачев А.М. Системный анализ и математическое моделирование процессов в машиностроении: учебное пособие / А.М. Щипачев; Уфимск. гос. авиац. техн. ун-т.- Уфа, 2008. -173 с.

ПРИЛОЖЕНИЯ

Приложение 1

Таблица функции $\varphi(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-z^2}{2}}$

(кривая вероятностей)

| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0,0 | 0,3989 | 0,3989 | 0,3989 | 0,3988 | 0,3986 | 0,3984 | 0,3982 | 0,3980 | 0,3977 | 0,3973 |
| 0,1 | 0,3970 | 0,3965 | 0,3961 | 0,3956 | 0,3951 | 0,3945 | 0,3939 | 0,3932 | 0,3925 | 0,3918 |
| 0,2 | 0,3910 | 0,3902 | 0,3894 | 0,3885 | 0,3876 | 0,3867 | 0,3857 | 0,3847 | 0,3836 | 0,3825 |
| 0,3 | 0,3814 | 0,3802 | 0,3790 | 0,3778 | 0,3765 | 0,3752 | 0,3739 | 0,3725 | 0,3712 | 0,3697 |
| 0,4 | 0,3683 | 0,3668 | 0,3653 | 0,3637 | 0,3621 | 0,3605 | 0,3589 | 0,3572 | 0,3555 | 0,3538 |
| 0,5 | 0,3521 | 0,3503 | 0,3485 | 0,3467 | 0,3448 | 0,3429 | 0,3410 | 0,3391 | 0,3372 | 0,3352 |
| 0,6 | 0,3332 | 0,3312 | 0,3292 | 0,3271 | 0,3251 | 0,3230 | 0,3209 | 0,3187 | 0,3166 | 0,3144 |
| 0,7 | 0,3123 | 0,3101 | 0,3079 | 0,3056 | 0,3034 | 0,3011 | 0,2989 | 0,2966 | 0,2943 | 0,2920 |
| 0,8 | 0,2897 | 0,2874 | 0,2850 | 0,2827 | 0,2803 | 0,2780 | 0,2756 | 0,2732 | 0,2709 | 0,2685 |
| 0,9 | 0,2661 | 0,2637 | 0,2613 | 0,2589 | 0,2565 | 0,2541 | 0,2516 | 0,2492 | 0,2468 | 0,2444 |
| 1,0 | 0,2420 | 0,2396 | 0,2371 | 0,2347 | 0,2323 | 0,2299 | 0,2275 | 0,2251 | 0,2227 | 0,2203 |
| 1,1 | 0,2179 | 0,2155 | 0,2131 | 0,2107 | 0,2083 | 0,2059 | 0,2036 | 0,2012 | 0,1989 | 0,1965 |
| 1,2 | 0,1942 | 0,1919 | 0,1895 | 0,1872 | 0,1849 | 0,1826 | 0,1804 | 0,1781 | 0,1758 | 0,1736 |
| 1,3 | 0,1714 | 0,1691 | 0,1669 | 0,1647 | 0,1626 | 0,1604 | 0,1582 | 0,1561 | 0,1539 | 0,1518 |
| 1,4 | 0,1497 | 0,1476 | 0,1456 | 0,1435 | 0,1415 | 0,1394 | 0,1374 | 0,1354 | 0,1334 | 0,1315 |
| 1,5 | 0,1295 | 0,1276 | 0,1257 | 0,1238 | 0,1219 | 0,1200 | 0,1182 | 0,1163 | 0,1145 | 0,1127 |
| 1,6 | 0,1109 | 0,1092 | 0,1074 | 0,1057 | 0,1040 | 0,1023 | 0,1006 | 0,0989 | 0,0973 | 0,0957 |
| 1,7 | 0,0940 | 0,0925 | 0,0909 | 0,0893 | 0,0878 | 0,0863 | 0,0848 | 0,0833 | 0,0818 | 0,0804 |
| 1,8 | 0,0790 | 0,0775 | 0,0761 | 0,0748 | 0,0734 | 0,0721 | 0,0707 | 0,0694 | 0,0681 | 0,0669 |
| 1,9 | 0,0656 | 0,0644 | 0,0632 | 0,0620 | 0,0608 | 0,0596 | 0,0584 | 0,0573 | 0,0562 | 0,0551 |
| 2,0 | 0,0540 | 0,0529 | 0,0519 | 0,0508 | 0,0498 | 0,0488 | 0,0478 | 0,0468 | 0,0459 | 0,0449 |
| 2,1 | 0,0440 | 0,0431 | 0,0422 | 0,0413 | 0,0404 | 0,0396 | 0,0387 | 0,0379 | 0,0371 | 0,0363 |
| 2,2 | 0,0355 | 0,0347 | 0,0339 | 0,0332 | 0,0325 | 0,0317 | 0,0310 | 0,0303 | 0,0297 | 0,0290 |
| 2,3 | 0,0283 | 0,0277 | 0,0270 | 0,0264 | 0,0258 | 0,0252 | 0,0246 | 0,0241 | 0,0235 | 0,0229 |
| 2,4 | 0,0224 | 0,0219 | 0,0213 | 0,0208 | 0,0203 | 0,0198 | 0,0194 | 0,0189 | 0,0184 | 0,0180 |
| 2,5 | 0,0175 | 0,0171 | 0,0167 | 0,0163 | 0,0158 | 0,0154 | 0,0151 | 0,0147 | 0,0143 | 0,0139 |
| 2,6 | 0,0136 | 0,0132 | 0,0129 | 0,0126 | 0,0122 | 0,0119 | 0,0116 | 0,0113 | 0,0110 | 0,0107 |
| 2,7 | 0,0104 | 0,0101 | 0,0099 | 0,0096 | 0,0093 | 0,0091 | 0,0088 | 0,0086 | 0,0084 | 0,0081 |
| 2,8 | 0,0079 | 0,0077 | 0,0075 | 0,0073 | 0,0071 | 0,0069 | 0,0067 | 0,0065 | 0,0063 | 0,0061 |
| 2,9 | 0,0060 | 0,0058 | 0,0056 | 0,0055 | 0,0053 | 0,0051 | 0,0050 | 0,0048 | 0,0047 | 0,0046 |
| 3,0 | 0,0044 | 0,0043 | 0,0042 | 0,0040 | 0,0039 | 0,0038 | 0,0037 | 0,0036 | 0,0035 | 0,0034 |
| 3,1 | 0,0033 | 0,0032 | 0,0031 | 0,0030 | 0,0029 | 0,0028 | 0,0027 | 0,0026 | 0,0025 | 0,0025 |
| 3,2 | 0,0024 | 0,0023 | 0,0022 | 0,0022 | 0,0021 | 0,0020 | 0,0020 | 0,0019 | 0,0018 | 0,0018 |
| 3,3 | 0,0017 | 0,0017 | 0,0016 | 0,0016 | 0,0015 | 0,0015 | 0,0014 | 0,0014 | 0,0013 | 0,0013 |
| 3,4 | 0,0012 | 0,0012 | 0,0012 | 0,0011 | 0,0011 | 0,0010 | 0,0010 | 0,0010 | 0,0009 | 0,0009 |
| 3,5 | 0,0009 | 0,0008 | 0,0008 | 0,0008 | 0,0008 | 0,0007 | 0,0007 | 0,0007 | 0,0007 | 0,0006 |
| 3,6 | 0,0006 | 0,0006 | 0,0006 | 0,0005 | 0,0005 | 0,0005 | 0,0005 | 0,0005 | 0,0005 | 0,0004 |
| 3,7 | 0,0004 | 0,0004 | 0,0004 | 0,0004 | 0,0004 | 0,0004 | 0,0003 | 0,0003 | 0,0003 | 0,0003 |
| 3,8 | 0,0003 | 0,0003 | 0,0003 | 0,0003 | 0,0003 | 0,0002 | 0,0002 | 0,0002 | 0,0002 | 0,0002 |
| 3,9 | 0,0002 | 0,0002 | 0,0002 | 0,0002 | 0,0002 | 0,0002 | 0,0002 | 0,0002 | 0,0001 | 0,0001 |
| 4,0 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
| 4,1 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
| 4,2 | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |

Приложение 2

Таблица функции $\mathbf{\Phi_0(z)=\int\limits_0^z \frac{1}{\sqrt{2\pi}}\cdot e^{\frac{-x^2}{2}}dx}$ (функция Лапласа)

| z | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0,0 | 0,00000 | 0,00399 | 0,00798 | 0,01197 | 0,01595 | 0,01994 | 0,02392 | 0,02790 | 0,03188 | 0,03586 |
| 0,1 | 0,03983 | 0,04380 | 0,04776 | 0,05172 | 0,05567 | 0,05962 | 0,06356 | 0,06749 | 0,07142 | 0,07535 |
| 0,2 | 0,07926 | 0,08317 | 0,08706 | 0,09095 | 0,09483 | 0,09871 | 0,10257 | 0,10642 | 0,11026 | 0,11409 |
| 0,3 | 0,11791 | 0,12172 | 0,12552 | 0,12930 | 0,13307 | 0,13683 | 0,14058 | 0,14431 | 0,14803 | 0,15173 |
| 0,4 | 0,15542 | 0,15910 | 0,16276 | 0,16640 | 0,17003 | 0,17364 | 0,17724 | 0,18082 | 0,18439 | 0,18793 |
| 0,5 | 0,19146 | 0,19497 | 0,19847 | 0,20194 | 0,20540 | 0,20884 | 0,21226 | 0,21566 | 0,21904 | 0,22240 |
| 0,6 | 0,22575 | 0,22907 | 0,23237 | 0,23565 | 0,23891 | 0,24215 | 0,24537 | 0,24857 | 0,25175 | 0,25490 |
| 0,7 | 0,25804 | 0,26115 | 0,26424 | 0,26730 | 0,27035 | 0,27337 | 0,27637 | 0,27935 | 0,28230 | 0,28524 |
| 0,8 | 0,28814 | 0,29103 | 0,29389 | 0,29673 | 0,29955 | 0,30234 | 0,30511 | 0,30785 | 0,31057 | 0,31327 |
| 0,9 | 0,31594 | 0,31859 | 0,32121 | 0,32381 | 0,32639 | 0,32894 | 0,33147 | 0,33398 | 0,33646 | 0,33891 |
| 1,0 | 0,34134 | 0,34375 | 0,34614 | 0,34849 | 0,35083 | 0,35314 | 0,35543 | 0,35769 | 0,35993 | 0,36214 |
| 1,1 | 0,36433 | 0,36650 | 0,36864 | 0,37076 | 0,37286 | 0,37493 | 0,37698 | 0,37900 | 0,38100 | 0,38298 |
| 1,2 | 0,38493 | 0,38686 | 0,38877 | 0,39065 | 0,39251 | 0,39435 | 0,39617 | 0,39796 | 0,39973 | 0,40147 |
| 1,3 | 0,40320 | 0,40490 | 0,40658 | 0,40824 | 0,40988 | 0,41149 | 0,41308 | 0,41466 | 0,41621 | 0,41774 |
| 1,4 | 0,41924 | 0,42073 | 0,42220 | 0,42364 | 0,42507 | 0,42647 | 0,42785 | 0,42922 | 0,43056 | 0,43189 |
| 1,5 | 0,43319 | 0,43448 | 0,43574 | 0,43699 | 0,43822 | 0,43943 | 0,44062 | 0,44179 | 0,44295 | 0,44408 |
| 1,6 | 0,44520 | 0,44630 | 0,44738 | 0,44845 | 0,44950 | 0,45053 | 0,45154 | 0,45254 | 0,45352 | 0,45449 |
| 1,7 | 0,45543 | 0,45637 | 0,45728 | 0,45818 | 0,45907 | 0,45994 | 0,46080 | 0,46164 | 0,46246 | 0,46327 |
| 1,8 | 0,46407 | 0,46485 | 0,46562 | 0,46638 | 0,46712 | 0,46784 | 0,46856 | 0,46926 | 0,46995 | 0,47062 |
| 1,9 | 0,47128 | 0,47193 | 0,47257 | 0,47320 | 0,47381 | 0,47441 | 0,47500 | 0,47558 | 0,47615 | 0,47670 |
| 2,0 | 0,47725 | 0,47778 | 0,47831 | 0,47882 | 0,47932 | 0,47982 | 0,48030 | 0,48077 | 0,48124 | 0,48169 |
| 2,1 | 0,48214 | 0,48257 | 0,48300 | 0,48341 | 0,48382 | 0,48422 | 0,48461 | 0,48500 | 0,48537 | 0,48574 |
| 2,2 | 0,48610 | 0,48645 | 0,48679 | 0,48713 | 0,48745 | 0,48778 | 0,48809 | 0,48840 | 0,48870 | 0,48899 |
| 2,3 | 0,48928 | 0,48956 | 0,48983 | 0,49010 | 0,49036 | 0,49061 | 0,49086 | 0,49111 | 0,49134 | 0,49158 |
| 2,4 | 0,49180 | 0,49202 | 0,49224 | 0,49245 | 0,49266 | 0,49286 | 0,49305 | 0,49324 | 0,49343 | 0,49361 |
| 2,5 | 0,49379 | 0,49396 | 0,49413 | 0,49430 | 0,49446 | 0,49461 | 0,49477 | 0,49492 | 0,49506 | 0,49520 |
| 2,6 | 0,49534 | 0,49547 | 0,49560 | 0,49573 | 0,49585 | 0,49598 | 0,49609 | 0,49621 | 0,49632 | 0,49643 |
| 2,7 | 0,49653 | 0,49664 | 0,49674 | 0,49683 | 0,49693 | 0,49702 | 0,49711 | 0,49720 | 0,49728 | 0,49736 |
| 2,8 | 0,49744 | 0,49752 | 0,49760 | 0,49767 | 0,49774 | 0,49781 | 0,49788 | 0,49795 | 0,49801 | 0,49807 |
| 2,9 | 0,49813 | 0,49819 | 0,49825 | 0,49831 | 0,49836 | 0,49841 | 0,49846 | 0,49851 | 0,49856 | 0,49861 |
| 3,0 | 0,49865 | 0,49869 | 0,49874 | 0,49878 | 0,49882 | 0,49886 | 0,49889 | 0,49893 | 0,49896 | 0,49900 |
| 3,1 | 0,49903 | 0,49906 | 0,49910 | 0,49913 | 0,49916 | 0,49918 | 0,49921 | 0,49924 | 0,49926 | 0,49929 |
| 3,2 | 0,49931 | 0,49934 | 0,49936 | 0,49938 | 0,49940 | 0,49942 | 0,49944 | 0,49946 | 0,49948 | 0,49950 |
| 3,3 | 0,49952 | 0,49953 | 0,49955 | 0,49957 | 0,49958 | 0,49960 | 0,49961 | 0,49962 | 0,49964 | 0,49965 |
| 3,4 | 0,49966 | 0,49968 | 0,49969 | 0,49970 | 0,49971 | 0,49972 | 0,49973 | 0,49974 | 0,49975 | 0,49976 |
| 3,5 | 0,49977 | 0,49978 | 0,49978 | 0,49979 | 0,49980 | 0,49981 | 0,49981 | 0,49982 | 0,49983 | 0,49983 |
| 3,6 | 0,49984 | 0,49985 | 0,49985 | 0,49986 | 0,49986 | 0,49987 | 0,49987 | 0,49988 | 0,49988 | 0,49989 |
| 3,7 | 0,49989 | 0,49990 | 0,49990 | 0,49990 | 0,49991 | 0,49991 | 0,49992 | 0,49992 | 0,49992 | 0,49992 |
| 3,8 | 0,49993 | 0,49993 | 0,49993 | 0,49994 | 0,49994 | 0,49994 | 0,49994 | 0,49995 | 0,49995 | 0,49995 |
| 3,9 | 0,49995 | 0,49995 | 0,49996 | 0,49996 | 0,49996 | 0,49996 | 0,49996 | 0,49996 | 0,49997 | 0,49997 |
| 4,0 | 0,499968 | | | | | | | | | |
| 4,5 | 0,499997 | | | | | | | | | |
| 5,0 | 0,4999997 | | | | | | | | | |

Приложение 3

Таблица значений функции Пуассона: $\mathbf{P(X=m)=P_{n,m}\approx\frac{\lambda^m}{m!}\cdot e^{-\lambda}}$.

| m \ λ | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 | 0,6 | 0,7 | 0,8 | 0,9 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0,9048 | 0,8187 | 0,7408 | 0,6703 | 0,6065 | 0,5488 | 0,4966 | 0,4493 | 0,4066 |
| 1 | 0,0905 | 0,1638 | 0,2222 | 0,2681 | 0,3033 | 0,3293 | 0,3476 | 0,3596 | 0,3659 |
| 2 | 0,0045 | 0,0164 | 0,0333 | 0,0536 | 0,0758 | 0,0988 | 0,1217 | 0,1438 | 0,1647 |
| 3 | 0,0002 | 0,0011 | 0,0033 | 0,0072 | 0,0126 | 0,0198 | 0,0284 | 0,0383 | 0,0494 |
| 4 | - | 0,0001 | 0,0003 | 0,0007 | 0,0016 | 0,0030 | 0,0050 | 0,0077 | 0,0111 |
| 5 | - | - | - | 0,0001 | 0,0002 | 0,0004 | 0,0007 | 0,0012 | 0,0020 |
| 6 | - | - | - | - | - | - | 0,0001 | 0,0002 | 0,0003 |

| m \ λ | 1,0 | 2,0 | 3,0 | 4,0 | 5,0 | 6,0 | 7,0 | 8,0 | 9,0 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0,3679 | 0,1353 | 0,0498 | 0,0183 | 0,0067 | 0,0025 | 0,0009 | 0,0003 | 0,0001 |
| 1 | 0,3679 | 0,2707 | 0,1494 | 0,0733 | 0,0337 | 0,0149 | 0,0064 | 0,0027 | 0,0011 |
| 2 | 0,1839 | 0,2707 | 0,2240 | 0,1465 | 0,0842 | 0,0446 | 0,0223 | 0,0107 | 0,0050 |
| 3 | 0,0613 | 0,1805 | 0,2240 | 0,1954 | 0,1404 | 0,0892 | 0,0521 | 0,0286 | 0,0150 |
| 4 | 0,0153 | 0,0902 | 0,1680 | 0,1954 | 0,1755 | 0,1339 | 0,0912 | 0,0572 | 0,0337 |
| 5 | 0,0031 | 0,0361 | 0,1008 | 0,1563 | 0,1755 | 0,1606 | 0,1277 | 0,0916 | 0,0607 |
| 6 | 0,0005 | 0,0120 | 0,0504 | 0,1042 | 0,1462 | 0,1606 | 0,1490 | 0,1221 | 0,0911 |
| 7 | 0,0001 | 0,0034 | 0,0216 | 0,0595 | 0,1045 | 0,1377 | 0,1490 | 0,1396 | 0,1171 |
| 8 | - | 0,0009 | 0,0081 | 0,0298 | 0,0655 | 0,1033 | 0,1304 | 0,1396 | 0,1318 |
| 9 | - | 0,0002 | 0,0027 | 0,0132 | 0,0363 | 0,0689 | 0,1014 | 0,1241 | 0,1318 |
| 10 | - | - | 0,0008 | 0,0053 | 0,0181 | 0,0413 | 0,0710 | 0,0993 | 0,1186 |
| 11 | - | - | 0,0002 | 0,0019 | 0,0082 | 0,0225 | 0,0452 | 0,0722 | 0,0970 |
| 12 | - | - | 0,0001 | 0,0006 | 0,0034 | 0,0113 | 0,0264 | 0,0481 | 0,0728 |
| 13 | - | - | - | 0,0002 | 0,0013 | 0,0052 | 0,0142 | 0,0296 | 0,0504 |
| 14 | - | - | - | 0,0001 | 0,0005 | 0,0022 | 0,0071 | 0,0169 | 0,0324 |
| 15 | - | - | - | - | 0,0002 | 0,0009 | 0,0033 | 0,0090 | 0,0194 |
| 16 | - | - | - | - | - | 0,0003 | 0,0015 | 0,0045 | 0,0109 |
| 17 | - | - | - | - | - | 0,0001 | 0,0006 | 0,0021 | 0,0058 |
| 18 | - | - | - | - | - | - | 0,0002 | 0,0009 | 0,0029 |
| 19 | - | - | - | - | - | - | 0,0001 | 0,0004 | 0,0014 |
| 20 | - | - | - | - | - | - | - | 0,0002 | 0,0006 |
| 21 | - | - | - | - | - | - | - | 0,0001 | 0,0003 |
| 22 | - | - | - | - | - | - | - | - | 0,0001 |

Приложение 4

Критические точки распределения $\chi^2$

| Число степеней свободы $k$ | Уровень значимости $\alpha$ | | | | | |
|---|---|---|---|---|---|---|
| | 0,01 | 0,025 | 0,05 | 0,95 | 0,975 | 0,99 |
| 1 | 6,6 | 5,0 | 3,8 | 0,0039 | 0,00098 | 0,00016 |
| 2 | 9,2 | 7,4 | 6,0 | 0,103 | 0,051 | 0,020 |
| 3 | 11,3 | 9,4 | 7,8 | 0,352 | 0,216 | 0,115 |
| 4 | 13,3 | 11,1 | 9,5 | 0,711 | 0,484 | 0,297 |
| 5 | 15,1 | 12,8 | 11,1 | 1,15 | 0,831 | 0,554 |
| 6 | 16,8 | 14,4 | 12,6 | 1,64 | 1,24 | 0,872 |
| 7 | 18,5 | 16,0 | 14,1 | 2,17 | 1,69 | 1,24 |
| 8 | 20,1 | 17,5 | 15,5 | 2,73 | 2,18 | 1,65 |
| 9 | 21,7 | 19,0 | 16,9 | 3,33 | 2,70 | 2,09 |
| 10 | 23,2 | 20,5 | 18,3 | 3,94 | 3,25 | 2,56 |
| 11 | 24,7 | 21,9 | 19,7 | 4,57 | 3,82 | 3,05 |
| 12 | 26,2 | 23,3 | 21,0 | 5,23 | 4,40 | 3,57 |
| 13 | 27,7 | 24,7 | 22,4 | 5,89 | 5,01 | 4,11 |
| 14 | 29,1 | 26,1 | 23,7 | 6,57 | 5,63 | 4,66 |
| 15 | 30,6 | 27,5 | 25,0 | 7,26 | 6,26 | 5,23 |
| 16 | 32,0 | 28,8 | 26,3 | 7,96 | 6,91 | 5,81 |
| 17 | 33,4 | 30,2 | 27,6 | 8,67 | 7,56 | 6,41 |
| 18 | 34,8 | 31,5 | 28,9 | 9,39 | 8,23 | 7,01 |
| 19 | 36,2 | 32,9 | 30,1 | 10,1 | 8,91 | 7,63 |
| 20 | 37,6 | 34,2 | 31,4 | 10,9 | 9,59 | 8,26 |
| 21 | 38,9 | 35,5 | 32,7 | 11,6 | 10,3 | 8,90 |
| 22 | 40,3 | 36,8 | 33,9 | 12,3 | 11,0 | 9,54 |
| 23 | 41,6 | 38,1 | 35,2 | 13,1 | 11,7 | 10,2 |
| 24 | 43,0 | 39,4 | 36,4 | 13,8 | 12,4 | 10,9 |
| 25 | 44,3 | 40,6 | 37,7 | 14,6 | 13,1 | 11,5 |
| 26 | 45,6 | 41,9 | 38,9 | 15,4 | 13,8 | 12,2 |
| 27 | 47,0 | 43,2 | 40,1 | 16,2 | 14,6 | 12,9 |
| 28 | 48,3 | 44,5 | 41,3 | 16,9 | 15,3 | 13,6 |
| 29 | 49,6 | 45,7 | 42,6 | 17,7 | 16,0 | 14,3 |
| 30 | 50,9 | 47,0 | 43,8 | 18,5 | 16,8 | 15,0 |

Приложение 5

## Критические точки распределения Стьюдента

| Число степеней свободы *k* | Уровень значимости α(двусторонняя критическая область) | | | | | |
|---|---|---|---|---|---|---|
| | 0,10 | 0,05 | 0,02 | 0,01 | 0,002 | 0,001 |
| 1 | 6,31 | 12,7 | 31,82 | 63,7 | 318,3 | 637,0 |
| 2 | 2,92 | 4,30 | 6,97 | 9,92 | 22,33 | 31,6 |
| 3 | 2,35 | 3,18 | 4,54 | 5,84 | 10,22 | 12,9 |
| 4 | 2,13 | 2,78 | 3,75 | 4,00 | 7,17 | 8,61 |
| 5 | 2,01 | 2,57 | 3,37 | 4,03 | 5,89 | 6,86 |
| 6 | 1,94 | 2,45 | 3,14 | 3,71 | 5,21 | 5,96 |
| 7 | 1,89 | 2,36 | 3,00 | 3,50 | 4,79 | 5,40 |
| 8 | 1,86 | 2,31 | 2,90 | 3,36 | 4,50 | 5,04 |
| 9 | 1,83 | 2,26 | 2,82 | 3,25 | 4,30 | 4,70 |
| 10 | 1,81 | 2,23 | 2,76 | 3,17 | 4,14 | 4,59 |
| 11 | 1,80 | 2,28 | 2,72 | 3,11 | 4,03 | 4,44 |
| 12 | 1,78 | 2,18 | 2,68 | 3,05 | 3,93 | 4,32 |
| 13 | 1,77 | 2,16 | 2,65 | 3,01 | 3,85 | 4,22 |
| 14 | 1,76 | 2,14 | 2,62 | 2,98 | 3,79 | 4,14 |
| 15 | 1,75 | 2,13 | 2,60 | 2,95 | 3,73 | 4,07 |
| 16 | 1,75 | 2,12 | 2,58 | 2,92 | 3,69 | 4,01 |
| 17 | 1,74 | 2,11 | 2,57 | 2,90 | 3,65 | 3,96 |
| 18 | 1,73 | 2,10 | 2,55 | 2,88 | 3,61 | 3,92 |
| 19 | 1,73 | 2,09 | 2,54 | 2,86 | 3,58 | 3,88 |
| 20 | 1,73 | 2,09 | 2,53 | 2,85 | 3,55 | 3,85 |
| 21 | 1,72 | 2,08 | 2,52 | 2,83 | 3,53 | 3,82 |
| 22 | 1,72 | 2,07 | 2,51 | 2,82 | 3,51 | 3,79 |
| 23 | 1,71 | 2,07 | 2,50 | 2,81 | 3,49 | 3,77 |
| 24 | 1,71 | 2,06 | 2,49 | 2,80 | 3,47 | 3,74 |
| 25 | 1,71 | 2,06 | 2,49 | 2,79 | 3,45 | 3,72 |
| 26 | 1,71 | 2,06 | 2,48 | 2,78 | 3,44 | 3,71 |
| 27 | 1,71 | 2,05 | 2,47 | 2,77 | 3,42 | 3,69 |
| 28 | 1,70 | 2,05 | 2,46 | 2,76 | 3,40 | 3,66 |
| 29 | 1,70 | 2,05 | 2,46 | 2,76 | 3,40 | 3,66 |
| 30 | 1,70 | 2,04 | 2,46 | 2,75 | 3,39 | 3,65 |
| 40 | 1,68 | 2,02 | 2,42 | 2,70 | 3,31 | 3,55 |
| 60 | 1,07 | 2,00 | 2,39 | 2,66 | 3,23 | 3,46 |
| 120 | 1,66 | 1,98 | 2,36 | 2,62 | 3,17 | 3,37 |
| Число степеней свободы k | 0,05 | 0,025 | 0,01 | 0,005 | 0,001 | 0,0005 |
| | Уровень значимости α(односторонняя критическая область) | | | | | |

Приложение 6

Критические точки распределения Фишера-Снедекора
($К_1$ - число степеней свободы большей дисперсии,
$К_2$ - число степеней свободы меньшей дисперсии)

| Уровень значимости $\alpha$ = 0,01 | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $K_1$ / $K_2$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1 | 4052 | 4999 | 5403 | 5625 | 5764 | 5889 | 5928 | 5981 | 6022 | 6056 | 6082 | 6106 |
| 2 | 98,49 | 99,01 | 90,17 | 99,25 | 99,33 | 99,30 | 99,34 | 99,36 | 99,36 | 99,40 | 99,41 | 99,42 |
| 3 | 34,12 | 38,81 | 29,46 | 28,71 | 28,24 | 27,91 | 27,67 | 27,49 | 27,34 | 27,23 | 27,13 | 27,05 |
| 4 | 21,20 | 18,00 | 16,69 | 15,98 | 15,52 | 15,21 | 14,96 | 14,80 | 14,66 | 14,54 | 14,45 | 14,37 |
| 5 | 16,26 | 13,27 | 12,06 | 11,39 | 10,97 | 10,67 | 10,45 | 10,27 | 10,15 | 10,05 | 9,96 | 9,89 |
| 6 | 13,74 | 10,92 | 9,78 | 9,15 | 8,75 | 8,47 | 8,26 | 8,10 | 7,98 | 7,87 | 7,79 | 7,72 |
| 7 | 12,25 | 9,55 | 8,45 | 7,85 | 7,46 | 7,19 | 7,00 | 6,84 | 6,71 | 6,62 | 6,54 | 6,47 |
| 8 | 11,26 | 8,65 | 7,59 | 7,01 | 6,63 | 6,37 | 6,19 | 6,03 | 5,91 | 5,82 | 5,74 | 5,67 |
| 9 | 10,56 | 8,02 | 6,99 | 6,42 | 6,06 | 5,80 | 5,62 | 5,47 | 5,35 | 5,26 | 5,18 | 5,11 |
| 10 | 10,04 | 7,56 | 6,55 | 5,99 | 5,64 | 5,39 | 5,21 | 5,06 | 4,95 | 4,85 | 4,78 | 4,71 |
| 11 | 9,86 | 7,20 | 6,22 | 5,67 | 5,32 | 5,07 | 4,88 | 4,72 | 4,63 | 4,54 | 4,46 | 4,40 |
| 12 | 9,33 | 6,93 | 5,95 | 5,41 | 5,06 | 4,82 | 4,65 | 4,50 | 4,39 | 4,30 | 4,22 | 4,16 |
| 13 | 9,07 | 6,70 | 5,74 | 5,20 | 4,86 | 4,62 | 4,44 | 4,30 | 4,19 | 4,10 | 4,02 | 3,96 |
| 14 | 8,86 | 6,51 | 5,56 | 5,03 | 4,69 | 4,46 | 4,28 | 4,14 | 4,03 | 3,94 | 3,86 | 3,80 |
| 15 | 8,68 | 6,36 | 5,42 | 4,89 | 4,56 | 4,32 | 4,14 | 4,00 | 3,89 | 3,80 | 3,73 | 3,67 |
| 16 | 8,53 | 6,23 | 5,29 | 4,77 | 4,44 | 4,20 | 4,03 | 3,89 | 3,78 | 3,69 | 3,61 | 3,55 |
| 17 | 8,40 | 6,11 | 5,18 | 4,67 | 4,44 | 4,10 | 3,93 | 3,79 | 3,68 | 3,59 | 3,52 | 3,45 |

| Уровень значимости $\alpha$ = 0,05 | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $K_1$ / $K_2$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1 | 161 | 200 | 216 | 225 | 230 | 234 | 237 | 239 | 241 | 242 | 243 | 244 |
| 2 | 18,51 | 19,00 | 19,16 | 19,25 | 19,30 | 19,33 | 19,36 | 19,37 | 19,38 | 19,39 | 19,40 | 19,41 |
| 3 | 10,13 | 9,55 | 9,28 | 9,12 | 9,01 | 8,94 | 8,88 | 8,84 | 8,81 | 8,78 | 8,76 | 8,74 |
| 4 | 7,71 | 6,94 | 6,59 | 6,39 | 6,26 | 6,16 | 6,09 | 6,04 | 6,00 | 5,96 | 5,93 | 5,91 |
| 5 | 6,61 | 5,79 | 5,41 | 5,19 | 5,05 | 4,95 | 4,88 | 4,82 | 4,78 | 4,74 | 4,70 | 4,68 |
| 6 | 5,99 | 5,14 | 4,76 | 4,53 | 4,39 | 4,28 | 4,21 | 4,15 | 4,10 | 4,06 | 4,03 | 4,00 |
| 7 | 5,59 | 4,74 | 4,35 | 4,12 | 3,97 | 3,87 | 3,79 | 3,73 | 3,68 | 3,63 | 3,60 | 3,57 |
| 8 | 5,32 | 4,46 | 4,07 | 3,84 | 3,69 | 3,58 | 3,50 | 3,44 | 3,39 | 3,34 | 3,31 | 3,28 |
| 9 | 5,12 | 4,26 | 3,86 | 3,63 | 3,48 | 3,37 | 3,29 | 3,23 | 3,18 | 3,13 | 3,10 | 3,07 |
| 10 | 4,96 | 4,10 | 3,71 | 3,48 | 3,33 | 3,22 | 3,14 | 3,07 | 3,02 | 2,97 | 2,94 | 2,91 |
| 11 | 4,84 | 3,98 | 3,59 | 3,36 | 3,20 | 3,09 | 3,01 | 2,95 | 2,90 | 2,86 | 2,82 | 2,79 |
| 12 | 4,75 | 3,88 | 3,49 | 3,26 | 3,11 | 3,00 | 2,92 | 2,85 | 2,80 | 2,76 | 2,72 | 2,69 |
| 13 | 4,67 | 3,80 | 3,41 | 3,18 | 3,02 | 2,92 | 2,84 | 2,77 | 2,72 | 2,67 | 2,63 | 2,60 |
| 14 | 4,60 | 3,74 | 3,34 | 3,11 | 2,96 | 2,85 | 2,77 | 2,70 | 2,65 | 2,60 | 2,56 | 2,53 |
| 15 | 4,54 | 3,68 | 3,29 | 3,06 | 2,90 | 2,79 | 2,70 | 2,64 | 2,59 | 2,55 | 2,51 | 2,48 |
| 16 | 4,49 | 3,63 | 3,24 | 3,01 | 2,85 | 2,74 | 2,66 | 2,59 | 2,54 | 2,49 | 2,45 | 2,42 |
| 17 | 4,45 | 3,59 | 3,20 | 2,96 | 2,81 | 2,70 | 2,62 | 2,55 | 2,50 | 2,45 | 2,41 | 2,38 |

Приложение 7

## Параметрические критерии значимости

| Нулевая гипотеза $H_0$ | Дополнитель-ные условия | Критерий проверки $H_0$ (критериальная статистика) | Используемое распределение | Конкур гип. $H_1$ | Критическая область и формулы для нахождения её границ | | Гипотеза $H_0$ не отвергается, если |
|---|---|---|---|---|---|---|---|
| $H_0 : \tilde{x} = a_0$ | $\sigma_г$ известно | $U_{набл} = \frac{\tilde{x} - a_0}{\sigma}\sqrt{n}$ | Нормальный закон Функция Лапласа | $\tilde{x} > a_0$ | ПКО | $\Phi_0(u_{кр}) = \frac{1-2\alpha}{2}$ | $\lvert U_{набл} \rvert < u_{кр}$ |
| | | | | $\tilde{x} < a_0$ | ЛКО | | |
| | | | | $\tilde{x} \neq a_0$ | ДКО | $\Phi_0(u_{кр}) = \frac{1-\alpha}{2}$ | |
| $H_0 : \tilde{x} = a_0$ | $\sigma_г$ неизвестно | $T_{набл} = \frac{\tilde{x} - a_0}{S}\sqrt{n-1}$ | Стьюдента с $k_{св} = n-1$ степенями свободы | $\tilde{x} > a_0$ | ПКО | $t_{правкр}(\alpha;k)$ | $T_{набл} < t_{правкр}$ |
| | | | | $\tilde{x} < a_0$ | ЛКО | $t_{лев.кр} = -t_{правкр}$ | $T_{набл} > -t_{правкр}$ |
| | | | | $\tilde{x} \neq a_0$ | ДКО | $t_{двуст.кр}(\alpha;k)$ | $\lvert T_{набл} \rvert < t_{двуст.кр}$ |
| $H_0 : \bar{x}_1 = \bar{x}_2$ | $n \geq 30$ | $z_{набл} = \frac{\tilde{x}_1 - \tilde{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ | Нормальный закон Функция Лапласа | $x_1 > x_2$ | ПКО | $\Phi_0(z_{кр}) = \frac{1-2\alpha}{2}$ | $z_{набл} < z_{кр}$ |
| | | | | $x_1 < x_2$ | ЛКО | | $z_{набл} > -z_{кр}$ |
| | | | | $x_1 \neq x_2$ | ДКО | $\Phi_0(z_{кр}) = \frac{1-\alpha}{2}$ | $\lvert z_{набл} \rvert < z_{кр}$ |
| $H_0 : \bar{x}_1 = \bar{x}_2$ | $x_1 \to N(a_1;\sigma_1^2)$ $x_2 \to N(a_2;\sigma_2^2)$ $\sigma_{1г}^2 = \sigma_{2г}^2$ | $T_{набл} = \frac{\tilde{x}_1 - \tilde{x}_2}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ | Стьюдента с $k_{св} = n_1 + n_2 - 2$ степенями свободы | $x_1 > x_2$ | ПКО | $t_{правкр}(\alpha;k)$ | $T_{набл} < t_{правкр}$ |
| | | | | $x_1 < x_2$ | ЛКО | $t_{лев.кр} = -t_{правкр}$ | $T_{набл} > -t_{правкр}$ |
| | | | | $x_1 \neq x_2$ | ДКО | $t_{двуст.кр}(\alpha;k)$ | $\lvert T_{набл} \rvert < t_{двуст.кр}$ |
| $H_0 : \sigma_{ген}^2 = \sigma_0^2$ | $x \to N(a;\sigma^2)$ | $\chi_{набл}^2 = \frac{(n-1)S^2}{\sigma_0^2}$ | $\chi^2$ распред. с $k_{св} = n-1$ степенями свободы | $\sigma_{ген}^2 > \sigma_0^2$ | ПКО | $\chi_{кр}^2(\alpha;k)$ | $\chi_{набл}^2 < \chi_{кр}^2$ |
| | | | | $\sigma_{ген}^2 < \sigma_0^2$ | ЛКО | $\chi_{кр}^2(1-\alpha;k)$ | $\chi_{набл}^2 > \chi_{кр}^2$ |
| | | | | $\sigma_{ген}^2 \neq \sigma_0^2$ | ДКО | $\chi_{правкр}^2(\alpha/2;k)$ $\chi_{лев.кр}^2(1-\alpha/2;k)$ | $\chi_{л.кр}^2 < \chi_{набл}^2 < \chi_{пр.кр}^2$ |

| Нулевая гипотеза $H_0$ | Дополнитель-ные условия | Критерий проверки $H_0$ (критериальная статистика) | Используемое распределение | Конкур гип. $H_1$ | Критическая область и формулы для нахождения её границ | | Гипотеза $H_0$ не отвергается, если |
|---|---|---|---|---|---|---|---|
| $H_0:\sigma_1^2=\sigma_2^2$ | $x_1\to N(a_1;\sigma_1^2)$ $x_2\to N(a_2;\sigma_2^2)$ | $F_{набл}=\dfrac{S_1^2(наиб)}{S_2^2(наим)}$ | F распред. с $k_1=n_1-1$ $k_2=n_2-1$ степенями свободы | $\sigma_1^2>\sigma_2^2$ | ПКО | $F_{кр}(\alpha;k_1;k_2)$ | $F_{набл}<F_{кр}$ |
| | | | | $\sigma_1^2\neq\sigma_2^2$ | ДКО | $F_{кр}(\alpha/2;k_1;k_2)$ | $\lvert F_{набл}\rvert<F_{кр}$ |
| $H_0:p=p_0$ | $n\to\infty$ $(n>30)$ $p$ неизвестна | $U_{набл}=\dfrac{m/n-p_0}{\sqrt{\dfrac{p_0q_0}{n}}}$ $q_0=1-p_0$ | Нормальный закон Функция Лапласа | $p>p_0$ | ПКО | $\Phi_0(u_{кр})=\dfrac{1-2\alpha}{2}$ | $U_{набл}<u_{кр}$ |
| | | | | $p<p_0$ | ЛКО | | $U_{набл}>-u_{кр}$ |
| | | | | $p\neq p_0$ | ДКО | $\Phi_0(u_{кр})=\dfrac{1-\alpha}{2}$ | $\lvert U_{набл}\rvert<u_{кр}$ |
| $H_0:p_1=p_2$ | $n\to\infty$ $(n>30)$ | $U_{\text{íàáë}}=\dfrac{\dfrac{m_1}{n_1}-\dfrac{m_2}{n_2}}{\sqrt{\dfrac{\overline{p}\overline{q}}{n}}}$ $\overline{p}=\dfrac{m_1+m_2}{n_1+n_2}$ $\overline{q}=1-\overline{p}$ $n=\dfrac{n_1n_2}{n_1+n_2}$ | Нормальный закон Функция Лапласа | $p>p_0$ | ПКО | $\Phi_0(u_{кр})=\dfrac{1-2\alpha}{2}$ | $U_{набл}<u_{кр}$ |
| | | | | $p<p_0$ | ЛКО | | $U_{набл}>-u_{кр}$ |
| | | | | $p\neq p_0$ | ДКО | $\Phi_0(u_{кр})=\dfrac{1-\alpha}{2}$ | $\lvert U_{набл}\rvert<u_{кр}$ |

Ефременкова Ольга Валентиновна

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ТЕОРИИ НАДЕЖНОСТИ

Учебное пособие для студентов всех форм обучения
направления «Строительство»

Редактор Е.Ф. Изотова

Подписано к печати 12.10.16. Формат 60х84/16.
Усл. печ. л. 4,44. Тираж 50 экз. Заказ 161580. Рег. № 33.

Отпечатано в ИТО Рубцовского индустриального института
658207, Рубцовск, ул. Тракторная, 2/6.