ПРЕДМЕТ ТЕОРИИ ВЕРОЯТНОСТЕЙ.

Теория вероятностей – это математическая наука, изучающая закономерности, которые возникают при взаимодействии большого числа случайных факторов.

Если в обыденных представлениях считается, что случайные события представляют собой нечто крайне редкое, идущее вразрез закономерному, то в теории вероятностей отказываются от этих представлений. Случайные события, как они понимаются в теории вероятностей, обладают рядом характерных особенностей, в частности, все они происходят в массовых явлениях. Под массовыми явлениями понимаются такие, которые имеют место в совокупностях большого числа почти равноправных объектов и определяются именно этим массовым характером явления и лишь в незначительной мере зависят от природы составляющих объектов.

Во всех случаях, когда применяются вероятностные методы исследования, цель их в том, чтобы, минуя слишком сложное (и зачастую практически невозможное) изучение отдельного явления, обусловленного большим количеством факторов, обратиться к законам, управляющими массами случайных явлений. Изучение этих законов позволяет не только осуществить научный прогноз в своеобразной области случайных явлений, но в ряде случаев помогает целенаправленно влиять на ход случайных явлений, контролировать их, ограничивать сферу действия случайности.

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

Характерным для современного этапа развития любой науки является широкое и плодотворное применение вероятностных и статистических методов. Это вполне естественно, так как при углубленном изучении любого круга явлений неизбежно наступает этап, когда требуется не только выявления основных закономерностей, но и анализ основных отклонений от них. В одних науках, в силу специфики предмета и исторических условий, внедрение статистических методов наблюдается раньше, в других – позже. В настоящее время нет почти не одной науки, в которой так или иначе не применялись бы вероятностные и статистические методы.

Математические законы теории вероятностей – отражение реальных статистических законов, объективно существующих в массовых случайных явлениях природы. К изучению этих явлений теория вероятностей применяет математический метод и по своему методу является одним из разделов математики, столь же логически точным и строгим, как и другие математические науки.

Из истории теории вероятностей.

Возникновение теории вероятностей относится к середине XVII века и связано с исследованиями Паскаля (1623 – 1662), Ферма (1601 – 1665) и Гюйгенса (1629 – 1695) в области азартных игр. В переписке Паскаля и Ферма, относящейся к 1654 году постепенно определились такие важные понятия, как вероятность и математическое ожидание.

Непосредственное практическое применение вероятностные методы нашли, прежде всего, в вопросах страхования. Возникавшие в то время в теории вероятностей задачи решались исключительно элементарно – арифметическими и комбинаторными методами.

Серьезные требования со стороны естествознания (теория ошибок наблюдений, задачи теории стрельбы, проблемы статистики народонаселения) привели к необходимости дальнейшего развития теории вероятностей и привлечения более развитого аналитического аппарата. Особенно значительную роль в развитии аналитических методов теории вероятностей сыграли Муавр (1667 – 1754), Лаплас (1747 –1827), Гаусс (1777 – 1855), Пуассон (1781 –1840).

Для всего XVIII и начала XIX века характерны бурное развитие теории вероятностей и повсеместное увлечение ею. Теория вероятностей становится «модной» наукой. Ее начинают применять не только там, где это применение правомерно, но и там, где оно ничем не оправдано. Для этого периода характерны многочисленные попытки применить теорию вероятностей к так называемым «нравственным» наукам. Появились работы, посвященные вопросам судопроизводства, истории, политики, богословия, в которых применялся аппарат теории вероятностей. Для всех псевдонаучных исследований характерен упрощенный подход к общественным явлениям. Подобные попытки обречены на неудачу, их косвенным результатом стало разочарование в теории вероятностей как в науке 20 – 30 годах XIX века в Западной Европе.

Именно в это время в России создается знаменитая Петербургская математическая школа, трудами которой теория вероятностей была поставлена на прочную логическую и математическую основу и сделана надежным, точным и эффективным методом познания. Успех русской науки был подготовлен деятельностью В.Я. Буняковского (1804- 1889). Им был написан первый в Российской Империи курс теории вероятностей, оказавший большое влияние на развитие этой области науки.

Работы П.Л. Чебышева (1821-18889), А.А. Маркова (1856-1922), А.М. Ляпунова (1857-1918), относящиеся к предельным теоремам теории вероятностей и случайным процессам, вывели теорию вероятностей с задворков науки и поставили ее в ряд точных математических наук.

Современное развитие теории вероятностей характеризуется всеобщим подъемом интереса к ней. А также расширением круга ее практических приложений. В России, США, Франции, Швеции, Италии, Японии, Великобритании, Польше, Венгрии и других странах мира имеется немало ученых, обогащающих теорию вероятностей важными результатами.

Украинская школа теории вероятностей, унаследовав традиции Петербургской математической школы, занимает в мировой науке одно из ведущих мест.

Основные понятия теории вероятностей.

Математическая теория вероятностей приобретает наглядный смысл в связи с такими действительными или мыслимыми опытами, как, например, бросание монеты 100 раз, бросание игральных костей, сдача колоды карт, игра в рулетку, наблюдение продолжительности жизни человека, скрещивание двух сортов растений и наблюдение фенотипов потомков. Сюда же относятся такие явления как пол новорожденных, колебание числа телефонных вызовов, наличие случайных шумов в системах связи, результаты выборочного контроля качества продукции, положение частицы при диффузии и т.д.

Любая теория обязательно предполагает некоторую идеализацию. Начнем ее с возможных исходов «опыта» или «наблюдения».

Например, при бросании монеты не обязательно выпадает герб или решетка; монета может куда–либо закатиться или встать на ребро. Тем не менее, мы условимся рассматривать герб и решетку как единственно возможные исходы бросания монеты. Это соглашение упрощает теорию и не сказывается на возможностях ее применения. Идеализация подобного рода проводится постоянно.

Результаты стохастических опытов или наблюдений будем называть **событиями**. Так «опыт», состоящий в подбрасывании игральной кости, может иметь своим результатом «событие», состоящее в том, что выпавшее число очков – четно.

Будем различать **элементарные** (неразложимые) события (их называют элементарными исходами) и составные (или разложимые) события.

Например, сказать, что «сумма очков, выпавших при бросании двух игральных костей, равна шести» все равно, что сказать, что произошло событие «(1,5) или (2,4), или (3,3), или (4,2), или (5,1)» и это перечисление разлагает событие «сумма очков равна шести» на пять элементарных исходов.

Любое событие может быть разложено на элементарные исходы, иначе говоря, *событие есть совокупность элементарных исходов.*

Элементарные исходы, представляющие собой мыслимые результаты опыта или наблюдения и определяют этот (идеализированный) опыт. Заметим, что термин «элементарный исход» остается столь же неопределенным, как и термин «точка» в геометрии.

Совокупность всех элементарных исходов будем называть **пространством элементарных исходов**, а сами элементарные исходы – точками этого пространства. Все события, связанные с данным (идеализированным) опытом, могут быть описаны как совокупность элементарных исходов, поэтому слово «событие» означает то же самое, что и некоторое «множество элементарных исходов».

Достоверным будем называть событие, которое обязательно произойдет в данном опыте.

Невозможное событие – это то, которое в данном опыте произойти не может.

Ясно, что каждое событие обладает той или иной степенью возможности. Чтобы количественно сравнивать между собой события по степени их возможности, очевидно, нужно с каждым событием связать определенное число, которое тем больше, чем более возможно событие. Таким числом является **вероятность события**, о которой поговорим чуть позже.

Операции над событиями.

Пусть задано произвольное, но фиксированное пространство элементарных исходов $\boldsymbol{\Omega}$, такой же символ используют для достоверного события. Будем использовать заглавные латинские буквы для обозначения событий. Невозможное событие будем обозначать символом $\varnothing$ (пустое множество).

Два события называют **тождественными** друг другу ($\boldsymbol{A=B}$) тогда и только тогда, когда эти события состоят из одних и тех же точек.

Событие, состоящее в наступлении хотя бы одного из событий $\boldsymbol{A}$ или $\boldsymbol{B}$, называют **суммой** событий $\boldsymbol{A}$ и $\boldsymbol{B}$ и обозначают $\boldsymbol{A+B}$.

Событие, состоящее в наступлении обоих событий $\boldsymbol{A}$ и $\boldsymbol{B}$, называют **произведением** событий $\boldsymbol{A}$ и $\boldsymbol{B}$ и обозначают $\boldsymbol{A \cdot B}$.

Событие, состоящее в том, что событие $\boldsymbol{A}$ происходит, а событие $\boldsymbol{B}$ не происходит, называют **разностью** событий $\boldsymbol{A}$ и $\boldsymbol{B}$ и обозначают $\boldsymbol{A-B}$.

Событие $\bar{\boldsymbol{A}}$ называют **противоположным** событию $\boldsymbol{A}$, если одновременно выполнены два соотношения: $\boldsymbol{A+\bar{A}=\Omega; \ A \cdot \bar{A}=\varnothing}$. То есть событие $\bar{\boldsymbol{A}}$ содержит все элементарные исходы, не содержащиеся в $\boldsymbol{A}$.

Два события называют **несовместными**, если их совместное появление невозможно, т.е. $\boldsymbol{A \cdot B=\varnothing}$.

События $\boldsymbol{B_1, B_2, \dots, B_n}$ образуют **полную группу событий**, если хотя – бы одно из них непременно произойдет в результате опыта, т.е.

$\boldsymbol{B_1+B_2+\dots+B_n=\Omega}$.

Примеры событий и действий над ними.

Пример 1.1. *Бросание монеты.* Пусть монета подбрасывается 3 раза. **Г** - выпадение герба при одном бросании, **Р** – выпадение решетки. Пространство элементарных исходов состоит из 8 точек: ГГГ, ГГР, ГРГ, РГГ, ГРР, РГР, РРГ, РРР. Событие **А** – «выпало не менее двух гербов» ={ГГГ, ГГР, ГРГ, РГГ}. Событие **В** – «выпала ровно одна решетка» = {ГГР, ГРГ, РГГ}. В этом случае **А+В=А, А·В=В, А-В=**{ГГГ}.

Пример 1.2. *Возраст супругов.* Страховые компании интересуются распределением возрастов супругов. Пусть событие **А**-«мужу свыше 40 лет», событие **В**- «муж старше жены», событие **С**-«жене свыше 40 лет». Тогда **А·В –** «мужу свыше 40 лет и он старше жены», **А-В –**«мужу свыше 40 лет, но он не старше своей жены», **А·С** –«каждому из супругов свыше 40 лет», **А+С** –«хотя бы одному из супругов свыше 40 лет».

Аксиомы теории вероятностей.

В современной математике принято аксиомами называть те предложения, которые принимаются за истинные и в пределах данной теории не доказываются. Все остальные положения этой теории должны выводиться чисто логическим путем из принятых аксиом. Формулировка аксиом представляет собой не начальную стадию развития математической науки, а являются результатом длительного накопления фактов и логического анализа полученных результатов с целью выявления действительно основных первичных фактов. Именно так складывались аксиомы геометрии. Подобный же путь прошла и теория вероятностей, в которой аксиоматическое построение ее основ явилось делом сравнительно недавнего прошлого. Впервые задача аксиоматического построения теории вероятностей была решена в 1917 году С.Н.Бернштейном.

В настоящее время общепринята аксиоматика А.Н.Колмогорова (1903 - 1987), опубликованная в 1933 году, которая связывает теорию вероятностей с теорией множеств и метрической теорией функций. В несколько упрощенном виде система аксиом выглядит следующим образом.

Аксиома 1 *(аксиома существования вероятности). Каждому событию* ***A*** *поставлено в соответствие неотрицательное число* ***P(A),*** *называемое вероятностью.*

Аксиома 2 *(вероятность достоверного события). Вероятность достоверного события равна единице:* $P(\Omega)=1$.

Аксиома 3 *(аксиома сложения). Если события* ***A*** *и* ***B*** *несовместны, то* $P(A+B)=P(A)+P(B)$.

Аксиома 4 *(расширенная аксиома сложения). Если событие* ***A*** *равносильно наступлению хотя бы одного из попарно несовместных событий* $A_1, A_2, \ldots$, *т.е.* $A=A_1+A_2+\ldots$, *то* $P(A)=P(A_1)+P(A_2)+\ldots$.

Вероятность события, определяемая аксиоматически, выражает численную меру степени объективной возможности этого события.

Первые три аксиомы определяют вероятность. Необходимость 4-ой аксиомы связана с тем, что в теории вероятностей постоянно приходится рассматривать события, разделяющиеся на бесконечное число частных случаев. Выведем несколько важных следствий.

Следствие 1. *Вероятность противоположного события* $P(\bar{A})=1-P(A)$.

Доказательство. По определению противоположного события

$A+\bar{A}=\Omega$. Используем вторую аксиому : $P(A+\bar{A}) = 1$. События A и $\bar{A}$ несовместны, следовательно по третьей аксиоме $P(A)+P(\bar{A})=1$. □

Следствие 2. *Вероятность невозможного события равна нулю*: $P(\varnothing)=0$.

Доказательство. Противоположное к невозможному событию – достоверное. Используя вторую аксиому и первое следствие, получаем требуемое.

Следствие 3. *Вероятность любого события* $P(A)\in[0,1]$.

Доказательство следует из второй аксиомы и второго следствия.

Теорема сложения.

Для произвольных событий A *и* B *верно:*

$P(A+B)=P(A)+P(B)-P(A\cdot B)$.

Доказательство. Представим событие $A+B$ в виде суммы несовместных событий: $A+B=A+(B-A\cdot B)$, тогда в силу аксиомы 3 имеем

$P(A+B)=P(A+\{B-A\cdot B\})=P(A)+P(B-A\cdot B)$.

Аналогично $P(B)=P(A\cdot B+\{B-A\cdot B\})=P(A\cdot B)+P(B-A\cdot B)$.

Из последнего равенства следует, что $P(B-A\cdot B)=P(B)+P(A\cdot B)$.

Таким образом $P(A+B)=P(A)+P(B)-P(A\cdot B)$.

Теорема доказана. □

Теорему сложения можно обобщить на случай 3-х событий:

$P(A+B+C)=P(A)+P(B)+P(C)-P(A\cdot\cdot B)-P(A\cdot C)-P(B\cdot C)+P(A\cdot B\cdot C)$.

Классическое определение вероятности.

Классическое определение вероятности основано на равновозможности (равновероятности) элементарных исходов. Например, при бросании игральной кости, которая имеет точную форму куба и изготовлена из однородного материала, равновероятными элементарными исходами будет выпадение какого – либо определенного числа очков (от 1 до 6), обозначенного на гранях этого куба, поскольку в силу наличия симметрии ни одна из граней не имеет объективного преимущества перед другими.

В общем случае рассмотрим полную группу, состоящую из конечного числа элементарных, равновозможных, несовместных исходов некоторого опыта. Такую группу называют группой возможных результатов испытания. Те из возможных результатов испытания, на которые подразделяется событие A, называют результатами испытания, благоприятствующими A.

Классическое определение вероятности может быть сформулировано так:

*Вероятность **P(A)** события **A** равняется отношению числа возможных результатов испытания, благоприятствующих **A**, к числу всех возможных результатов испытания.*

$$P\{A\} = \frac{M}{N}$$

Определенная таким образом вероятность удовлетворяет всем аксиомам вероятности Колмогорова.

Полезность пространств с равновероятными элементарными исходами проявляется при изучении азартных игр и в комбинаторном анализе.

Примеры вычисления вероятности события по классической формуле.

Пример 1.3. *Подбрасывание игральной кости 1 раз.* Событие ***A***=«выпавшее число очков – четно». В этом случае $N = 6$ – число граней куба, $M = 3$ – число граней с четными номерами; тогда $P(A)=1/2$.

Пример 1.4. *Вытягивание шара из урны, содержащей 2 белых и 5 черных шаров.* Событие ***A***=«вытянули черный шар». $N = 2+5=7$(общее число шаров в урне), $M= 5$ (число черных шаров), тогда $P(A)= 5/7$.

Вычисление вероятности по классической формуле вызывает в некоторых случаях затруднения, связанные с тем, что при вычислении чисел ***M*** и ***N*** требуется знание комбинаторных формул. Приведем некоторые из них.

Число всевозможных **перестановок** из n различных элементов равно ***n!***(читается «эн факториал») и обозначает произведение всех целых чисел от 1 до n, т.е. $n! = 1 \cdot 2 \cdot \ldots \cdot n$. Например, число способов рассадить 4-х человек на 4-х местах равно $4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24$.

Число всевозможных способов выбрать ***m*** элементов из ***n*** (порядок, в котором выбирались элементы, роли не играет) называют числом **сочетаний** из ***n*** по ***m*** и обозначают C_n^m.

Справедлива формула $C_n^m = \dfrac{n!}{m!(n-m)!}$.

Например, число способов выбрать 3-х дежурных из группы в 20 человек равно

$$C_{20}^3 = \frac{20!}{3! \cdot 17!} = \frac{20 \cdot 19 \cdot 18}{1 \cdot 2 \cdot 3} = 20 \cdot 19 \cdot 3 = 1140.$$

Число всевозможных способов выбрать ***m*** элементов из ***n*** в определенном порядке называют числом **размещений** из ***n*** по ***m*** и обозначают A_n^m.

Справедлива формула $A_n^m = \dfrac{n!}{(n-m)!}$.

Например, число способов выбрать председателя и секретаря собрания, если в нем участвуют 20 человек, равно $A_{20}^{2} = \frac{20!}{18!} = 20 \cdot 19 = 380.$

УСЛОВНЫЕ ВЕРОЯТНОСТИ.

Понятие условной вероятности является основным инструментом теории вероятностей.

Вероятность события $\boldsymbol{A}$ в предположении, что уже произошло событие $\boldsymbol{B}$, называют **условной вероятностью** события $\boldsymbol{A}$ при условии $\boldsymbol{B}$ и обозначают $\boldsymbol{P(A|B)}$.

Например, в урне два белых шара и один черный. Два человека вынимают из урны по одному шару. Рассмотрим события:

$\boldsymbol{A}$ – появление белого шара у первого человека,

$\boldsymbol{B}$ – появление белого шара у второго человека.

Тогда $\boldsymbol{P(A)=2/3}$; *а* $\boldsymbol{P(A|B)=1/2}$.

Условные вероятности вычисляют (по определению) по формуле

$\boldsymbol{P(A|B)=P(A \cdot B)/P(B)}$, если $\boldsymbol{P(B) \neq 0}$.

Условная вероятность обладает всеми свойствами вероятности. Рассмотрение условных вероятностей при одном и том же данном событии $\boldsymbol{B}$ равносильно выбору $\boldsymbol{B}$ в качестве нового пространства элементарных исходов с вероятностями, пропорциональными первоначальным. Коэффициент пропорциональности $\boldsymbol{P(B)}$ необходим для того, чтобы сделать вероятность нового пространства равной единице. Из формул для условных вероятностей $\boldsymbol{P(A|B)}$ и $\boldsymbol{P(B|A)}$ легко получить следующую теорему.

Теорема умножения. *Вероятность произведения двух событий равна произведению вероятности одного из этих событий на условную вероятность другого при условии, что первое произошло:*

$$\boldsymbol{P(A \cdot B)=P(B) \cdot P(A|B)=P(A) \cdot P(B|A)}.$$

Теорема умножения применима и в том случае, когда одно из событий $\boldsymbol{A}$ или $\boldsymbol{B}$ является невозможным, так в этом случае вместе с $\boldsymbol{P(A)=0}$ имеют место равенства $\boldsymbol{P(A|B)=0}$ и $\boldsymbol{P(A \cdot B)=0}$.

Все основные теоремы о вероятностях остаются справедливыми для условных вероятностей, взятых относительно некоторого фиксированного события $\boldsymbol{B}$.

События $\boldsymbol{A}$ и $\boldsymbol{B}$ называют **независимыми**, если

$$\boldsymbol{P(A \cdot B)=P(A) \cdot P(B)}.$$

Если события $\boldsymbol{A}$ и $\boldsymbol{B}$ независимы, то условные вероятности совпадают с безусловными:

$$P(A|B)=P(A),\ P(B|A)=P(B).$$

Понятие независимости событий играет значительную роль в теории вероятностей и ее приложениях. В практических вопросах для определения независимости данных событий редко обращаются к проверке выполнения равенства, данного в определении. Обычно пользуются интуитивными соображениями, основанными на опыте. Так, например, ясно, что выпадение герба на одной монете не изменяет вероятности появления герба на другой монете, если только эти монеты во время бросания не связаны между собой. Точно также рождение мальчика у одной матери не изменяет вероятности появления мальчика у другой матери. Это – независимые события.

События $B_1, B_2, \ldots, B_k$ называют **независимыми в совокупности,** если для любых $1 \le i < j < \ldots < r \le k$ выполнено:

$$P(B_i \cdot B_j \cdot \ldots \cdot B_r) = P(B_i) \cdot P(B_j) \cdot \ldots \cdot P(B_r).$$

Заметим, что для независимости в совокупности нескольких событий не достаточно их попарной независимости.

Примеры использования теорем сложения и умножения вероятностей при решении задач.

Пример 2.1. Прибор состоит из 3-х узлов, каждый из которых может выйти из строя. Отказ хотя бы одного узла приводит к отказу прибора в целом. Надежность (вероятность безотказной работы) для каждого из узлов соответственно равна $p_1=0{,}9;\ p_2=0{,}8;\ p_3=0{,}7$. Найти надежность прибора в целом.

Решение. Рассмотрим события: A_1- безотказная работа 1-го узла; A_2- безотказная работа 2-го узла; A_3- безотказная работа 3-го узла; A- безотказная работа прибора. Ясно, что $A= A_1 \cdot A_2 \cdot A_3$. По теореме умножения для независимых событий

$$P(A)=P(A_1\ A_2 \cdot A_3)=P(A_1) \cdot P(A_2) \cdot P(A_3)=0{,}9 \cdot 0{,}8 \cdot 0.7=0{,}504.$$

Пример 2.2. Студент пришел сдавать зачет, зная из 30 вопросов только одну треть. Какова вероятность сдать зачет, если после отказа отвечать на первый вопрос, преподаватель задает еще один?

Решение. Обозначим события: A-студент сдал зачет, B- студент ответил на первый вопрос преподавателя, C- студент ответил на второй вопрос преподавателя. Очевидно, что $A = B + \overline{B} \cdot C$, т.е. студент сдаст зачет, если он либо ответит на первый вопрос, либо не ответит на первый, но ответит на второй.

По теореме сложения $P(A)=P(B+\overline{B}\cdot C)= P(B)+P(\overline{B}\cdot C)- P(B\cdot\overline{B}\cdot C);$ но $B\cdot\overline{B}\cdot C=\varnothing$ (так как события $\boldsymbol{B}$ и $\overline{B}$ не могут осуществиться одновременно, поэтому $P(A)= P(B)+P(\overline{B}\cdot C)$.

По условию задачи $P(B)=10/30=1/3$. По теореме умножения $P(\overline{B}\cdot C)=P(\overline{B})P(C|\overline{B})$.

Далее, $P(\overline{B})=1-P(B)=2/3;$ $\mathrm{P}(C|\overline{B})=10/29$ (так как осталось 29 вопросов, из которых студент знает 10) .

Таким образом, $\boldsymbol{P(A)}=\frac{1}{3}+\frac{2}{3}\cdot\frac{10}{29}=\frac{49}{87}\approx\boldsymbol{0{,}56}.$

Пример 2.3. Вероятность попадания в цель первым стрелком равна 0,7, а вторым -0,6. Стрелки делают по одному выстрелу по цели одновременно. Определить вероятность того, что цель будет поражена, если они стреляют независимо друг от друга.

***Решение*.** Обозначим события: $\boldsymbol{A_1}$ – цель поражена первым стрелком, $\boldsymbol{A_2}$ – цель поражена вторым стрелком, $\boldsymbol{A}$ – цель поражена. Ясно, что $\boldsymbol{A=A_1+A_2}$. По теореме сложения

$$\boldsymbol{P(A)=P(A_1+A_2)=P(A_1)+P(A_2)-P(A_1\cdot A_2)}.$$

Так как события $\boldsymbol{A_1}$ и $\boldsymbol{A_2}$ – независимы, то $\boldsymbol{P(A_1\cdot A_2)=P(A_1)P(\cdot A_2)}$. По условию задачи $\boldsymbol{P(A_1)=0{,}7,\ P(\cdot A_2)=0{,}6}$. Таким образом, $\boldsymbol{P(A)= P(A_1)+P(A_2)-P(A_1)P(\cdot A_2)=0{,}7+0{,}6-0{,}7\cdot 0{,}6=0{,}88}$.

Формула полной вероятности.

Следствием теорем сложения вероятностей и умножения вероятностей является формула полной вероятности.

Пусть требуется найти вероятность некоторого события $\boldsymbol{A}$, которое может произойти одновременно с одним из событий $\boldsymbol{H_1,H_2,\ldots,H_K}$, образующих полную группу несовместных событий. Эти события называют **гипотезами.**

Так как гипотезы образуют полную группу, то событие $\boldsymbol{A}$ может осуществиться только в комбинации с какими – либо из этих гипотез:

$$\boldsymbol{A=A\cdot H_1+A\cdot H_2+\ldots+A\cdot H_K}\ .$$

Так как гипотезы $\boldsymbol{H_1,\ H_2,\ \ldots,H_K}$ несовместны, то и события $\boldsymbol{A\cdot H_1,\ A\cdot H_2,\ldots,A\cdot H_K}$ также несовместны, используя теорему сложения, получим:

$$P(A)=P(A\cdot H_1)+P(A\cdot H_2)+\ldots+P(A\cdot H_K).$$

Применяя к событиям $A\cdot H_1, A\cdot H_2,\ldots,A\cdot H_K$ теорему умножения, получим

$$P(A)=P(H_1)P(A|H_1)+P(H_2)P(A|H_2)+\ldots+P(H_K)P(A|H_K),$$

или

$$P(A)=\sum_{j=1}^{N} P(H_J)P(A|H_J).$$

Последнюю формулу называют формулой полной вероятности.

Пример 2.4. Группа студентов состоит из 3-х отличников , 9-ти хорошистов и 18 студентов, занимающихся слабо. Отличники на экзамене могут получить «5» с вероятностью 0,9 и «4» с вероятностью 0,1; хорошо успевающий студент может получить «5» с вероятностью 0,3, «4» с вероятностью 0,5 и «3» с вероятностью 0,2; слабо успевающий студент может получить «4» с вероятностью 0,2, «3» с вероятностью 0,4 и «2» с вероятностью 0,4. Найти вероятность того, что случайно выбранный студент получит «5» или «4».

Решение. Событие A - случайно выбранный студент получит на экзамене «5» или «4». Гипотезы: H_1- студент успевает отлично, H_2- студент успевает хорошо, H_3- студент успевает слабо.

Вероятности гипотез: $P(H_1)=3:30=0,1; P(H_2)=9:30=0,3; P(H_3)=18:30=0,6.$

Условные вероятности: $P(A|H_1)=1, P(A|H_2)=0,8, P(A|H_3)=0,2.$

По формуле полной вероятности: $P(A)=0,1\cdot 1+0,3\cdot 0,8+0,6\cdot 0,2=0,46.$

Формула Байеса.

Следствием теоремы умножения и формулы полной вероятности является формула Байеса, названная по имени установившего её в 1763 году Т.Бейеса (Thomas Bayes 1702 - 1761).

Пусть имеется полная группа несовместных событий – гипотез $H_1,H_2,\ldots,H_K$. Вероятности этих гипотез до проведения опыта известны и равны $P(H_1), P(H_2),\ldots, P(H_K)$. Эти вероятности называют **априорными** (или вероятностями *a priori* – до опыта). Произведен опыт, в результате которого произошло некоторое событие A. Требуется пересчитать вероятности гипотез в связи с появлением этого события, т.е. вычислить условную вероятность $P(H_J|A)$ для каждой гипотезы. Условные вероятности гипотез после проведения опыта и реализации события А называют **апостериорными** (или вероятностями *a posteriori* – после опыта).

По теореме умножения имеем:

P(A·H_J)= P(A|H_J)·P(H_J)= P(H_J|A)·P(A) для J=1,2,…,K.

Следовательно, $\boldsymbol{P(H_J|A)}=\dfrac{P(H_J)\cdot P(A|H_J)}{P(A)}$ для J=1,2,…,K.

Выражая P(A) с помощью формулы полной вероятности, имеем

$$P(H_J|A)=\frac{P(H_J)\cdot P(A|H_J)}{\sum\limits_{N=1}^{K}P(A|H_N)\cdot P(H_N)} \quad \text{для } J=1,2,\ldots,K.$$

Последняя формула и носит название формулы Байеса.

Значение формулы Байеса состоит в том, что при наступлении нового события, т.е. по мере получения новой информации, мы можем проверять и корректировать выдвинутые до испытания гипотезы. Такой подход, называемый б*айесовским*, даёт возможность корректировать управленческие решения в экономике.

Пример 2.5. На предприятии изготавливаются изделия определенного вида на трех поточных линиях. На первой линии производится 20% изделий от всего объема их производства, на второй – 30%, на третьей – 50%. Каждая из линий характеризуется соответственно следующими процентами брака: 5%, 2%, 3%. Наугад взятое изделие оказалось бракованным, требуется определить вероятность того, что оно изготовлено на первой линии.

Решение. Обозначим $\boldsymbol{H_1, H_2, H_3}$ события, состоящие в том, что наугад взятое изделие произведено на первой, второй или третьей линиях. Согласно условиям задачи $\boldsymbol{P(H_1)=0{,}2,\ P(H_2)=0{,}3,\ P(H_3)=0{,}5.}$ Обозначим через $\boldsymbol{A}$ – событие, состоящее в том, что наугад взятое изделие оказалось бракованным. По условиям задачи $\boldsymbol{P(A|H_1)=0{,}05,\ P(A|H_2)=0{,}02,\ P(A|H_3)=0{,}03.}$

По формуле Байеса имеем

$$\boldsymbol{P(H_1|A)}=\frac{P(H_1)\cdot P(A|H_1)}{\sum\limits_{N=1}^{3}P(A|H_N)\cdot P(H_N)}=\frac{0{,}05\cdot 0{,}2}{0{,}05\cdot 0{,}2+0{,}02\cdot 0{,}3+0{,}03\cdot 0{,}5}=\frac{10}{31}\,.$$

Случайные величины и законы их распределения.

Одним из основных понятий теории вероятностей является понятие случайной величины. Сначала рассмотрим примеры.

Число вызовов, поступивших от абонентов в течение определенного времени на телефонную станцию, является случайным и принимает те или иные значения в зависимости от случайных обстоятельств.

Число отличных оценок у студентов одной группы на экзамене; периметр перпендикулярного сечения ствола дерева; расстояние точки падения диска от точки метания; вес наугад взятого зерна пшеницы; число избирателей, которые могут отдать свои голоса определенному политическому блоку, - примеры случайных величин, относящихся к различным областям жизни.

Несмотря на разнородность конкретного содержания приведенных примеров, все они с точки зрения математики представляют одну и ту же картину. Каждая из этих величин под влиянием случайных обстоятельств, способна принимать различные значения. Заранее указать, какое значение примет эта величина, нельзя, так как оно меняется от зависимости от результата стохастического опыта (эксперимента) случайным образом. В самом общем смысле случайная величина – это некоторая переменная величина, принимающая в зависимости от случая те или иные значения с определенными вероятностями.

Считают, что случайная величина известна, если все её возможные значения, и вероятности, с которыми случайная величина принимает эти значения.

Разнообразие случайных величин велико. Число принимаемых ими значений может быть конечным, счетным или несчетным; значения могут быть расположены дискретно или заполнять интервалы (конечные или бесконечные). Для того чтобы задавать вероятности значений случайных величин, столь различных по своей природе, и притом задавать их одним и тем же способом, в теории вероятностей используют функцию распределения случайной величины.

Случайная величина есть некоторая измеримая функция, заданная на пространстве элементарных исходов. Для каждой случайной величины определена функция распределения.

Пусть X – случайная величина, x – действительное число. Вероятность того, что X примет значение, меньшее, чем x, называется **функцией распределения вероятностей случайной величины:**

$$F_X(x) = P\{X < x\}.$$

Свойства функции распределения.

1) При помощи функции распределения можно вычислить вероятность попадания случайной величины в полуинтервал:

$$P\{X \in [x_1, x_2)\} = F_X(x_2) - F_X(x_1).$$

Действительно, пусть ***A***- событие, состоящее в том, что X примет значение, меньшее, чем x_2; ***B***- событие, состоящее в том, что $X<x_1$, и,

наконец ***C***- событие $\{x_1 \le X < x_2\}$; тогда, очевидно, ***A=B+C***. Так как события ***B*** и ***C*** несовместны, то ***P(A)=P(B)+P(C).***

Но $P(A)=F_X(x_2); P(B)=F_X(x_1); P(C)=P\{x_1 \le X < x_2\}$, поэтому

$P\{x_1 \le X < x_2\}=F_X(x_2)-F_X(x_1)$.

2) Функция распределения $F(x)$ есть неубывающая функция, т.е. при $x_2>x_1$ имеет место $F(x_2)\ge F(x_1)$.

Действительно, так как, определению, вероятность есть неотрицательное число, то из первого свойства следует второе.

3) Предел функции распределения на минус бесконечности равен нулю. Предел функции распределения на плюс бесконечности равен единице.

Действительно, так как неравенство $\{X < +\infty\}$ достоверно, то $P\{X < +\infty\}=1$.

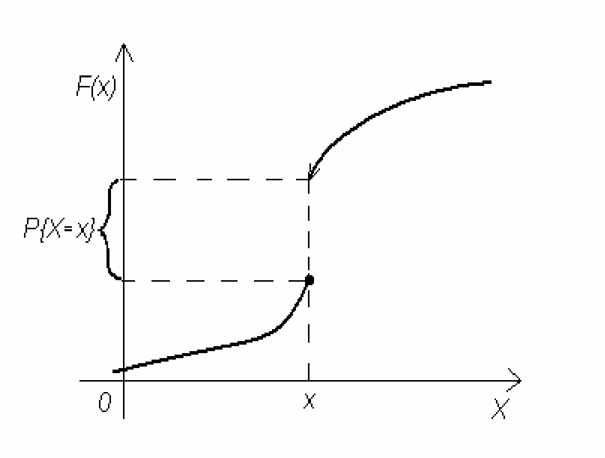
4) Функция распределения непрерывна слева, то есть

$P\{X < x\}=F(x)=F(x-0);$

$P\{X \le x\}=F(x+0).$

Таким образом, $P\{X = x\}=F(x+0)-F(x)$.

Это означает, что вероятность того, что случайная величина примет значение x, равна скачку функции распределения в данной точке.



В дальнейшем будем рассматривать два типа случайных величин – непрерывные и дискретные. Кроме того, существуют, так называемые сингулярные распределения. В общем случае, произвольное распределение может быть представлено в виде смеси непрерывного, дискретного и сингулярного:

$$\boldsymbol{F} = \boldsymbol{p}_1 \cdot \boldsymbol{F}_1 + \boldsymbol{p}_2 \cdot \boldsymbol{F}_2 + \boldsymbol{p}_3 \cdot \boldsymbol{F}_3, \qquad \text{где} \qquad \boldsymbol{p}_1 + \boldsymbol{p}_2 + \boldsymbol{p}_3 = 1.$$

Дискретные случайные величины.

Дискретные случайные величины – это случайные величины, которые могут принимать только конечное или счетное (бесконечное множество, все члены которого можно занумеровать натуральными числами) множество значений.

Например, число появлений герба при трех подбрасываниях монеты (возможные значения 1, 2, 3); число отказавших элементов в приборе, состоящем из 5 элементов (возможные значения 0,1, 2, 3, 4, 5); число выстрелов до первого попадания в цель (возможные значения 1, 2, …, К; где К-число имеющихся патронов) – все это дискретные случайные величины.

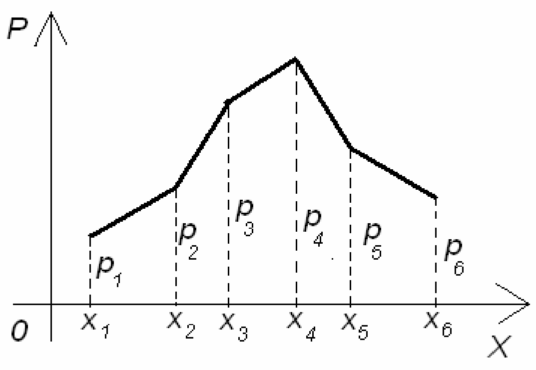
Для полной вероятностной характеристики дискретной случайной величины X, принимающей с положительными вероятностями значения x_1, $x_2, x_3, \ldots$, достаточно знать вероятности, с которыми дискретная случайная величина принимает свои значения

$$\boldsymbol{p_к = P\{X = x_к\}}.$$

Совокупность значений случайной величины $\{x_к\}$ и их вероятностей $\{p_к\}$ называют **рядом распределения.** Ряд распределения часто записывают в виде таблицы, где в верхней строчке перечисляют значения случайной величины, а нижней - вероятности этих значений.

X	x_1	x_2	…	x_k
P	p_1	p_2	…	p_k

Чтобы придать ряду распределения более наглядный вид, часто прибегают к его графическому изображению – **многоугольнику распределения.** Многоугольник распределения строиться следующим образом: по оси абсцисс откладывают возможные значения случайной величины $x_1, x_2, x_3, \ldots$, а по оси ординат – вероятности этих значений $p_1, p_2, p_3, \ldots$. Полученные точки с координатами $(x_1,p_1), (x_2,p_2), (x_3,p_3), \ldots$ соединяют отрезками прямых.



Функция распределения любой дискретной случайной величины определяется через ряд распределения с помощью равенства ***F(x)*** **=** $\sum_{x_R \prec x} p_K$, из которого следует важнейшее свойство ряда распределения: *сумма всех вероятностей, составляющих ряд распределения, равна единице:*

$$\sum_K p_K = \mathbf{1.}$$

Исходя из определения функции распределения дискретной случайной величины, можно записать некоторые характерные *свойства* такой функции:

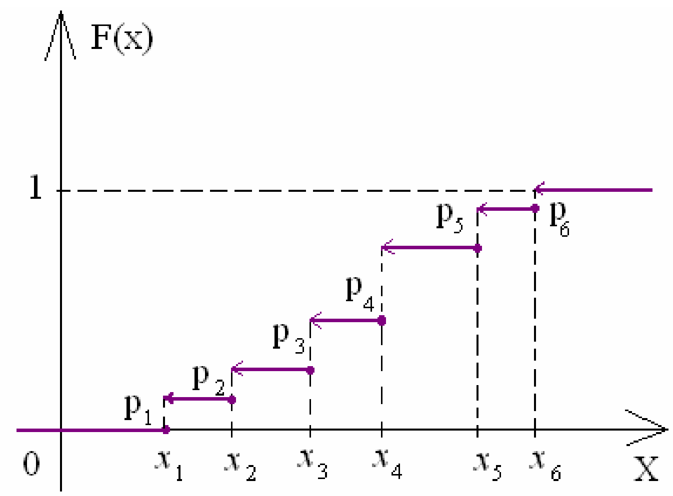
1) функция распределения дискретной случайной величины – разрывная ступенчатая функция;

2) функция распределения дискретной случайной величины возрастает скачками при тех значениях x, которые являются возможными значениями этой случайной величины;

3) величина скачков функции распределения дискретной случайной величины равна

$p_к = P\{X = x_к\}$.

4)сумма всех скачков функции распределения дискретной случайной величины равна единице.



Непрерывные случайные величины.

Непрерывные случайные величины – это величины, возможные значения которых образуют некоторый конечный или бесконечный интервал.

Функция распределения непрерывной случайной величины везде непрерывна. Из последнего положения следует, что вероятность любого отдельного значения непрерывной случайной величины равна нулю.

Остановимся на последнем утверждении подробнее.

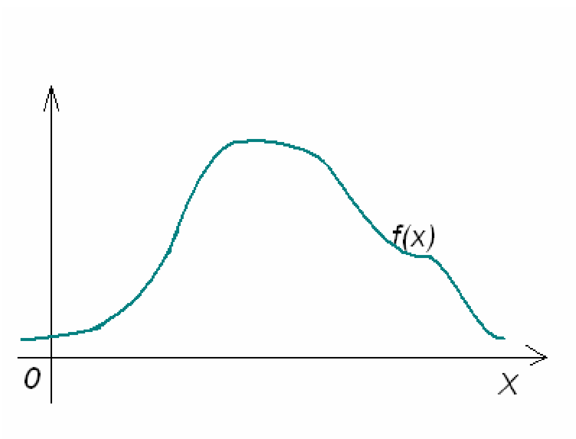
Событие, состоящее в том, что непрерывная случайная величина примет значение *a*, возможно; однако его вероятность равна нулю. Такие события – возможные, но с нулевой вероятностью – появляются только в том случае, когда пространство элементарных исходов не является конечным или счетным. При непрерывном распределении вероятность попадания на сколь угодно малый участок может быть отлична от нуля.

Вычислим вероятность попадания непрерывной случайной величины X в интервал $(x;\ x+\Delta x)$, предположив, что соответствующая функция распределения $F(x)$ непрерывна и дифференцируема: $P\{x < X < x+\Delta x\} = F(x+\Delta x) - F(x)$. Рассмотрим отношение этой вероятности к длине участка, т.е. среднюю вероятность. Устремим Δx к нулю, тогда в пределе получим производную от функции распределения:

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

Функцию $f(x)$ называют **плотностью распределения** случайной величины X.

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



Используя формулу Ньютона – Лейбница, можно записать вероятность попадания случайной величины в заданный интервал ***(а,в)*** через плотность распределения:

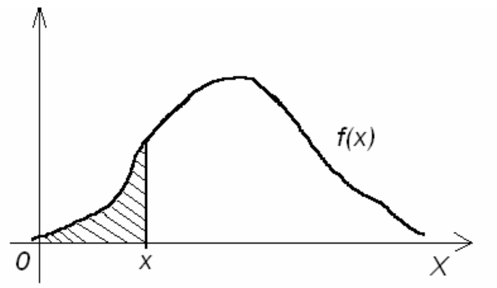
$$P\{a < X < b\} = \int_a^b f(x)dx.$$

Так как вероятность любого отдельного значения непрерывной случайной величины равна нулю, то можно рассматривать здесь открытый интервал, не включая в него левый конец (***в***).

По теореме Барроу можно выразить функцию распределения непрерывной случайной величины через её плотность распределения:

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

Геометрически ***F(x)*** представляет собой площадь под кривой распределения, лежащую левее точки *x*.



Основные свойства плотности распределения:

1) плотность распределения является неотрицательной функцией;
2) интеграл от минус до плюс бесконечности от плотности распределения равен единице:

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

Числовые характеристики случайных величин.

Функция распределения случайной величины, так же, как и плотность распределения, и ряд распределения, являются исчерпывающими характеристиками случайной величины.

Однако во многих вопросах практики нет необходимости характеризовать случайную величину полностью. Зачастую достаточно указать только отдельные числовые параметры, характеризующие основные черты распределения случайной величины: например, какое-то среднее значение, около которого группируются возможные значения случайной величины; какое-либо число, характеризующее степень разбросанности этих значений относительно среднего. Пользуясь такими характеристиками, все существенные сведения относительно случайной величины можно выразить наиболее компактно. Такие характеристики называют числовыми характеристиками случайной величины.

В теории вероятностей числовые характеристиками и операции с ними играют большую роль. Часто удается решить задачу до конца, оставляя в стороне законы распределения и используя только числовые характеристики.

1.Математическое ожидание случайной величины.

Среднее значение случайной величины – некоторое число, являющееся как бы её представителем и заменяющее её при грубо ориентировочных расчетах. Когда говорят «средняя продолжительность жизни в России равна 60 годам» или «средняя заработная плата в городе N равна 3000 руб.», то этим указывают определенную числовую характеристику случайной величины, описывающую её положение на числовой оси, т.е. характеристику положения.

Из характеристики положения в теории вероятностей важнейшую роль играет математическое ожидание (или среднее значение) случайной величины.

Рассмотрим дискретную случайную величину $\boldsymbol{X}$, принимающую возможные значения $\boldsymbol{x_1}$, $\boldsymbol{x_2}$,$\boldsymbol{x_3}$,…, с вероятностями $\boldsymbol{p_1}$, $\boldsymbol{p_2}$, $\boldsymbol{p_3}$,…. **Математическим ожиданием** дискретной случайной величины называется сумма произведений всех возможных значений случайной величины на вероятности этих значений:

$$M(X) = \sum_K x_K p_K \ .$$

Для непрерывной случайной величины $\boldsymbol{X}$ с плотностью $\boldsymbol{f(x)}$ математическое ожидание выражается уже не суммой, а интегралом:

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

Заметим, что математическое ожидание существует не для всех случайных величин, так как сумма и интеграл в определении должны сходиться абсолютно.

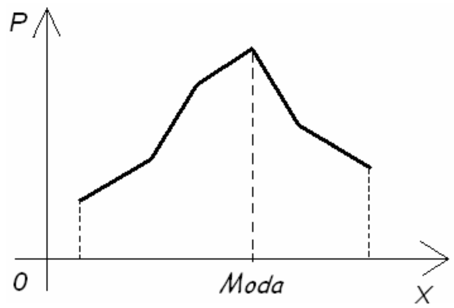
Простейшие свойства математического ожидания:

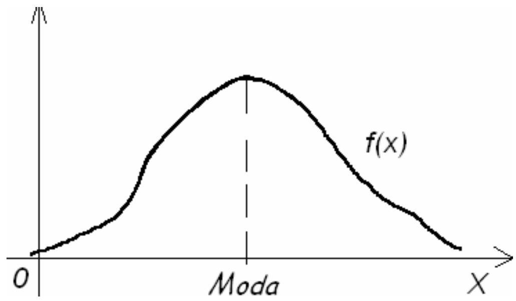
1) математическое ожидание постоянной равно этой постоянной: $\boldsymbol{M(C)=C}$;

2) постоянную величину можно выносить за знак математического ожидания, т.е. $\boldsymbol{M(C \cdot X)=C \cdot X}$.

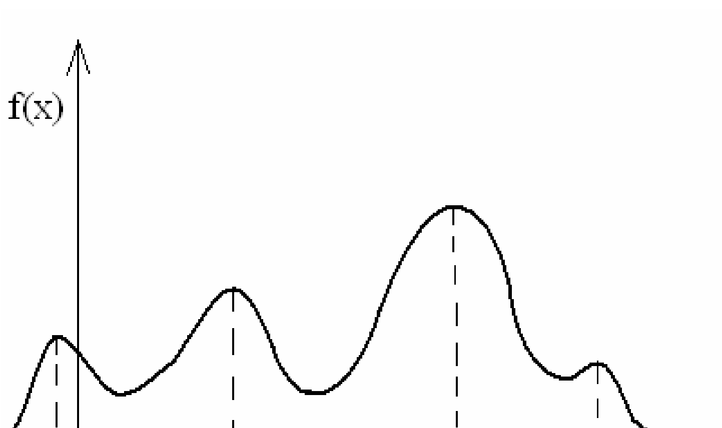
2.Медиана и мода случайной величины.

Кроме важнейших из характеристик положения – математического ожидания – практике применяют и другие характеристики положения, в частности, мода и медиана случайной величины. **Модой дискретной случайной величины** называется её наиболее вероятное значение. Для непрерывной величины модой является то значение, в котором плотность вероятности максимальна.





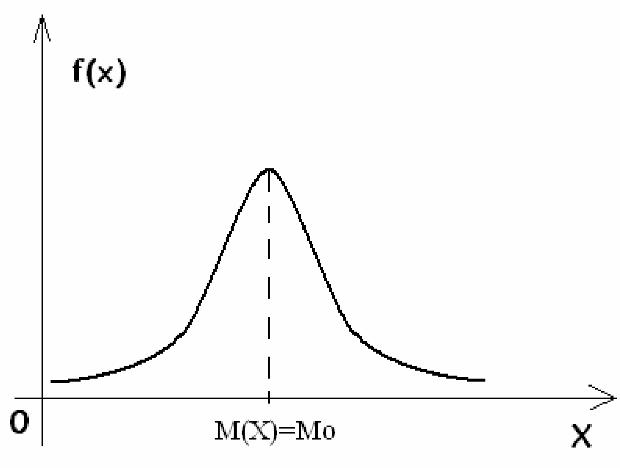
Если многоугольник распределения или кривая распределения имеют более одного максимума, то распределение называют <u>*полимодальным.*</u>



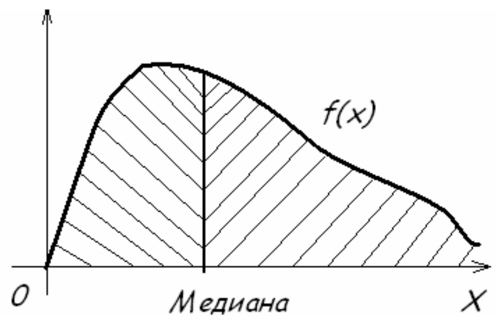
Если же многоугольник распределения или кривая распределения имеют ровно один максимум, то распределение называют *унимодальным*.

В общем случае математическое ожидание и мода не совпадают. В частных случаях, когда распределение является симметричным, существуют мода и математическое ожидание, то они совпадают друг с другом и с центром симметрии распределения.

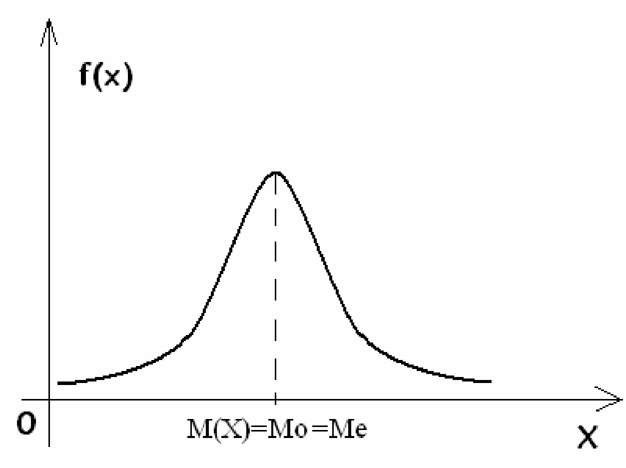
На рисунке – кривая симметричного унимодального распределения.



Для непрерывных случайных величин применяют ещё одну характеристику положения - медиану случайной величины. **Медиана** – это абсцисса точки, в которой площадь под кривой распределения делиться пополам, т.е. одинаково вероятно, окажется ли случайная величина меньше или больше медианы.



В случае симметричного унимодального распределения медиана совпадает с математическим ожиданием и модой.



3. Дисперсия и среднеквадратическое отклонение случайной величины.

Дисперсия и среднеквадратическое отклонение относятся к характеристикам вариации. Характеристики вариации уточняют представление о распределении случайной величины, давая сведение о степени рассеивания случайной величины относительно центра группирования.

Математическое ожидание и дисперсия случайной величины являются её основными характеристиками.

Дисперсия случайной величины – это характеристика рассеивания значений случайной величины около её математического ожидания, определяемая по формуле:

$$D(X)=M(X-M(X))^2.$$

Т.е. дисперсия равна математическому ожиданию квадрата отклонения случайной величины от своего математического ожидания.

Дисперсию часто вычисляют по формуле

$$D(X)=M(X^2)-(M(X))^2,$$

которая следует из определения дисперсии и свойств математического ожидания:

$$D(X)=M(X-M(X))^2 = M(X^2-2X\cdot M(X)+(M(X))^2) =$$

$$=M(X^2)- 2M(X)\cdot M(X)+(M(X))^2 = M(X^2)-M(X))^2.$$

Если случайная величина X – дискретна и известен её ряд распределения, то

$$D(X) = \sum_K x_K^2 \cdot p_K - (\sum_K x_K \cdot p_K)^2;$$

Если же случайная величина X – непрерывна и задана её плотность, то

$$D(X) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx - \left(\int_{-\infty}^{\infty} x \cdot f(x)dx \right)^2.$$

Простейшие свойства дисперсии:

1) дисперсия любой случайной величины неотрицательна: $D(X)\geq 0$ (это сразу следует из определения);

2) дисперсия постоянной равна нулю: $D(C)=M(C-M(C))^2=M(C-C)^2= M(0)=0$;

3) дисперсия произведения случайной величины не постоянную равна произведению дисперсии случайной величины на квадрат постоянной: $D(C\cdot X)=C^2\cdot D(X)$;

4) дисперсия случайной величины не изменится, если к случайной величине прибавить постоянную:

$D(X+C)=D(X)$.

Дисперсия случайной величины имеет размерность квадрата случайной величины, в то время как математическое ожидание имеет размерность самой случайной величины. Для наглядной характеристики рассеивания удобнее пользоваться величиной, размерность которой совпадает с размерностью случайной величины. Для этого из дисперсии извлекают квадратный корень. Полученную величину называют **среднеквадратическим отклонением случайной величины.** Среднеквадратическое отклонение обозначают $\sigma(X)$.

В математической модели случайна величина описывает те или иные параметры изучаемого явления. Числовые значения параметров зависят от выбора масштаба его измерения (например, рубли, тысячи рублей, миллионы рублей). При этом числовые характеристики случайной величины зависят от выбора масштаба измерения исходного параметра.

Для изучения свойств случайных величин, не зависящих от выбора масштаба измерения и положения центра группирования, исходную случайную величину приводят к некоторому стандартному, нормированному виду.

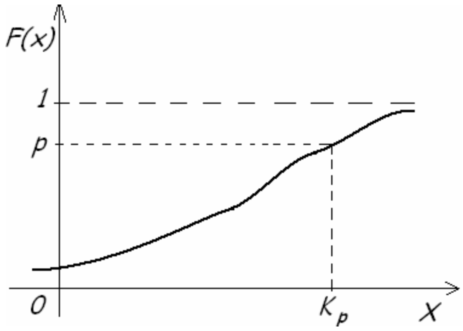
Если $M(X)=0$ и $D(X)=1$, то случайную величину X называют **нормированной.** Для того, чтобы отнормировать случайную величину, из неё надо вычесть математическое ожидание и поделить на среднеквадратическое отклонение: $X^* = \dfrac{X - M(X)}{\sigma(X)}$.

Нормируя случайную величину, мы как бы меняем начало отсчета и масштаб измерения исходного параметра. При этом случайная величина X^* является уже безразмерной и не зависит от выбора масштаба измерения.

4. Квантиль распределения случайной величины.

Название квантиль произошло от латинского *quantum* – сколько.

Квантиль распределения K_p случайной величины уровня p для непрерывных распределений определяется как решение уравнения $F(K_p)=p$, где F- функция распределения случайной величины.



На рисунке – график функции некоторого распределения. По оси ординат откладываем значение уровня p, проводим прямую $y=p$, абсцисса точки пересечения кривой и прямой и есть квантиль уровня p.

Квантиль уровня ½ - это медиана распределения.

Если положить p=0,95, то получим квантиль уровня 0,95.

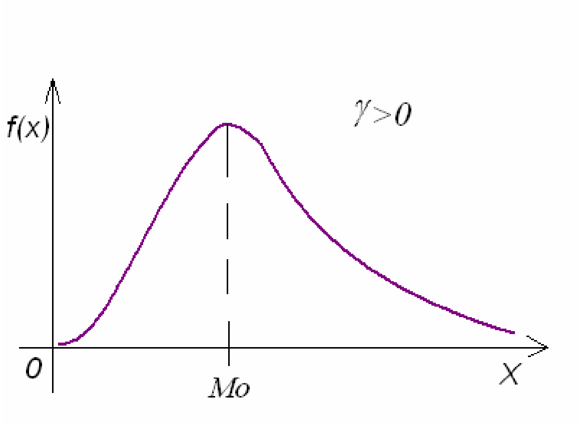
Квантили различных уровней находят по таблицам. Для применения в математической статистике составлены таблицы квантилей наиболее важных распределений.

5. Асимметрия и эксцесс распределения.

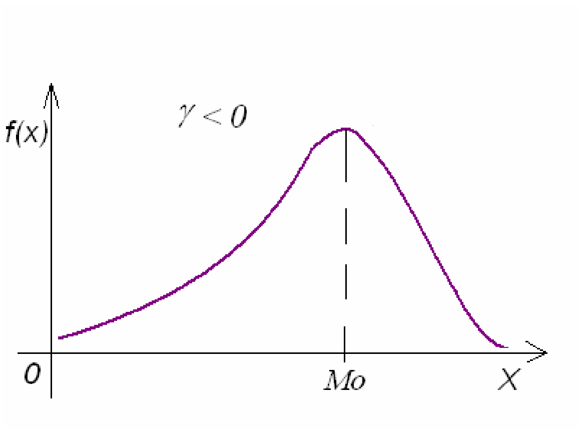
Асимметрия распределения – это качественное свойство кривой распределения, указывающее на отличие от симметричного распределения. Коэффициент асимметрии определяется отношением:

$$\gamma = \frac{M(X - M(X))^3}{(\sigma(X))^3}.$$

Если коэффициент асимметрии положителен, более «длинная» часть кривой плотности распределения лежит правее моды, что видно на рисунке.



Если коэффициент асимметрии отрицателен, более «длинная» часть кривой плотности распределения лежит левее моды, что также видно на рисунке.



Эксцесс – числовая характеристика островершинности графика плотности вероятности унимодального распределения. Коэффициент эксцесса определяется по формуле

$$\varsigma = \frac{M(X - M(X))^4}{(D(X))^2} - 3.$$

Дискретные случайные величины.

Рассмотрим некоторые дискретные распределения, часто встречающиеся на практике.

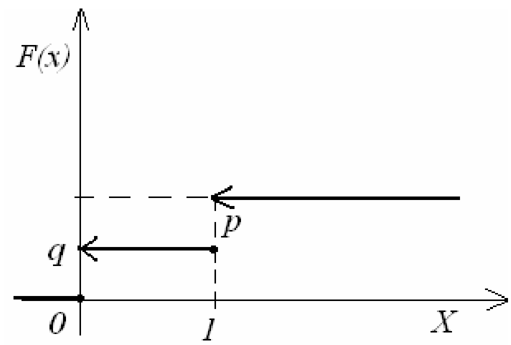
<u>Распределение Бернулли.</u>

Производится один опыт (или наблюдение), в котором может произойти или не произойти событие ***A***. Вероятность того, что событие ***A*** произойдёт, равна числу ***p***. Один такой опыт, в котором возможны лишь два исхода, называемые «успех» и «неудача», называют **испытанием Бернулли**.

Пусть случайная величина ***X*** является индикатором события ***A*** в данном опыте, т.е. ***X=1***, если ***A*** произошло и ***X=0***, если ***A*** не произошло; тогда

$\boldsymbol{P}\{X=1\}=\boldsymbol{P}\{A\}=\boldsymbol{p};\ \ \boldsymbol{P}\{X=0\}=\boldsymbol{P}\{\bar{A}\}=1-p=\boldsymbol{q};$

И говорят, что случайная величина ***X*** распределена по Бернулли. Такую случайную величину называют также **альтернативной.** Функция распределения такой случайной величины имеет вид:



Пример 4.1. При подбрасывании монеты может выпасть “орел” ($X=1$) или “решка” ($X=0$). Если монета симметрична и однородна, то $\boldsymbol{p}=0{,}5$.

<u>Математическое ожидание</u> случайной величины, имеющей распределение Бернулли:

$M(X)=1\cdot p+0\cdot q=\boldsymbol{p},$

(математическое ожидание альтернативной случайной величины равно вероятности положительного исхода).

<u>Дисперсия</u> такой случайной величины:

$D(X)=M(X^2)-(M(X))^2=1^2\cdot p+0^2\cdot q-p^2=\boldsymbol{p\cdot q},$

(дисперсия альтернативной случайной величины равна произведению вероятностей положительного и отрицательного исходов).

Биномиальное распределение.

Проводится n независимых опытов, в каждом из которых может произойти (с вероятностью p) или не произойти (с вероятностью $1-p=q$) некоторое событие A, т.е. производится n независимых испытаний Бернулли. Повторные независимые испытания Бернулли называют **схемой Бернулли** (в честь швейцарского математика Якоба Бернулли 1654 – 1705, который доказал важную теорему, относящуюся к таким испытаниям).

Рассмотрим случайную величину X, равную числу «успехов» в схеме Бернулли.

Найдем вероятность $p_к$ того, что в n испытаниях Бернулли будет к «успехов», или, что то же самое, что случайная величина X примет значение, равное $к$.

Рассмотрим событие $B_к$, состоящее в том, что $X=к$, т.е. событие A появится в опытах ровно $к$ раз.

Событие $B_к$ может осуществиться разными способами, разложим его на сумму произведений событий, состоящих в появлении или не появлении события A в отдельном опыте. Будем обозначать A_i появление события A в i-ом опыте; $\bar{A}_i$- не появление события A в i-ом опыте.

Каждый вариант появления события $B_к$ (т.е. каждый член суммы разложения события $B_к$) должен состоять из $к$ появлений события A и $n-к$ непоявлений, т.е.

$$B_к=A_1A_2\ldots A_к\bar{A}_{к+1}\ldots\bar{A}_n + \ldots + \bar{A}_1\,\bar{A}_2\ldots\bar{A}_{n-к+1}A_{n-к}\ldots A_n,$$

Причем в каждое произведение событие A_i должно входить $к$ раз, а событие $\bar{A}_j$ должно входить $n-к$ раз.

Число всех комбинаций такого рода равно $C_n^к=\dfrac{n!}{к!(n-к)!}$ - числу способов, каким можно из n опытов выбрать $к$, в которых осуществилось событие A.

Вероятность каждой такой комбинации по теореме умножения для независимых событий, равна $p^к(1-p)^{n-к}$.

Так как эти комбинации между собой несовместны, то по теореме сложения, вероятность события $B_к=\{X=к\}$ равна $p_к=C_n^к p^к(1-p)^{n-к}$.

Последнюю формулу называют **формулой Бернулли**.

Случайная величина X имеет ***биномиальное распределение*** *с параметрами* **n** *и* ***p****, если она принимает целочисленные значения от* ***0*** *до* ***n*** *с вероятностями*

$$p_к=P(X=к)=C_n^к p^к(1-p)^{n-к}.$$

<u>Математическое ожидание</u> случайной величины, имеющей биномиальное распределение с параметрами ***n*** и ***p***:

$$M(X)=\sum_{к=0}^{n} к\cdot C_n^к\cdot p^к\cdot(1-p)^{n-к}=n\cdot p\cdot\sum_{к=1}^{n-1}\frac{(n-1)!}{(к-1)!(n-к)!}p^{к-1}(1-p)^{n-к}=n\cdot p.$$

(последняя сумма равна 1, так как состоит из вероятностей биномиального распределения с параметрами ***n-1*** и ***p***).

Таким образом, математическое ожидание биномиальной случайной величины равно произведению числа испытаний на вероятность положительного исхода.

<u>Дисперсия</u> такой случайной величины:

$$D(X)=M(X^2)-(M(X))^2=n\cdot p\cdot q,$$

т.е. дисперсия биномиальной случайной величины равна произведению числа испытаний на вероятности положительного и отрицательного исходов.

Пример 4.2. Длительной проверкой качества стандартных деталей установлено, что 75% деталей не имеют дефектов. Какова вероятность, что из взятых наудачу 6 деталей ровно 5 не имеют дефектов?

Решение. Из условия задачи следует, что ***X***-число стандартных деталей из 6 взятых – имеет биномиальное распределение с параметрами ***n***=6 и ***p***=0,75. По формуле Бернулли

$$P(X=5)=C_6^5\cdot 0{,}75^5\cdot 0{,}25=0{,}356.$$

Пример 4.3. Всхожесть семян данного сорта растений оценивается вероятностью 0,8. Какова вероятность того, что из 5 посеянных зерен взойдет не менее 4? Найти среднее число взошедших семян.

Решение. 1) Обозначим ***X***- число взошедших семян из 5 посеянных, тогда случайная величина X имеет биномиальное распределение с параметрами ***n***=5 и ***p***=0,8. Поэтому

$$P(X\ge 4)=P(X=4)+P(X=5)=C_5^4\cdot 0{,}8^4\cdot 0{,}2+C_5^5\cdot 0{,}8^5=0{,}73728.$$

2) Среднее число взошедших семян: $M(X)=5\cdot 0{,}8=4$.

<u>Геометрическое распределение.</u>

Рассмотрим последовательность испытаний Бернулли. Пусть случайная величина X – число произведенных испытаний до первого «успеха». Найдем для $\boldsymbol{\kappa}$= 1, 2,…вероятность того, что успех наступит при $\boldsymbol{\kappa}$-ом испытании. Событие $\boldsymbol{B_\kappa=(X=\kappa)}$ можно представить как произведение

$$\boldsymbol{B_\kappa=\bar{A}_1\cdot\ldots\cdot\bar{A}_{\kappa-1}\cdot A_\kappa},$$

По теореме умножения дл независимых событий

$$\boldsymbol{P(B_\kappa)=P(X=\kappa)=p\cdot(1-p)^{\kappa-1}}.$$

Случайная величина $\boldsymbol{X}$ *имеет геометрическое распределение с параметром* $\boldsymbol{p}$*, если она принимает натуральные значения* $\boldsymbol{\kappa}$*= 1, 2,… с вероятностями* $\boldsymbol{p_\kappa= p\cdot(1-p)^{\kappa-1}}$.

<u>Математическое ожидание </u>случайной величины, имеющей геометрическое распределение с параметром $\boldsymbol{p}$:

$$M(\boldsymbol{X})=\sum_{\kappa=1}^{\infty}\kappa\cdot p\cdot(1-p)^{\kappa-1}=\frac{1}{p}.$$

Таким образом , математическое ожидание геометрической случайной величины обратно пропорционально вероятности положительного исхода.

<u>Дисперсия</u> данной случайной величины

$$D(\boldsymbol{X})=\frac{1-p}{p^2}.$$

Пример 4.4. Симметричную монету подбрасывают до первого появления орла. Найти вероятность того, что первый раз орел выпадет при пятом подбрасывании.

Решение. Пусть $\boldsymbol{X}$-число подбрасываний монеты до первого появления орла. В силу симметричности монеты $\boldsymbol{X}$ имеет геометрическое распределение с параметром ½. Тогда

$$\boldsymbol{P}(\boldsymbol{X}=5)=0{,}5\cdot0{,}5^4=1/32.$$

Пример 4.5. Найти среднее значение и среднеквадратическое отклонение числа подбрасываний симметричной игральной кости до первого появления «6».

Решение. Случайная величина $\boldsymbol{X}$ – число подбрасываний игральной кости до первого появления «6» имеет геометрическое распределение с параметром 1/6 (так как кость симметрична, а граней всего 6). Тогда среднее значение

$M(X)=1:(1/6)=6;$ $\sigma(X)=\sqrt{\frac{5/6}{1/6^2}}=\sqrt{30}\approx 5{,}5.$

Распределение Пуассона.

Случайная величина X имеет распределение Пуассона с параметром λ ($\lambda>0$), если эта величина принимает целые неотрицательные значения κ=0, 1, 2,… с вероятностями $p_\kappa=\frac{\lambda^\kappa}{\kappa!}\cdot e^{-\lambda}$. (Это распределение впервые было рассмотрено французским математиком и физиком Симеоном Дени Пуассоном в 1837 г.)

Распределение Пуассона иногда называют законом редких событий, так как вероятности p_κ дают приближенное распределение числа наступлений некоторого маловероятного (редкого) события при большом числе независимых испытаний. В этом случае полагают $\lambda=n\cdot p$, где n- число испытаний Бернулли, p- вероятность осуществления события в одном испытании.

Правомерность использования закона Пуассона вместо биномиального распределения при большом числе испытаний дает следующая теорема.

Теорема Пуассона. Если в схеме Бернулли $n\to\infty$, $p\to 0$, так что $n\cdot p\to\lambda$ (конечному числу), то $C_n^k p^k(1-p)^{n-k}\to\frac{\lambda^k}{k!}e^{-\lambda}$ при любых k=0, 1, 2,…

Без доказательства.

Математическое ожидание случайной величины, имеющей распределение Пуассона с параметром λ:

$$M(X)=\sum_{k=0}^{\infty}k\cdot\frac{\lambda^k}{k!}e^{-\lambda}=\lambda\cdot e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^k}{k!}=\lambda\cdot e^{-\lambda}\cdot e^{\lambda}=\lambda.$$

Дисперсия случайной величины, имеющей распределение Пуассона параметром λ:

$D(X)=\lambda$.

Пример 4.6. Вероятность появления бракованного изделия при массовом производстве равна 0,002. Найти вероятность того, что в партии из 1500 изделий будет не более 3-х бракованных. Найти среднее число бракованных изделий.

Решение. 1) Пусть X-число бракованных изделий в партии из 1500 изделий. Тогда искомая вероятность, это вероятность того, что $X\le 3$. В данной задаче мы имеем схему Бернулли с n=1500 и p=0,002. Для применения теоремы Пуассона положим λ=1500·0,002=3. Тогда искомая вероятность

$$P = e^{-3}(1 + \frac{3}{1!} + \frac{3^2}{2!} + \frac{3^3}{3!}) = \frac{13}{e^3} \approx 0{,}65.$$

2) Среднее число бракованных изделий $M(\boldsymbol{X})=\lambda=3$.

Пример 4.7. Коммутатор учреждения обслуживает 100 абонентов. Вероятность того, что в течение 1 минуты абонент позвонит, равна 0,01. Найти вероятность того, что в течение 1 минуты никто не позвонит.

Решение. Пусть ***X***- число позвонивших на коммутатор в течение 1 минуты. Тогда искомая вероятность – это вероятность того, что ***X***=0. В данной задаче применима схема Бернулли с ***n***=100, ***p***=0,01. Для использования теоремы Пуассона положим λ=100·0,01=1. Тогда искомая вероятность ***P*** = $e^{-1} \approx 0{,}37$.

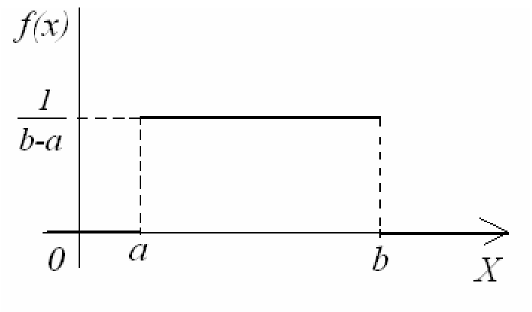
Непрерывные случайные величины.

1.Равномерное распределение.

Понятие равномерного распределения соответствует представлению о выборе точки из определённого отрезка наудачу.

Случайная величина ***X*** **распределена равномерно на отрезке *[a,в]***, если её плотность вероятности равна $f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a,b] \\ 0, & x \notin [a,b] \end{cases}$.

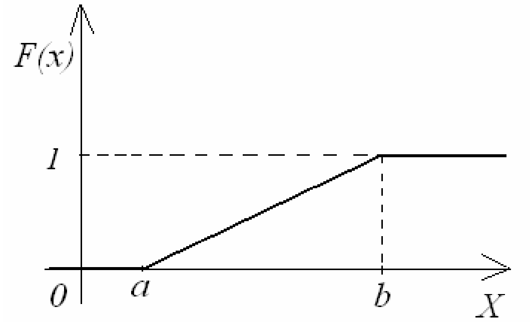
Из-за внешнего вида графиков плотности равномерные распределения называют прямоугольными.



При равномерном распределении отрезок *[а,в]* становится выборочным пространством, в котором вероятности интервалов, лежащих внутри *[а,в]* пропорциональны их длинам.

Функция распределения равномерного закона:

$$F(x) = \begin{cases} 0, \textit{при} \quad x \le a \\ \dfrac{x-a}{b-a}, \textit{при} \quad a < x \le b \\ 1, \textit{при} \quad x > a \end{cases}.$$



Математическое ожидание равномерно распределенной случайной величины

$$M(\boldsymbol{X}) = \int_{-\infty}^{\infty} x \cdot f(x)dx = \int_a^b \frac{x}{b-a}dx = \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} = \frac{a+b}{2}.$$

Дисперсия такой случайной величины

$$D(X) = M(X^2) - [M(X)]^2 = \int_a^b x^2 \frac{dx}{b-a} - \frac{(a+b)^2}{4} =$$

$$= \frac{b^3 - a^3}{3(b-a)} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}.$$

Пример 5.1. Интервал времени между отправлениями поездов в метрополитене равен 3 минутам. Найти вероятность того, что человек, пришедший на станцию метро в случайный момент времени, будет ждать не более 1 минуты.

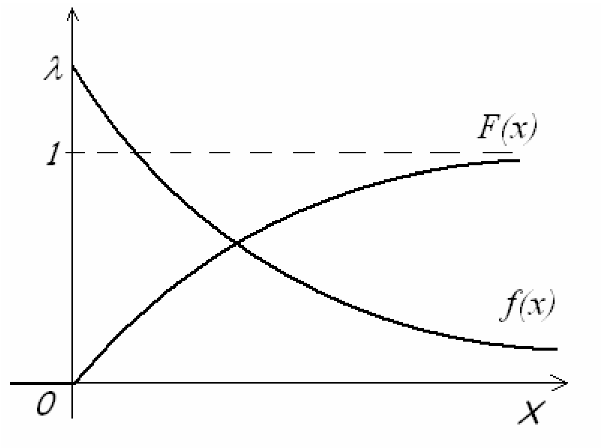
Решение. Время ожидания поезда метрополитена можно считать случайной величиной, имеющей равномерное распределение на отрезке

[0,3]. Вероятность того, что человек будет ждать не более 1 минуты, равна значению функции распределения в точке x=1, т.е. **P**=F(1)=1/3.

2.Показательное распределение.

Непрерывная случайная величина ***X***, принимающая неотрицательные значения, имеет **показательное распределение** с параметром $\boldsymbol{\lambda}$, если её плотность имеет вид: $f(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$,

а функция распределения $F(x) = \begin{cases} 1 - e^{-\lambda \cdot x} & x > 0 \\ 0 & x \leq 0 \end{cases}$.



Показательное распределение – единственное, наделённое «полной потерей памяти». Это свойство называют также свойством *отсутствия последействия*.

Аналитически это свойство записывают следующим образом: для любых чисел x и y

$$\boldsymbol{P\{X > x+y\} = P\{X>x\} \cdot P\{X>y\}}.$$

Считают, что время жизни атома имеет показательное распределение. Свойство отсутствия последействия имеет следующий смысл: *каков бы ни был настоящий возраст, оставшееся время жизни не зависит от прошлого и имеет то же самое распределение, что и само время жизни.*

Использование показательного распределения в математических моделях реальных явлений обычно связывают именно с этим характерным свойством.

Пример 5.2. Время обслуживания клиента на станции технического обслуживания имеет показательное распределение, причем, чем дольше обслуживают в среднем каждого клиента, тем меньше значения параметра λ.

Математическое ожидание случайной величины, имеющей показательное распределение с параметром λ:

$$M(X)=\int_0^{\infty} x\cdot\lambda\cdot e^{-\lambda\cdot x}dx=-x\cdot e^{-\lambda\cdot x}\big|_0^{\infty}-\frac{1}{\lambda}e^{-\lambda\cdot x}\big|_0^{\infty}=\frac{1}{\lambda}.$$

Дисперсия такой случайной величины:

$$D(X)=\int_0^{\infty} x^2\cdot\lambda\cdot e^{-\lambda\cdot x}dx-\frac{1}{\lambda^2}=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}.$$

3. Нормальное распределение.

Непрерывная случайная величина ***X*** имеет нормальное распределение с параметрами m и σ^2, если её плотность распределения $f(x)=\dfrac{1}{\sqrt{2\pi}\cdot\sigma}\exp\left\{-\dfrac{(x-m)^2}{2\sigma^2}\right\}$. Первый множитель в выражении для плотности является нормировочным.

Нормальный закон распределения (также называемый законом Гаусса) играет исключительную роль в теории вероятностей и занимает среди других законов распределения особое положение. Это наиболее часто встречающийся на практике закон распределения. Главная особенность, выделяющая нормальный закон среди других, состоит в том , что он является предельным законом, к которому приближаются другие законы другие законы при весьма часто встречающихся условиях. Теоремы, устанавливающие нормальный закон как предельный, будут рассмотрены в дальнейшем. Нормальный закон может появляться как точное решение некоторых задач. Классические примеры возникновения нормального распределения как точного принадлежат Карлу Фридриху Гауссу (1777-1855) - закон распределения ошибок наблюдения- и Джеймсу Клерку Максвеллу (1833-1879) - закон распределения скоростей молекул.

Во многих задачах, связанных с нормальным распределением, приходится определять вероятность попадания случайной величины ***X***, подчиненной нормальному закону с параметрами m и σ^2, на участок ***[а,в].***

Вероятность попадания случайной величины на заданный интервал выражается через плотность распределения

$$\boldsymbol{P\{X\in[a,b]\}}=\frac{1}{\sqrt{2\pi}}\int_a^b\frac{1}{\sigma}\exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}dx.$$

Последний интеграл не выражается через элементарные функции. Сделаем в нём замену переменных $t=\dfrac{x-m}{\sigma}$, тогда пределы интегрирования поменяются на $\alpha=\dfrac{a-m}{\sigma}$, $\beta=\dfrac{b-m}{\sigma}$. Итак, $\mathbf{P\{X\in[a,b]\}}=\dfrac{1}{\sqrt{2\pi}}\int_\alpha^\beta\exp\left\{-\dfrac{t^2}{2}\right\}dt=\Phi_N(\beta)-\Phi_N(\alpha)$, где $\Phi_N(x)$ – функция распределения *стандартно нормального* распределения (т.е. нормального распределения с параметрами 0 и 1). Эта функция табулирована.

Часто вместо функции стандартно нормального распределения используют также табулированную *функцию Лапласа*: $\Phi(x)=\dfrac{2}{\sqrt{2\pi}}\int_0^x\exp\left\{-\dfrac{t^2}{2}\right\}\,dt$ (для положительных $\boldsymbol{x}$, для отрицательных $\boldsymbol{x}$ функцию Лапласа считают нечетной).

Таким образом, $\mathbf{P\{X\in(a,b)\}}=\dfrac{1}{2}[\,\Phi(\dfrac{b-m}{\sigma})-\Phi(\dfrac{a-m}{\sigma})\,]$.

Математическое ожидание случайной величины, имеющей нормальное распределение с параметрами m и σ^2:

$$M(X)=\int_{-\infty}^{\infty}\frac{x}{\sqrt{2\pi}\cdot\sigma}\exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}dx=$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma\cdot t+m)\cdot\exp\left\{-\frac{t^2}{2}\right\}dt=m.$$

так как $\int_{-\infty}^{\infty}\exp\left\{-\dfrac{t^2}{2}\right\}dt=\sqrt{2\pi}$ и $\int_{-\infty}^{\infty}t\cdot\exp\left\{-\dfrac{t^2}{2}\right\}dt=0$.

Дисперсия такой случайной величины

$$D(X)=\frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{\infty}(x-m)^2\exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}dx=\sigma^2.$$

Часто при решении задач требуется оценить диапазон возможных значений случайной величины. Способ, позволяющий указать интервал практически возможных значений нормально распределенной случайной величины, называют "**правилом 3-х σ".** Рассмотрим вероятность того, что случайная величина отклоняется от своего математического ожидания m не больше, чем на 3σ, т.е.

$$\mathbf{P\{|X-m|<3\sigma\}}=P\{X\in(m-3\sigma;m+3\sigma)\}=\frac{1}{2}[\Phi(\frac{m+3\sigma-m}{\sigma})-\Phi(\frac{m-3\sigma-m}{\sigma})]=$$

$$=\frac{1}{2}[\Phi(3)-\Phi(-3)]=\Phi(3)\approx\mathbf{0{,}997}.$$

Пример 5.3. Ошибка взвешивания – случайная величина, имеющая нормальное распределение с параметрами m=1 и σ=5 (в граммах). Найти интервал практически возможных значений ошибки взвешивания.

Решение. По «правилу 3σ» интервал практически возможных значений равен

(1-3·5;1+3·5), т.е. (-14; 16).

Иногда в экономических расчетах используют «***правило 2-х σ***»:

С вероятностью ***0,95*** *нормально распределенная случайная величина принадлежит интервалу* ***(m-2σ; m+2σ).***

Мода и медиана нормального распределения совпадают со средним. Коэффициенты асимметрии и эксцесса равны нулю. Т.о. ***«крутость» других распределений определяется по отношению к нормальному.***

4.Функции от случайной величины. Логарифмически нормальное распределение.

Функцию от случайной величины иначе называют функцией случайного аргумента.

Функция случайного аргумента $Y=g(X)$ – это случайная величина, функционально зависящая от другой случайной величины, область значений которой совпадает с областью значений функции $y=g(x)$.

Начнем с самой простой функции – линейной.

Теорема. Линейная функция от аргумента, подчинённого нормальному закону $Y=a\cdot X+b$ – это случайная величина, также подчинённая нормальному закону с параметрами $m_Y=a\cdot m_X+b;\ \sigma_Y^2=a^2\cdot\sigma_X^2$.

Доказательство. Рассмотрим функцию распределения случайной величины Y:

$F_Y(x)=P\{Y<x\}=P\{a\cdot X+b<x\}=P\{X<\dfrac{x-b}{a}\}=F_X(\dfrac{x-b}{a})$, если $a>0$.

$F_Y(x)=P\{Y<x\}=P\{a\cdot X+b<x\}=P\{X>\dfrac{x-b}{a}\}=1-F_X(\dfrac{x-b}{a})$, если $a<0$.

Откуда плотность распределения случайной величины Y

$$f_Y(x)=F'_Y(x)==\frac{1}{|a|}\cdot f_X\left(\frac{x-b}{a}\right).$$

записывая явное выражение для плотности, отсюда сразу можно получить доказываемое утверждение.

При построении вероятностных моделей встречаются законы распределения случайных величин, представляющие собой нелинейные функции от нормально распределённых случайных величин. В частности, при решении различных экономических, биологических, геометрических и физических задач используют логарифмически нормальное распределение.

Неотрицательная случайная величина Y имеет **логарифмически нормальное распределение**, если $X=\ln Y$ имеет нормальное распределение. Т.е. $Y=exp\{X\}$.

Плотность логнормального распределения имеет вид

$$f(x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\cdot\sigma\cdot x}\exp\left\{-\dfrac{(\ln x-\ln m)^2}{2\sigma^2}\right\}, & x>0\\ 0, & x\le 0\end{cases}$$

Числовые характеристики логнормального закона можно вычислить, исходя из следующей теоремы.

Теорема. Математическое ожидание функции от случайной величины $Y=g(X)$ вычисляется по формуле: $M(Y)=M(g(X))$.

Таким образом, математическое ожидание логнормального распределения

$M(Y)=exp\{m+\sigma^2/2\}$.

Интересно отметить, что медиана логнормального распределения $Me(Y)=m$, а мода $Mo(X)=a\cdot exp\{-\sigma^2\}$.

Таким образом, если в нормальном распределении параметр m выступает в качестве среднего значения случайной величины, то в логнормальном – в качестве медианы.

Логнормальное распределение используется для описания распределения доходов, банковских вкладов, цен активов, месячной заработной платы, посевных площадей под различные культуры, долговечности вещей в режиме износа и старения.

Пример 5.4. проведенное исследование показало, что вклады населения в данном банке могут быть описаны случайной величиной X, имеющей логнормальное распределение с параметрами m=530, σ^2=0,64.

Найти средний размер вклада, моду и медиану X (пояснив их смысл), долю вкладчиков, размер вклада которых составляет не менее 1000 у.е.

Решение. Используя приведенные выше формулы, имеем:

- средний размер вклада $M(X)$=730 (у.е.);

- мода $Mo(X)$=280 (у.е.) – наиболее часто встречающийся банковский вклад;

- медиана $Me(X)$=530 (у.е.), т.е. половина вкладчиков имеют вклады до 530 у.е., а другая половина – сверх 530 у.е.;

- доля вкладчиков, размер вклада которых составляет не менее 1000 у.е., есть

$P\{X\geq 1000\}=1-P\{X<1000\}=1-F(1000)\approx 0,215$.

Системы случайных величин.

В практических применениях теории вероятностей приходится иметь дело с задачами, в которых результат опыта (стохастического эксперимента) описывается не одной случайной величиной, а двумя или более, образующими систему или случайный вектор.

Например, расходы случайно выбранной семьи зависят от затрат на питание, обувь, одежду, транспорт, удовлетворение духовных потребностей; на урожайность данной культуры влияют погодные условия, применяемые удобрения, характер почвы, качество посевного материала и т.д.

Случайным вектором или **n-мерной случайной величиной** называют упорядоченный набор из n случайных величин $(X_1, X_2,\ldots,X_n)$.

Многомерная случайная величина, случайный вектор, система случайных величин – всё это различные интерпретации одного и того же математического объекта, каждой из которых можно пользоваться для удобства изложения.

Многомерные случайные величины так же, как и одномерные, могут быть дискретными и непрерывными.

Наиболее простой случай – двумерная случайная величина.

Функцией распределения двумерной случайной величины *(X,Y)* называют вероятность совместного выполнения двух неравенств:

F(x,y)=P{X<x;Y<y}.

Используя геометрическую интерпретацию двумерной случайной величины как случайной точки в декартовой системе координат, можно сказать, что функция распределения ***F(x,y)*** есть вероятность попадания случайной точки ***(X,Y)*** в бесконечный квадрант с вершиной в точке ***(x,y),*** лежащий левее и ниже её (границы не входят).

<u>Свойства функции распределения.</u>

1) функция распределения есть неубывающая функция по каждому из своих аргументов;

2) повсюду на $-\infty$ функция распределения равна ***0***;

3) если оба аргумента $=+\infty$, то функция распределения ***=1***;

4) при одном из аргументов, $=+\infty$, функция распределения двумерной случайной величины превращается в функцию распределения компоненты, соответствующей другому аргументу.

Функция распределения существует как для непрерывных, так и для дискретных случайных величин.

Распределение многомерных непрерывных случайных величин обычно характеризуют плотностью распределения.

Плотность распределения двумерной непрерывной случайной величины называют предел отношения вероятности попадания случайной величины в малый прямоугольник к площади этого прямоугольника, когда оба его раздела стремятся к нулю:

$$f(x,y)=\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\frac{P\{(X,Y)\in R_{\Delta}\}}{\Delta x\cdot\Delta y}.$$

Если известна функция распределения, то плотность можно вычислить как вторую смешенную частную производную:

$$f(x,y)=F''_{x,y}(x,y).$$

Геометрически плотность распределения можно изобразить некоторой поверхностью, так называемой **поверхностью распределения**. Эта поверхность аналогична кривой распределения для одномерной случайной величины.

Свойства плотности распределения.

1) плотность распределения – неотрицательная функция;

2) двойной интеграл в бесконечных пределах от плотности двумерного распределения равен 1 (действительно, данный интеграл есть не что иное, как вероятность попадания во всю плоскость ***xOy***);

3) плотности распределения компонент могут быть получены по формулам:

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)dy, \qquad f_Y(y)=\int_{-\infty}^{\infty} f(x,y)dx,$$

Последнее свойство следует из соответствующего свойства для функций распределения и формулы, связывающей функцию распределения функцию распределения случайного вектора и плотность.

Обратимся к дискретным двумерным случайным векторам.

Закон распределения дискретного случайного вектора ***(X,Y)*** – это совокупность всех возможных значений $\{x_i, y_j\}_{i,j}$ случайного вектора и их вероятностей $p_{ij}=P\{X=x_i\,,\ Y=y_j\}$.

Сумма вероятностей для всех возможных значений равна 1:

$$\sum_{i,j} p_{ij}=1.$$

Так как и в непрерывном случае, зная закон распределения случайного вектора, можно найти распределение координат, просуммировав по соответствующему индексу:

$$p_{i\bullet}=P\{X=x_i\}=\sum_j p_{ij}; \qquad p_{\bullet j}=P\{Y=y_j\}=\sum_i p_{ij}.$$

Пример 6.1. Качество продукции характеризуется двумя случайными величинами. Закон распределения случайного вектора ***(X,Y)*** задан таблицей:

	$x_1 = 0$	$x_2 = 0,1$	$x_3 = 0,2$	$x_4 = 0,3$	$p_{i\bullet}$
$y_1 = 5$	**0,2**	**0,1**	**0,05**	**0,05**	**0,4**
$y_2 = 6$	**0**	**0,15**	**0,15**	**0,1**	**0,4**
$y_3 = 7$	**0**	**0**	**0,1**	**0,1**	**0,2**
$p_{\bullet j}$	**0,2**	**0,25**	**0,3**	**0,25**	**1**

На пересечении ***i***-ой строки и ***j***-го столбца стоят вероятности

$$p_{ij}=P\{X=x_i\,,\ Y=y_j\}.$$

Вероятности $p_{i\bullet}=P\{X=x_i\}$ есть сумма вероятностей, стоящих в i-ой строке, а $p_{\bullet j}=P\{Y=y_j\}$ - в j-ом столбце.

Зависимые и независимые случайные величины.

Для того, чтобы полностью описать систему случайных величин, недостаточно знать распределение каждой из величин, входящих в эту систему; нужно знать еще зависимость между величинами, входящими в систему. Эта зависимость характеризуется с помощью условных законов распределения.

Для непрерывных величин большую роль играют **условные плотности распределения** – это плотности распределения одной случайной величины, вычисленные при условии, что другая случайная величина приняла определённое значение:

$$f_X(x \mid y_0) = \frac{f(x, y_0)}{\int\limits_{-\infty}^{+\infty} f(x, y_0)dx};$$

$$f_Y(y \mid x_0) = \frac{f(y, x_0)}{\int\limits_{-\infty}^{+\infty} f(y, x_0)dy}.$$

Условное распределение компонент дискретного случайного вектора – это ряд распределения одной случайной величины, вычисленный при условии, что другая случайная величина приняла определённое значение:

$$p_X(x_i \mid y_j) = P\{X = x_i \mid Y = y_j\} = \frac{p_{ij}}{\sum\limits_{i=1}^{n} p_{ij}};$$

$$p_Y(y_j \mid x_i) = P\{Y = y_j \mid X = x_i\} = \frac{p_{ij}}{\sum\limits_{j=1}^{m} p_{ij}}.$$

Здесь $\{p_{ij}\}_{i=\overline{1,n};\, j=\overline{1,m}}$ - закон распределения дискретного случайного вектора ***(X,Y)***;

n и ***m*** могут быть конечны или бесконечны.

Пример 6.2. Закон распределения дискретного случайного вектора ***(X,Y)*** задан таблицей:

	$x_1 = 0$	$x_2 = 0{,}1$	$x_3 = 0{,}2$	$x_4 = 0{,}3$	$p_{i\bullet}$
$y_1 = 5$	**0,2**	**0,1**	**0,05**	**0,05**	**0,4**
$y_2 = 6$	**0**	**0,15**	**0,15**	**0,1**	**0,4**
$y_3 = 7$	**0**	**0**	**0,1**	**0,1**	**0,2**
$p_{\bullet j}$	**0,2**	**0,25**	**0,3**	**0,25**	**1**

Найдем условное распределение случайной величины $\boldsymbol{X}$ при условии, что случайная величина $\boldsymbol{Y}$ приняла значение $\boldsymbol{y_2=0,1}$. Выбрав значения $\boldsymbol{p_{i2}}$ из столбца таблицы, соответствующего значению $\boldsymbol{y_2=0,1}$, и разделив их на $\boldsymbol{p_{\bullet 2}=0,25}$, получаем условное распределение $\boldsymbol{X}$ при условии, что $\boldsymbol{Y=0,1}$:

x_i	5	6
$p_X(x_i \vert y_2)$	0,4	0,6

При изучении систем случайных величин всегда следует обращать внимание на степень и характер их зависимости. Эта зависимость может быть более или менее ярко выражена, более или менее тесная. В некоторых случаях зависимость между случайными величинами может быть настолько тесной, что зная значение одной величины, можно в точности указать значение другой. В другом крайнем случае зависимость между случайными величинами является настолько слабой, что их можно считать практически независимыми.

Перейдем к точным формулировкам.

Случайные величины $\boldsymbol{X}$ и $\boldsymbol{Y}$ называют **независимыми**, если

$$F(x,y)=F_X(x)\cdot F_Y(y),$$

т.е. события $\{X<x\}$ и $\{Y<y\}$ независимы при любых x и y.

Понятие независимости легко обобщить на случай n величин.

Случайные величины $X_1, X_2, \ldots, X_n$ называют **независимыми в совокупности**, если события $\{X_1<x_1\}, \{X_2<x_2\}, \ldots, \{X_n<x_n\}$ независимы при любых $x_1, x_2, \ldots, x_n$.

Многие результаты, относящиеся к предельным теоремам теории вероятностей, получены в предположении, что исходные случайные величины независимы в совокупности.

Обратимся к зависимым величинам.

Вероятностная зависимость между случайными величинами очень часто встречается на практике. Если случайные величины ***X*** и ***Y*** находятся в вероятностной зависимости, это не означает, что изменением величины одной величины, другая тоже изменится; это лишь означает, что изменением величины ***X*** величина ***Y*** имеет тенденцию также изменятся (например, возрастать или убывать с ростом ***X***). Эта тенденция соблюдается лишь в общих чертах, и в каком-то отдельном случае от неё возможны отступления.

Примерами случайных величин, находящихся в вероятностной зависимости, являются рост и возраст ребенка; затраты и прибыль при производстве определенной продукции; затраты на рекламу и объем продаваемой продукции и т.д.

Степень зависимости между случайными величинами обычно оценивают с помощью числовых характеристик зависимости.

Числовые характеристики зависимости.

Ковариация случайных величин X и Y – математическое ожидание произведения центрированных величин:

$\boldsymbol{Cov\ (X,Y)=M[(X-m_X)\cdot(Y-m_Y)]}$.

Если случайные величины дискретны, то их ковариацию вычисляют по формуле:

$$\operatorname{cov}(X,Y)=\sum_{i,j}(x_i-m_X)\cdot(y_j-m_Y)\cdot p_{ij} \quad .$$

Для непрерывных величин сумму заменяют на интеграл.

Ковариация описывает как рассеивание случайных величин, так и связь между ними.

Теорема. Если случайные величины независимы, то ковариация между ними равна ***0.***

Доказательство проведем для дискретного случая (для непрерывного – аналогично).

Так как случайные величины X и Y - независимы, то

$p_{ij}=P\{X=x_i\ ,\ Y=y_j\}=P\{X=x_i\}\cdot P\{Y=y_j\}=p_i\cdot p_j$.

Следовательно,

$$\operatorname{cov}(X,Y)=\sum_{i,j}(x_i-m_X)\cdot(y_j-m_Y)\cdot p_{ij}=\sum_{i,j}(x_i-m_X)\cdot(y_j-m_Y)\cdot p_i\cdot p_j \quad =$$

$$=[\sum_i (x_i - m_X)\cdot p_i]\cdot[\sum_j (y_j - m_Y)\cdot p_j] =$$

$$=[\sum_i x_i \cdot p_i - \sum_i m_X \cdot p_i]\cdot[\sum_j y_j \cdot p_j - \sum_j m_Y \cdot p_j] = [m_X - m_X]\cdot[m_Y - m_Y] = 0.$$

Здесь было использовано определение математического ожидания дискретной случайной величины и условие нормировки для вероятностей.□

Для характеристики зависимости в чистом виде используют безразмерную величину – корреляцию.

Коэффициентом корреляции случайных величин называют их нормированную ковариацию:

$$\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma(X)\cdot\sigma(Y)}.$$

Очевидно, что коэффициент корреляции обращается в ноль вместе с ковариацией случайных величин, т.е. для независимых случайных величин коэффициент корреляции равен $\boldsymbol{0}$.

Случайные величины, имеющие нулевой коэффициент корреляции, называют **некоррелированными**.

Из независимости случайных величин следует их некоррелированность. Обратное неверно. Условие независимости более жесткое, чем некоррелированности.

Коэффициент корреляции характеризует *линейную* зависимость между случайными величинами.

Теорема. Если случайные величины связаны линейной зависимостью $\boldsymbol{Y=a\cdot X+b}$, то коэффициент корреляции $\boldsymbol{\rho(X,Y)=|1|}$.

*Доказательство***.** Используем свойства математического ожидания и дисперсии:

$\boldsymbol{Cov(X,Y)=M[(X-m_X)\cdot(a\cdot X+b-a\cdot m_X-b)]=a\cdot M[(X-m_X)^2]=a\cdot\sigma^2(X),}$

$\boldsymbol{\sigma(Y)=|a|\cdot\sigma^2(X).}$

Можно показать, что обратное утверждение также верно.

Свойство коэффициента корреляции.

Величина коэффициента корреляции заключена в пределах [-1;1].

Если коэффициент корреляции больше (меньше) нуля, то говорят о положительной (отрицательной) корреляции между случайными величинами, это означает, что при возрастании одной из случайных величин другая имеет тенденцию в среднем возрастать (убывать).

Например, рост и вес человека связаны положительной корреляцией; а время, потраченное на регулировку прибора и количество неисправностей, обнаруженных при работе прибора – отрицательной.

В качестве характеристики зависимости системы n случайных величин используют корреляционную матрицу $\boldsymbol{K}=||cov(X_i,X_j)||_{i,j}$. Из определения ковариации следует, что эта матрица является симметричной.

Функции нескольких случайных аргументов.Свертка.

Рассмотрим важный частный случай: нахождение распределения суммы двух независимых случайных величин.

Пусть случайная величина ***Z=X+Y*** (является суммой двух непрерывных случайных величин). Тогда её функция распределения :

$$G(z)=P\{Z<z)=P\{X+Y<z\}=\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{z-x} f(x,y)dy\right\}dx$$

Теперь рассмотрим плотность распределения случайной величины ***Z***:

$$g(z)=\int_{-\infty}^{+\infty} f_X(x)\cdot f_Y(z-x)dx=\int_{-\infty}^{+\infty} f_X(z-y)\cdot f_Y(y)dy.$$

Для двух последних интегралов, выражающих плотность распределения суммы независимых случайных величин, используют специальное обозначение:

$$\boldsymbol{g=f_X * f_Y}$$

и называют **композицией** (или **свёрткой**) плотностей распределения.

Пример 6.3. Если случайные величины ***X*** и ***Y*** равномерно распределены на ***[a,b]***, то плотность распределения их суммы

$$g(z)=\left\{\begin{matrix} \dfrac{z-2a}{(b-a)^2} & при & 2a<z\le a+b \\ \dfrac{2b-z}{(b-a)^2} & при & a+b<z\le 2b \\ 0 & при & z\le 2a, z>2b \end{matrix}\right\},$$

функцию ***g(z)*** называют плотностью распределения Симпсона.

Пример 6.4. Закон распределения суммы ***n*** независимых случайных величин, имеющих показательное распределение, называют законом Эрланга (***n-1***) порядка. Показательное распределение является законом Эрланга нулевого порядка. Закон Эрланга был получен при моделировании работы телефонных сетей в первых работах по теории массового обслуживания. Плотность распределения Эрланга ***n***-го порядка есть:

$$f(x) = \begin{cases} \dfrac{\lambda}{n!}(\lambda \cdot x)^n e^{-\lambda \cdot x}, & x \geq 0 \\ 0, & x < 0 \end{cases}.$$

Для того, что найти распределение суммы двух независимых дискретных величин, используют аналогичную формулу, но вместо интеграла пишут сумму.

Пример 6.5. Сумма независимых случайных величин, распределенных по закону Пуассона с параметрами λ_1 и λ_2 , также имеет распределение Пуассона с параметром $\lambda_1 + \lambda_2$.

Свойства числовых характеристик суммы случайных величин.

1) ***M(X+Y)=M(X)+M(Y);***

2) ***D(X+Y)=D(X)+D(Y)+2cov(X,Y);***

3) Если случайные величины независимы, то ***D(X+Y)=D(X)+D(Y);***

4) Если случайные величины независимы, то ***D(X-Y)=D(X)+D(Y);***

5) Если случайные величины независимы, то

$$\sigma^2(X+Y)=\sqrt{\sigma^2(X)+\sigma^2(Y)}.$$

Многомерные распределения.

Двумерное нормальное распределение.

В теории вероятностей и её приложениях большую роль играет двумерное нормальное распределение. Плотность двумерной нормальной случайной величины имеет вид:

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-r^2}} \times$$

$$\times \exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_X)^2}{\sigma_X^2} - \frac{2r(x-m_X)(y-m_Y)}{\sigma_X\sigma_Y} + \frac{(y-m_Y)^2}{\sigma_Y^2}\right]\right\}.$$

Это распределение зависит от 5 параметров:

m_X, m_Y – математические ожидания случайных величин;

σ_X, σ_Y – среднеквадратические отклонения;

r – коэффициент корреляции.

Если компоненты двумерной нормальной случайной величины не коррелированны, т.е. ***r=0,*** то

$$f(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y}\exp\left\{-\left[\frac{(x-m_X)^2}{2\sigma_X^2}+\frac{(y-m_Y)^2}{2\sigma_Y^2}\right]\right\}=f_X(x)\cdot f_Y(y).$$

Таким образом, из некоррелированности нормально распределенных случайных величин следует их независимость. Так как для любых случайных величин из независимости следует некоррелированность, то термины «некоррелированные» и «независимые» величины для случая нормального распределения эквивалентны.

При ***r≠0*** компоненты двумерного нормального вектора зависимы. Условные распределения имеют вид:

$$f(x\,|\,y)=\frac{1}{\sigma_X\sqrt{1-r^2}\sqrt{2\pi}}\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{x-m_X}{\sigma_X}+r\frac{y-m_Y}{\sigma_Y}\right]^2\right\};$$

$$f(y\,|\,x)=\frac{1}{\sigma_Y\sqrt{1-r^2}\sqrt{2\pi}}\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{y-m_Y}{\sigma_Y}+r\frac{x-m_X}{\sigma_X}\right]^2\right\}.$$

Из последних формул следует, что условные распределения также являются нормальными с параметрами

$$m_{X|Y}=m_X+r\frac{\sigma_X}{\sigma_Y}(y-m_Y),\quad \sigma_{X|Y}=\sigma_X(1-r);$$

$$m_{Y|X}=m_Y+r\frac{\sigma_Y}{\sigma_X}(x-m_X),\quad \sigma_{Y|X}=\sigma_Y(1-r).$$

Величину $m_{X|Y}$ называют **условным математическим ожиданием** случайной величины ***X*** при условии, что ***Y= y*** (или **регрессией** случайной величины ***X*** по величине ***Y)***.

К числу немногих плоских фигур, вероятность попадания в которые может быть вычислена в явном виде, принадлежит эллипс равных вероятностей:

$$G_C=\left\{(x,y):\left[\frac{(x-m_X)^2}{\sigma_X^2}-\frac{2r(x-m_X)(y-m_Y)}{\sigma_X\sigma_Y}+\frac{(y-m_Y)^2}{\sigma_Y^2}=C^2\right]\right\}.$$

Вероятность попадания точки ***(X,Y)*** внутрь эллипса равной вероятности G_C есть

$$P(C) = 1 - \exp\{-\frac{C^2}{2(1-r^2)}\}.$$

Пример 7.1. Случайные величины ***X*** и ***Y*** – независимы и стандартно нормально распределены. Найти вероятность того, что случайная точка (X,Y) попадёт в кольцо $K=\{(x,y): 4 \le x^2 + y^2 \le 9\}$.

Ясно, что $P\{(X,Y)\in K\}=P\{(X,Y)\in G_3\}- P\{(X,Y)\in G_2\}=P(3)-P(2)=$
$=e^{-2}-e^{-4,5}\approx 0,1242.$

Общий случай n-мерного нормального распределения.

Плотность нормального распределения в пространстве n измерений имеет вид:

$$f(x_1, x_2, ..., x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(K)}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} K_{ij}^{-1}(x_i - m_i)(x_j - m_j)\right\},$$

где m_i – математическое ожидание i – ой компоненты нормального случайного вектора;

K_{ij}^{-1}- элементы матрицы, обратной по отношению к корреляционной матрице ***K***.

Из общего выражения вытекают все формы нормального закона для любого числа измерений и любых видов зависимости между случайными величинами.

Функции от нормально распределенных случайных величин.

1. Линейная функция.

Линейная функция является наиболее простой. Плотность распределения суммы двух независимых случайных величин, подчинённых нормальному закону распределения, можно найти, используя формулу свёртки.

Правило. При композиции нормальных распределений получается снова нормальных распределение, причём математические ожидания и дисперсии суммируются. Т.е., если

$X_1 \sim \mathcal{N}(m_1, \sigma_1^2)$, $X_2 \sim \mathcal{N}(m_2, \sigma_2^2)$, то $X_1 + X_2 \sim \mathcal{N}(m_1 + m_2,;\ \sigma_1^2 + \sigma_2^2)$.

Если слагаемые зависимы, то сумма снова является нормально распределенной случайной величиной со средним $m_1 + m_2$ и дисперсией, равной $\sigma_1^2+\sigma_2^2+2r\sigma_1\sigma_2$, где ***r***- коэффициент корреляции между слагаемыми.

Рассмотрим общий случай. $Y=\sum_{i=1}^{n} a_i \cdot X_i + b$, где $X_i \sim \mathcal{N}(m_i, \sigma_i^2)$. Случайная величина Y также распределена нормально с параметрами $m_Y=\sum_{i=1}^{n} a_i \cdot m_i + b$;

$\sigma_Y^2=\sum_{i=1}^{n} a_i^2 \cdot \sigma_i^2 + 2\sum_{i<j} a_i a_j r_{ij} \sigma_i \sigma_j$.

Следующие распределения играют большую роль в статистике.

2. Распределение хи-квадрат с n степенями свободы $\chi^2(n)$.

Пусть $X_1, X_2,\dots, X_n$ – независимые случайные величины, распределенные по одному и тому же нормальному закону с параметрами m и σ^2. Функция распределения суммы

$$\chi^2(n)=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-m)^2$$

носит название хи-квадрат распределение с n степенями свободы. Плотность такого распределения для $x>0$ имеет вид:

$$f(x)=\frac{x^{n/2-1}\cdot x^{-x/2}}{2^{n/2}\Gamma(n/2)},$$ где $\Gamma(\lambda)$ – гамма-функция.

Распределение $\chi^2(n)$ табулировано.

$M[\chi^2(n)]=n, D[\chi^2(n)]=2n.$

3. Распределение Стьюдента *St(n)*.

[Стьюдент – псевдоним английского статистика Уильяма Сили Госсета, впервые нашедшего этот закон в 1908 году эмпирическим путём.]

Распределение Стьюдента с n степенями свободы определяется как распределение отношения:

$$St=\frac{X}{\sqrt{\chi^2(n)/n}},$$ где $X \sim \mathcal{N}(0, 1)$.

В силу определения это распределение является симметричным.

Важная роль распределения Стьюдента в математической статистике объясняется следующим фактом: если случайные величины $X_0, X_1, X_2,\dots, X_n$ независимы и одинаково нормально распределены с параметрами m и σ^2, то отношение

$$St = \frac{X_0 - m}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2}}$$

имеет распределение Стьюдента с $\boldsymbol{n}$ степенями свободы, которое не зависит от параметров $\boldsymbol{m}$ и $\boldsymbol{\sigma^2}$. Распределение Стьюдента для небольших n табулировано. Для достаточно больших $\boldsymbol{n}$ распределение Стьюдента близко к стандартно нормальному, поэтому можно пользоваться таблицами нормального распределения.

4. Распределение Фишера – Снедекора $\boldsymbol{F(n_1,n_2)}$.

Распределение Фишера – Снедекора имеет отношение двух случайных величин, распределённых по $\boldsymbol{\chi^2}$:

$$F(n_1, n_2) = \frac{\chi^2(n_1)}{n_1} : \frac{\chi^2(n_2)}{n_2}.$$

В математической статистике используют следующий факт: если случайные величины $\boldsymbol{X_1, X_2, \ldots, X_n}$ ***и*** $\boldsymbol{Y_1, Y_2, \ldots, Y_k}$ независимы и одинаково нормально распределены с параметрами $\boldsymbol{m}$ и $\boldsymbol{\sigma^2}$, то случайная величина

$$F(n,k) = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2 \Big/ \frac{1}{k}\sum_{j=1}^{k}(Y_j - m)^2$$

имеет распределение Фишера – Снедекора, не зависящее от параметров $\boldsymbol{m}$ и $\boldsymbol{\sigma^2}$.

Распределение Фишера – Снедекора также табулировано.

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

В практической деятельности большое значение имеют события с вероятностями, близкими к 1 или 0. Поэтому одной из задач теории вероятностей является установление закономерностей, происходящих с вероятностями, близкими к 1; при этом особую роль играют закономерности, возникающие в результате наложения большого числа случайных факторов. Группа теорем, известная под названием «закона больших чисел», является одним из таких предложений теории вероятностей.

Установленное при помощи закона больших чисел свойство случайных величин при определённых условиях вести себя как неслучайные позволяет предсказывать результаты массовых случайных явлений почти с полной определенностью.

Возможности таких предсказаний в области массовых случайных явлениях ещё больше расширяются наличием другой группы теорем, известных под названием «центральная предельная теорема». При суммировании достаточно большого числа случайных величин закон распределения суммы приближается к нормальному при соблюдении некоторых условий. Формы центральной предельной теоремы различаются между собой теми условиями на случайные величины, для которых устанавливается это свойство суммы.

Совокупность различных форм закона больших чисел вместе с различными формами центральной предельной теоремы образуют **предельные теоремы теории вероятностей.**

Для доказательства теорем, относящихся к группе «закона больших чисел», необходимо одно общее неравенство, называемое неравенством Чебышёва. Это неравенство было открыто независимо друг от друга Пафнутием Львовичем Чебышёвым в 1867 г. и Жюлем Бьенемэ в 1853 г.

Неравенство Чебышёва. Каково бы ни было положительное число $\boldsymbol{\varepsilon}$, вероятность того, что случайная величина $\boldsymbol{X}$ отклониться от своего математического ожидания не больше, чем на $\boldsymbol{\varepsilon}$, ограничена сверху величиной $\frac{D(X)}{\varepsilon^2}$:

$$P\{|X - M(X)| \geq \varepsilon\} \leq \frac{D(X)}{\varepsilon^2}.$$

Доказательство проведём для случая, когда случайная величина $\boldsymbol{X}$ – непрерывна. Используя свойство плотности распределения, имеем:

$$P\{|X - M(X)| \geq \varepsilon\} = \int\limits_{|x-M(X)|>\varepsilon} f(x)dx,$$

где ***f(x)*** – плотность распределения случайной величины $\boldsymbol{X}$. Далее:

$$D(X) = \int\limits_{-\infty}^{+\infty} (x - M(X))^2 f(x)dx \geq \int\limits_{|x-M(X)|\geq\varepsilon} (x - M(X))^2 f(x)dx \geq$$

$$\geq \varepsilon^2 \int\limits_{|x-M(X)|\geq\varepsilon} f(x)dx = = \varepsilon^2 P\{|X - M(X)| \geq \varepsilon\}. \quad \square$$

Пример 8.1. *M(X)=m, D(X)=*σ^2. Оценим вероятность того, что случайная величина ***X*** отклониться от своего математического ожидания не больше, чем на 3σ. Для этого в неравенстве Чебышёва надо положить $\varepsilon=3\sigma$:

$$P\{|X-m|\geq 3\sigma\}\leq \frac{D(X)}{9\sigma^2}=\frac{1}{9}.$$

Заметим, что неравенство Чебышёва даёт только верхнюю границу вероятности данного отклонения. Больше этой границы вероятность быть не может ни при каком законе распределения. На практике, в большинстве случаев вероятность того, что величина ***X*** выйдет за пределы участка $m\pm 3\sigma$, значительно меньше, чем 1/9. Например, для нормального закона эта вероятность равна 0,003.

Закон больших чисел.

Одной из наиболее простых, но вместе с тем наиболее важных форм закона больших чисел является теорема, доказанная Чебышевым в 1867 г.

Теорема Чебышева. Если X_1 , X_2 ,…, X_n ,…– последовательность независимых случайных величин, имеющих конечные дисперсии, то

$$P\left\{\left|\frac{X_1+X_2+...+X_n}{n}-\frac{M(X_1)+M(X_2)+...+M(X_n)}{n}\right|>\varepsilon\right\}\xrightarrow[n\to\infty]{}0$$

для любого положительного ε (т.е. среднее арифметическое случайных величин становится близко к постоянной величине – среднему арифметическому их математических ожиданий).

Важным частным случаем теоремы Чебышёва является теорема Бернулли. (Яков Бернулли в конце XVII века доказал теорему, носящую теперь его имя; эта теорема впервые была опубликована в1713 г. после смерти автора.)

Теорема Бернулли. Пусть μ_n – число наступлений события ***A*** в n независимых опытах (испытаниях Бернулли) и ***p*** есть вероятность наступления события ***A*** в каждом из испытаний. Тогда

$$P\left\{\left|\frac{\mu_n}{n}-p\right|>\varepsilon\right\}\xrightarrow[n\to\infty]{}0$$

при любом ε > 0 (т.е. с ростом числа испытаний частота события приближается к его вероятности).

Доказательство. Рассмотрим случайные величины X_k –индикаторы появления события A в k-ом опыте, тогда

$\mu_n = X_1 + X_2 + \ldots + X_k$.

А так как $M(X_k)=p$; $D(X_k)=p(1-p)$, отсюда следует, что теорема Бернулли является простейшим частным случаем теоремы Чебышёва.

Так как на практике часто неизвестные вероятности приходится определять из опыта приближенно, то для проверки соответствия теоремы Бернулли было проведено большое число опытов. При этом рассматривались события, вероятности которых можно считать по тем или иным соображениям известными, относительно которых легко проводить испытания и обеспечить независимость этих испытаний, а также постоянство вероятностей в каждом из испытаний. Все такие опыты дали прекрасное совпадение с теорией. Приведём некоторые из них. Французский естествоиспытатель Жорж Луи Бюффон бросал монету 4040 раз, герб выпал при этом 2048 раз. Частота появления герба в опыте Бюффона равна 0,507. Английский статистик Эгон Шарп Пирсон бросал монету 12000 раз и при этом наблюдал 6019 выпадений герба. Частота выпадения герба в этом опыте Пирсона равна 0,5016. В другой раз он бросил монету 24000 раз, и герб выпал 12012 раз; частота оказалась равна 0,5005. Во всех приведенных опытах частоты лишь немного уклонялись от вероятности 0,5.

Закон больших чисел дает теоретическое основание правилу среднего арифметического, постоянно употребляющемуся в теории измерений. Предположим, что проводится измерение некоторой физической величины a. Повторив измерение n раз в одинаковых условиях, наблюдатель часто получает не вполне совпадающие результаты: $x_1, x_2, \ldots, x_n$. В качестве приближенного значения a принято брать среднее арифметическое из результатов наблюдений:

$$a \approx \frac{x_1 + x_2 + \ldots + x_n}{n}.$$

Если измерения лишены систематической ошибки (т.е. если математические ожидания величин $x_1, x_2, \ldots, x_n$ равны нулю), то согласно закону больших чисел при достаточно больших n с вероятностью, близкой к 1, указанным путем можно получить значение, достаточно близкое к a.

Центральная предельная теорема.

Все формы центральной предельной теоремы посвящены установлению условий, при которых возникает нормальный закон распределения. Так как на практике эти условия часто выполняются, нормальный закон является самым распространённым. Он возникает во всех случаях, когда исследуемая случайная величина может быть представлена в виде суммы достаточно большого числа независимых или слабо зависимых слагаемых. Например, при измерении физических величин возникающие ошибки складываются из многочисленных независимых элементарных ошибок, порождаемых различными причинами. Поэтому ошибки измерения в большинстве случаев распределены по нормальному закону. Некоторые показатели в экономике также представляют собой сумму большого числа слагаемых, каждое из которых имеет малый вклад в суммарный показатель. В этом случае исследуемый показатель имеет приближенно нормальное распределение.

Различные формы центральной предельной теоремы отличаются условиями, накладываемыми на распределение слагаемых, образующих сумму.

Одна из наиболее простых форм центральной предельной теоремы – теорема, относящаяся к случаю одинакового распределения слагаемых.

Центральная предельная теорема для одинаково распределенных слагаемых.

Если X_1 , X_2 ,…, X_n ,…– последовательность независимых одинаково распределённых случайных величин, с математическим ожиданием m и дисперсией σ^2, то при неограниченном увеличении n закон распределения нормированной суммы неограниченно приближается к стандартно нормальному распределению:

$$P\left\{\alpha < \frac{\sum_{k=1}^{n} X_k - n\cdot m}{\sqrt{n}\cdot\sigma} < \beta\right\} \xrightarrow[n\to\infty]{} \int_{\alpha}^{\beta} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} dx .$$

Частным случаем последней теоремы является теорема Муавра – Лапласа, которая в общем виде была доказана французским математиком и физиком Пьером Симоном Лапласом в 1812 году, а один частный случай был известен ещё английскому математику Абрахаму де Муавру (1730 г.).

Если производится n независимых опытов, в каждом из которых событие A появляется с вероятностью p , то справедливо соотношение

$$P\left\{\alpha < \frac{Z_n - n \cdot p}{\sqrt{n \cdot p \cdot q}} < \beta\right\} \approx \int_{\alpha}^{\beta} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} dx,$$

где Z_n – число появлений события A в n опытах; $q=1-p$; n – достаточно велико.

Доказательство. Представим случайную величину Z_n –общее число появлений события A в n опытах в виде суммы

$$Z_n = \sum_{k=1}^{n} X_k,$$

Где X_k – индикатор (число появлений) события A в k опыте.

Случайная величина X_k имеет биномиальное распределение с параметром p, поэтому

$M(X_k) = p; \; D(X_k) = p \cdot q.$

В силу центральной предельной теоремы для одинаково распределенных слагаемых отсюда сразу следует доказываемое утверждение.