Министерство образования и науки Российской Федерации

Федеральное государственное автономное
образовательное учреждение высшего образования
«ЮЖНЫЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Институт математики механики и компьютерных наук
им. академика РАН Воровича И.И.

**А.И.Луценко**

# Теория вероятностей

***учебник***

Ростов-на-Дону
2016

**Оглавление**

*«… большей частью важнейшие жизненные вопросы
являются на самом деле лишь
задачами из теории вероятностей»*
*П.Лаплас*

## Предисловие

Приступая к изучению теории вероятностей, мы должны иметь в виду, что по форме проявления причинных связей **законы природы** и **общества** делятся на два класса *детерминированные* и *статистические.*

На протяжении длительного времени человечество изучало и использовало только детерминистические закономерности. Например, при нагревании металлического стержня увеличение его длины описывается линейной функцией. Объём правильной четырёхугольной пирамиды, длина стороны основания которого равна $\boldsymbol{a}$, равен $\boldsymbol{V} = \frac{1}{3}\boldsymbol{a}^2 \cdot \boldsymbol{h}$. При увеличении длины стороны основания в два раза объём пирамиды увеличится в четыре раза. Путь, пройденный телом за $\boldsymbol{t}$ секунд при движении с постоянной скоростью $\boldsymbol{v}$ $^{м}/_{сек}$, равен $\boldsymbol{S} = \boldsymbol{v} \cdot \boldsymbol{t}$. Если функция $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$ непрерывная и неотрицательная на $[\boldsymbol{a},\boldsymbol{b}]$, то площадь криволинейной трапеции, ограниченной этой функцией, равна $\boldsymbol{S} = \int_{\boldsymbol{a}}^{\boldsymbol{b}} \boldsymbol{f}(\boldsymbol{x})\boldsymbol{dz}$.

Однако в разнообразии жизненных ситуаций есть такие, которые нельзя описать и проанализировать их, используя только детерминистические методы. Мы не можем записать функцию, описывающую изменение относительной частоты выпадений «герба» при многократном подбрасывании монеты. Нельзя предсказать число вызовов, которые в течение дня поступят на станцию скорой помощи. Нельзя точно определить необходимые количества каждого размера обуви, которые надо завезти с базы в магазин, чтобы удовлетворить спросы покупателей и, в то же время, избежать затоваривания. Мы не можем, отправляясь в путешествие на автомобиле, точно определить расход бензина на каждые 100 километров пути. Извлекая из партии готовых изделий несколько штук для проверки их качества, мы не можем сказать: ни сколько хороших изделий будет среди выбранных, ни сколько хороших изделий имеется во всей партии.

Во всех этих примерах мы имеем дело со случайными событиями и явлениями, которые необходимо учитывать в нашей повседневной жизни.

Мы, порой сами того не замечая, всё попадающее в сферу нашего внимания стремимся измерить, всему стараемся дать количественную или качественную оценку.

Согласно детерминистическим законам состояние изучаемого объекта или явления определяется чётко и точно. То есть в детерминистических случаях изучаемому, или просто рассматриваемому объекту мы даём точную оценку его состояния, которую назовём *мерой объекта*.

При рассмотрении статистических закономерностей нельзя дать однозначный ответ относительно настоящего или будущего состояния изучаемого объекта. Например, зная какие исходы возможны при подбрасывании монеты, мы никогда чётко и точно не сможем указать предстоящий результат, то есть - исход подбрасывания.

Но случайные события, которые мы наблюдаем постоянно, происходят без нашего желания и влияют на выбор варианта нашего поведения. А для этого нужна количественная или качественная оценка возможного события, нужно знание *меры объекта*.

Анализ азартных игр как результатов многократных повторений одного и того же опыта позволили в наблюдаемых статистических закономерностях заметить *статистическую устойчивость.* Это позволило определить *меру возможности объекта*, которым может закончиться опыт.

В письме Пьеру Ферма (18 октября 1654г) Блез Паскаль впервые употребляет слово ***вероятность (probabilite)***: «Я буду пользоваться термином *вероятность* для обозначения числа, означающего степень уверенности. Естественно принять степень уверенности в появлении достоверного события равной единице».

То есть, вместо *меры объекта*, которую мы определяем в детерминистических закономерностях, Паскаль предложил при изучении статистических закономерностей определять *меру возможности объекта*, которая будет числом между нулём и единицей.

*Теория вероятностей* – наука, изучающая статистические закономерности случайных событий и явлений.

Развитие естествознания и общественной практики (оценка ошибок измерений, задачи теории стрельбы, различные виды страхования, статистика народонаселения) привели к дальнейшему развитию теории вероятностей путём применения всё более серьёзного аналитического аппарата.

К концу XIX века сложился стандарт требований к логической строгости при построении математической теории, основанный на теоретико-множественной концепции.

Любая математическая теория имеет дело с одним или несколькими множествами объектов, связанных между собой некоторыми соотношениями. Все формальные свойства этих объектов и отношений, необходимые для развития теории, фиксируются в виде аксиом, не затрагивающих конкретной природы самих объектов и отношений.

Система аксиом, сделавшая теорию вероятностей настоящей математической наукой, была создана А.Н.Колмогоровым, на основе:

*а)* теории множеств (Георг Кантор),

*б)* теории интеграла (Анри Лебега),

*в)* теории борелевских множеств (Эмиль Борель).

Метрическая теория функций стала основой строгого построения теории вероятностей.

Знакомство с теорией вероятностей начинается с изучения статистических закономерностей, которые можно наблюдать при проведении испытаний с конечным или счётным множеством возможных исходов.

Словарь теории вероятностей – специфический. Он отличается от словаря математических дисциплин, изучаемых студентом на младших курсах. Это создаёт дополнительные трудности при знакомстве с основными терминами теории вероятностей, развитии интуиции, правилами действия при решении простейших задач, в большинстве которых инструментом для определения *меры возможности наступления события* является *комбинаторика.*

Аксиоматическое построение *вероятностного пространства* позволяет расширить интуитивные представления и практическое истолкование вероятности при знакомстве с более сложными моделями, в которых изучаются статистические закономерности. При этом следует иметь в виду, что целью теории является не вычисление значений вероятности, чем мы занимаемся на начальном этапе знакомства с теорией вероятностей, а построение абстрактных моделей, которые могли бы помочь в изучении различных явлений естествознания. Для этого вводятся понятия *случайной величины, её функции распределения и числовых характеристик.*

После этого становится возможным раскрытие сути общих законов и зависимостей теории вероятностей, которые можно назвать *предельной проблемой теории вероятностей.*

Такая последовательность изложения теории вероятностей, аналогичная последовательностям изложения, принятым в других математических дисциплинах, является плодотворной, и пока нет предложений другого пути изложения.

Книг, учебников, учебных пособий по теории вероятностей написано и издано много. В каждом издании автор стремится, учитывая специфику контингента, для которого оно предназначается, дать описание теории вероятностей и её приложений, в достаточной мере для удовлетворения разнообразных нужд и потребностей этого контингента.

Теорию вероятностей хочется сравнить с айсбергом в океане, большая часть которого скрыта под водой. Изучив его надводную часть можно составить суждение обо всём айсберге. Но эту надводную часть айсберга можно изучать с различными целями, можно приближаться к нему с разных сторон и на разные дистанции, рассматривать его при разной погоде и при разной освещённости солнцем.

Существует несколько экранизаций романов «Война и мир» Л.Н.Толстого и «Тихий Дон» М.А.Шолохова. Каждую экранизацию нельзя рассматривать как простое изложение содержания этих великих романов. В своих экранизациях сценарист и режиссер отражают своё понимание мыслей и чувств авторов этих великих романов XIX и XX веков. Но во всех случаях мы не думаем об этом, а следим за изложением исторических событий, потрясений, поражений и побед российского общества в Отечественной и в Гражданской войнах.

Хочется, чтобы примерно с этих позиций оценивалась и предлагаемая книга.

В основу этой книги положены лекции, которые читались автором в течение ряда лет на III-IV курсе Института математики, механики и компьютерных наук имени И.И.Воровича Южного федерального университета студентам бакалавриата по специальности «прикладная математика и информатика».

Во время лекции, приступая к рассмотрению очередной темы, лектор обычно сначала излагает суть возникающей проблемы, ставит задачу и намечает возможные пути её решения. Эта часть лекции обычно студентом просто выслушивается и остаётся за рамками его конспекта. При самостоятельном изучении материала и подготовке к экзамену с использованием конспекта лекций восстановление в памяти этих моментов лекции обычно весьма затруднительно. Поэтому в предлагаемом учебнике эти моменты теоретического материала излагаются под названием ***комментарий***. Под этим же названием излагается и обсуждение полученных результатов.

В конспекте лекций в основном записаны только определения, формулировки теорем, исходные положения, формулы, а также математические преобразования этих формул, проводимые в ходе доказательств теорем. Эти моменты теоретического материала излагаются в учебнике под названием **теория**.

Иллюстрация полученных результатов и их практическое применение приводятся под названием **практика**.

Такой способ изложения курса «теория вероятностей» должен помочь студенту в более осмысленном понимании и усвоении теоретического материала. Чего не бывает при механическом запоминании перед экзаменом определений и теорем курса из конспекта, которые в таких случаях быстро забываются после сдачи экзамена.

# ГЛАВА ПЕРВАЯ
# Случайное событие
## §1.I Основные понятия и определения теории вероятностей
### *Комментарий*

Напомним, как *происходит* начальное знакомство и *проходит* процесс изучения любого математического курса или предмета, построенного на теоретико-множественной концепции.

Сначала вводятся основные понятия, или термины. Этим понятиям-терминам нельзя дать точное определение. Мы просто договариваемся, описываем: что мы будем понимать, или иметь в виду называя то, или иное понятие-термин.

Используя описание основных понятий, мы формулируем определения объектов, которые изучаются в данном курсе.

Затем перечисляются аксиомы – основные исходные положения, основные утверждения о свойствах изучаемых объектов, которые принимаются без доказательства. Аксиомы – это правила, которым безоговорочно подчиняются.

Основываясь на аксиомах, в соответствии с правилами логической дедукции, формулируются и доказываются теоремы, следствия из теорем о свойствах и отличительных характеристиках объектов, как тех, которые были определены ранее, так и новых объектов, определения которых вводятся по мере расширения круга рассматриваемых вопросов изучаемого курса. Решаются различные задачи, в которых надо применять доказанные теоремы и следствия.

Например, приступая в средней школе к изучению геометрии, мы сначала знакомимся с основными понятиями геометрии: *точка, прямая, луч, отрезок, плоскость, параллельные прямые и т.д.*

Используя эти понятия, мы формулируем определения геометрических фигур - объектов, которые в геометрии являются предметами изучения: *угла, треугольника вообще и треугольников различных видов, параллелограмма, трапеции, окружности и т.д.*

Затем мы принимаем «как истины, которые не требуют доказательства» пять аксиом геометрии – пять постулатов Евклида.

После этого мы, используя правила логической дедукции, формулируем и доказываем различные теоремы и следствия, описывающие различные свойства объектов геометрии, которым ранее были даны определения.

Например: *«Сумма углов любого треугольника равна* $\pi$*»; «В любой треугольник можно вписать окружность, центр которой находится в точке пересечения биссектрис углов треугольника»; «Сумма квадратов длин сторон параллелограмма равна сумме квадратов длин его диагоналей» и т.д.*

Ясно, что по мере углубления в изучаемый предмет, по мере расширения знакомства с простейшими понятиями, определениями и свойствами, изучаемых объектов, мы вводим новые понятия, формулируем новые определения и рассматриваем и доказываем теоремы об их свойствах.

Познакомимся теперь (по аналогии с тем, как это делается в школьном курсе геометрии) с основными понятиями и определениями теории вероятностей, иллюстрируя их примерами.

**Теория**

Первым основным понятием теории вероятностей является понятие *испытания* или *опыта*.

***Испытание или опыт*** – некоторый комплекс условий и действий, направленных на достижение какой-то цели.

Вводя, таким образом, понятие испытания, мы лишь договариваемся: на что мы должны обращать внимание при стремлении или желании реализовать свою конкретную цель. Приведём простейшие примеры испытаний или опытов. При этом мы должны по имеющемуся описанию испытания или опыта иметь возможность мысленно воспроизвести это испытание, не имея в руках монет, колоды карт или игральной кости.

**Практика**

Пример 1. Монета наудачу подбрасывается один раз.

Что имеется в виду в описании этого эксперимента или опыта? Берётся «хорошая», не гнутая монета и «честно» подбрасывается вверх. В соответствии с законом всемирного тяготения монета падает на ровную горизонтальную поверхность.

Пример 2. Монета наудачу подбрасывается три раза (***n*** раз).

Здесь мы три раза (***n*** раз) повторяем условия и действия, записанные в первом примере.

Пример 3. Из колоды карт наудачу извлекается одна карта.

В описании этого эксперимента подразумевается, что берётся «хорошая», тщательно перемешанная колода «некрапленных» карт. Из этой колоды произвольно, случайно выбирается одна карта. Однако, такое описание испытания – не достаточное. Мы не сможем в точности мысленно воспроизвести это испытание, если не будем знать количество карт в колоде.

Пример 4. Игральная кость наудачу подбрасывается один раз.

Игральная кость – кубик, на шести гранях которого нанесены одна, две, три, …, шесть точек, подбрасывается вверх (как и монета) и падает на ровную горизонтальную поверхность.

Пример 5. Монета подбрасывается до первого выпадения «герба».

Мы подбрасываем монету и смотрим на результат – «герб» или «решётка» оказались на верхней грани. Если на верхней грани оказался «герб», то подбрасывания прекращаем, если на верхней грани оказалась «решётка», то снова и снова подбрасываем монету до тех пор, пока на верхней грани не увидим «герб».

**Теория**

Вторым основным понятием теории вероятностей является понятие «элементарно исхода».

***Элементарный исход*** – некоторый *простейший* результат испытания или опыта.

Будем рассматривать испытания, в которых все возможные элементарные исходы можно перечислить. Это позволяет нам каждый элементарный исход обозначать $\omega_i$.

**Практика**

Рассмотрим пять примеров записанных выше испытаний, и выпишем элементарные исходы этих испытаний.

Пример 1. Ясно, что элементарными исходами будут $\omega_1$ - «герб» и $\omega_2$ - «решётка».

В описании понятия «элементарный исход» термин *простейший результат* приводит нас к мысли о возможности нескольких вариантов перечисления результатов испытания или опыта. То есть, термин *простейший результат* предполагают возможность разных уровней подробной детализации результатов испытания, которые определяются целью проводимого испытания.

Во втором примере перечисление всех возможных исходов, как *простейших результатов* испытания, можно сделать двумя вариантами. Выбор варианта зависит от цели проводимого испытания.

Пример 2. *а)* Если нас интересует последовательность выпадений «гербов» и «решёток» при трёх подбрасываниях монеты, то мы запишем восемь элементарных исходов: $\omega_1 - (p,p,p)$, $\omega_2 - (p,p,г)$, $\omega_3 - (p,г,p)$, $\omega_4 - (г,p,p)$, $\omega_5 - (p,г,г)$, $\omega_6 - (г,p,г)$, $\omega_7 - (г,г,p)$, $\omega_8 - (г,г,г)$.

*б)* Если нас интересует только количество «гербов», появившихся в результате трёх подбрасываний монеты, и нам не важно, в какой последовательности они появлялись, то мы запишем четыре элементарных исхода: $\omega_0 - (\textit{ноль гербов})$, $\omega_1 - (\textit{один герб})$, $\omega_2 - (\textit{два герба})$, $\omega_3 - (\textit{три герба})$.

В этом примере комплекс условий и действий в обоих случаях одинаков, а возможные цели испытания позволяют записать два варианта названий элементарного исхода. Но и в первом и во втором случае мы видим, что перечисленные элементарные исходы являются *простейшими результатами* испытания или опыта.

В следующем примере мы увидим, что названия элементарных исходов как простейших результатов испытания определяются порой не только целью испытания, но и подготовленностью испытателя различать варианты результатов опыта.

Пример 3. *а)* Название элементарного исхода должно быть таким, чтобы оно однозначно воспринималось всеми интересующимися этим испытанием и не допускало различных толкований его. Поэтому будет логично определить в этом примере элементарный исход: $\omega_i$ - *номинал карты.* Но мы сразу замечаем, что мы не можем перечислить все элементарные исходы и указать их количество. Это является следствием того, что при определении в примере 3 сути испытания мы, перечисляя комплекс условий и действий с колодой карт, не указали количество карт в колоде. То есть, был перечислен не весь комплекс условий. Будем считать, что в колоде 36 карт. Теперь мы сможем записать элементарный исход: $\omega_i$ - *номинал карты*, где $i = 1,2,\ldots,36$.

Менее подготовленный игрок может не уметь отличать, например, «трефовую семёрку» от «трефовой девятки». Но он может отличить карту трефовой масти от карты пиковой масти, или - «картинку» от «туза», или от «цифры». Поэтому он может предложить два таких варианта названий элементарных исходов в этом испытании:

*б)* $\boldsymbol{\omega}_1 - "\text{карта бубновой масти}"$; $\boldsymbol{\omega}_2 - "\text{карта трефовой масти}"$; $\boldsymbol{\omega}_3 - "\text{карта пиковой масти}"$; $\boldsymbol{\omega}_4 - "\text{карта чирвовой масти}"$.

*в)* $\boldsymbol{\omega}_1 - "\text{цифра}"$; $\boldsymbol{\omega}_2 - "\text{картинка}"$; $\boldsymbol{\omega}_3 - "\text{туз}"$.

Но всегда, записав название элементарного исхода, мы должны убедиться, что при однократном проведении испытания всегда будет реализовываться один и только один элементарный исход.

Пример 4. В этом примере запись элементарных исходов не вызывает затруднений: $\boldsymbol{\omega}_i - "\text{на верхней грани кости } i \text{ точек}"$, где $i = 1,2,3,4,5,6$.

Пример 5. Естественно, название элементарного исхода должно отражать число сделанных подбрасываний монеты. Если «герб» не выпадает, а нам больше делать нечего, то подбрасывать монету мы будем очень долго. Поэтому элементарный исход запишем так: $\boldsymbol{\omega}_i - "\text{сделано } i \text{ подбрасываний монеты}"$, где $i = 1,2,3,\ldots.$, то есть $\boldsymbol{\omega}_i = (\underbrace{p,p,\ldots,p}_{i-1},г)$.

Основное понятие теории множеств - «***множество***» мы понимаем как совокупность объектов, элементов, объединённых по какому-то признаку.

Поэтому все элементарные исходы, которые могут реализоваться в результате проведения испытания, объединим во множество $\boldsymbol{\Omega}$, которое будем называть «*множество элементарных исходов*». Символом $|\boldsymbol{\Omega}|$ будем обозначать количество элементов этого множества.

**Практика**

В примере 1 множество элементарных исходов состоит из двух элементов: $\boldsymbol{\Omega} = \{\omega_1, \omega_2\}$ и $|\boldsymbol{\Omega}| = 2$.

В примере 2 комплекс условий и действий с монетой – испытание позволяет записать два варианта названия элементарного исхода. В первом варианте множество $\boldsymbol{\Omega}$ - множество возможных последовательностей, состоящих из букв «*г*» и «*р*», которые можно составить в результате трёх подбрасываний монеты: $\boldsymbol{\Omega} = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8\}$ и $|\boldsymbol{\Omega}| = 8$. Во втором варианте множество $\boldsymbol{\Omega}$ - множество возможных количеств выпадений «*герба*» в результате трёх подбрасываний монеты: $\boldsymbol{\Omega} = \{\omega_0, \omega_1, \omega_2, \omega_3\}$ и $|\boldsymbol{\Omega}| = 4$.

Пример 3*а).* Здесь $\boldsymbol{\Omega}$ - эквивалентно множеству возможных номиналов карт в колоде, состоящей из 36 карт: $\boldsymbol{\Omega} = \{\omega_i\}$, где $i = 1,2,\ldots 36$ и $|\boldsymbol{\Omega}| = 36$.

Пример 3*б).* Здесь $\boldsymbol{\Omega}$ - эквивалентно множеству возможных «*мастей*» карт в колоде: $\boldsymbol{\Omega} = \{\omega_i\}$, где $i = 1,2,3,4$ и $|\boldsymbol{\Omega}| = 4$.

Пример 3*в).* Здесь $\boldsymbol{\Omega}$ - эквивалентно множеству возможных типов рисунков на картах: $\boldsymbol{\Omega} = \{\omega_i\}$, где $i = 1,2,3$ и $|\boldsymbol{\Omega}| = 3$.

Формулирование и вид записи как элементарных исходов $\boldsymbol{\omega}_i$, так и самого множества элементарных исходов $\Omega$ должны быть чёткими, по возможности быть краткими и не допускать различных вариантов толкования. Умение формулировать и записывать названия этих терминов достигается практикой, накапливающейся при рассмотрении различных, порой внешне не похожих, описаний испытаний или опытов.

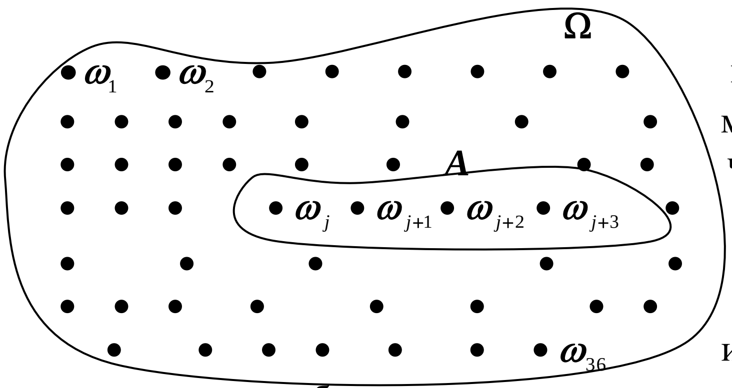
Пример 4. В этом примере запись множества элементарных исходов не вызывает больших затруднений. Сделайте это самостоятельно.

Пример 5. Множеством элементарных исходов $\Omega$ будет множество последовательностей $\omega_i = (\underbrace{p,p,\dots,p}_{i-1},г)$, где $i = 1,2,3,\dots.$, то есть $\Omega = \{\omega_i\}$. Это множество эквивалентно множеству натуральных чисел, то есть $\Omega$ - счётное множество или $|\Omega| = a$.

**Теория**

Рассмотрев несколько примеров основных понятий: испытаний и элементарных исходов, мы можем теперь дать первое определение теории вероятностей – определение ***случайного события***.

Но сначала обратимся к варианту *а)* выделения элементарных исходов в примере 3. Пусть реализация элементарного исхода $\omega_j$ означает извлечение «*бубнового короля*», соответственно, элементарные исходы $\omega_{j+1}, \omega_{j+2}$ и $\omega_{j+3}$ означают извлечение *чирвонного, трефового и пикового короля*.




Представив множество $\Omega$ как множество, состоящее из 36 точек мы видим, что выделенные нами четыре элементарных исхода являются подмножеством множества $\Omega$. С другой стороны, эти элементарные исходы объединены в одно подмножество по объявленному нами единому для всех них качественному признаку: «*короли*». Аналогично, объявляя интересующий нас качественный признак, например - «*картинки*», мы можем указать элементарные исходы, которые этим признаком объединяются в подмножество, состоящее из 12 элементов.

Сформулируем теперь важнейшее определение теории вероятностей.
Определение 1. ***Случайным событием*** называется любое подмножество множества элементарных исходов.

Очевидно, что формулировка случайного события должна содержать качественный признак, по которому мы некоторые элементарные исходы множества $\Omega$ объединяем в подмножество.

Случайные события обозначаются прописными буквами латинского алфавита $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, … .

В примере 3*а* мы можем сформулировать такие события:

случайное событие $\boldsymbol{A}$ – *появился «король»*;

случайное событие $\boldsymbol{B}$ – *появилась «картинка»*;

случайное событие $\boldsymbol{C}$ – *извлечена карта «бубновой масти»*;

случайное событие $\boldsymbol{D}$ – *извлечён «туз»*.

В примере 3*а* слова *появился*, *извлечён* имеют одинаковый смысл, определяемый комплексом условий и действий, при проведении испытания.

В общем случае мы должны иметь в виду, что при однократном проведении испытания, опыта реализуется один и только один элементарный исход $\boldsymbol{\omega}_i$ из множества $\boldsymbol{\Omega}$. В результате проведении испытания, опыта наступает случайное событие $\boldsymbol{A}$, являющееся некоторым подмножеством множества $\boldsymbol{\Omega}$.

Определение 2. Элементарный исход $\boldsymbol{\omega}_i$, реализация которого означает наступление случайного события $\boldsymbol{A}$, будем называть ***благоприятствующим исходом*** событию $\boldsymbol{A}$.

Так как случайное событие $\boldsymbol{A}$ является множеством, то, переписывая все элементарные исходы, благоприятствующие этому событию, можем, например, записать событие $\boldsymbol{A}$ как последовательность: $\boldsymbol{A}=\{\boldsymbol{\omega}_j, \boldsymbol{\omega}_{j+1}, \boldsymbol{\omega}_{j+2}, \boldsymbol{\omega}_{j+3}\}$ и подсчитать количество исходов, которые ему благоприятствуют: $|\boldsymbol{A}|=\boldsymbol{m}=4$.

Из курса линейной алгебры мы знаем, что можно проводить операции объединения, пересечения, вычитания (разности) множеств, в результате которых получаются новые множества. Случайные события, которые могут наступить в результате проведения испытания, являются множествами, следовательно, со случайными событиями можно проводить операции объединения, пересечения, вычитания. Но, так как теория вероятностей и её терминология начали своё развитие раньше теории множеств, то эти операции со случайными событиями имеют в теории вероятностей своё название.

Определение 3. Случайное событие $\boldsymbol{C}$ называется ***суммой*** событий $\boldsymbol{A}$ и $\boldsymbol{B}$, если ему благоприятствуют элементарные исходы, благоприятствующие ***или*** событию $\boldsymbol{A}$, ***или*** событию $\boldsymbol{B}$.

Ясно, что в теории множеств аналогом суммы событий в теории вероятностей будет *объединение* множеств. Поэтому можем записать: $\boldsymbol{C}=\boldsymbol{A}\bigcup\boldsymbol{B}=\{\boldsymbol{\omega_i}\in\boldsymbol{\Omega}: \vee\boldsymbol{\omega_i}\in\boldsymbol{A}, \vee\boldsymbol{\omega_i}\in\boldsymbol{B}\}$.

Если $\boldsymbol{A}$ и $\boldsymbol{C}$ – случайные события, записанные в примере 3*а*, то наступление случайного события $\boldsymbol{A}\bigcup\boldsymbol{C}$ мы понимаем как появление или карты «король», или карты «бубновая масть».

Определение 4. Случайное событие $\boldsymbol{D}$ называется ***произведением*** событий $\boldsymbol{A}$ и $\boldsymbol{B}$, если ему благоприятствуют элементарные исходы, благоприятствующие ***и*** событию $\boldsymbol{A}$, ***и*** событию $\boldsymbol{B}$.

В теории множеств аналогом произведения событий в теории вероятностей будет *пересечение* множеств. Поэтому можем записать: $\boldsymbol{D}=\boldsymbol{A}\bigcap\boldsymbol{B}=\{\boldsymbol{\omega}_i\in\boldsymbol{\Omega}: \wedge\boldsymbol{\omega}_i\in\boldsymbol{A}, \wedge\boldsymbol{\omega}_i\in\boldsymbol{B}\}$.

Для случайных событий $\boldsymbol{A}$ и $\boldsymbol{C}$, записанных в примере 3*а***,** наступление случайного события $\boldsymbol{A}\bigcap\boldsymbol{C}$ мы понимаем как появление карты «бубновый король».

<u>Определение 5</u>. Будем говорить: «Случайное событие $\boldsymbol{A}$ ***влечёт за собой*** случайное событие $\boldsymbol{B}$», если наступление случайного события $\boldsymbol{A}$ одновременно означает и наступление случайного события $\boldsymbol{B}$.

То есть все элементарные исходы, которые благоприятствуют наступлению случайного события $\boldsymbol{A}$, будут в это же время благоприятствовать и наступлению случайного события $\boldsymbol{B}$.

В теории множеств аналогом термину «$\boldsymbol{A}$ ***влечёт за собой*** $\boldsymbol{B}$» будет термин «*множество* $\boldsymbol{A}$ *является подмножеством множества* $\boldsymbol{B}$». Поэтому, если случайное событие $\boldsymbol{A}$ влечёт за собой случайное событие $\boldsymbol{B}$, будем применять обозначение $\boldsymbol{A} \subset \boldsymbol{B}$.

Ясно, что в примере 3*а* наступление случайного события $\boldsymbol{A}$ – появился «король» одновременно означает и наступление случайного события $\boldsymbol{B}$ – появилась «картинка».

<u>Определение 6</u>. Случайное событие $\boldsymbol{U}$ называется ***достоверным***, если ему благоприятствуют все элементарные исходы множества элементарных исходов $\boldsymbol{\Omega}$. То есть при однократном проведении испытания или опыта достоверное события обязательно наступит.

<u>Определение 7</u>. Случайное событие $\boldsymbol{V}$ называется ***невозможным***, если ему не благоприятствует ни один элементарный исход множества элементарных исходов $\boldsymbol{\Omega}$. Следовательно, случайное событие $\boldsymbol{V}$ это пустое подмножество $\boldsymbol{\varnothing}$ множества элементарных исходов и при проведении испытания или опыта невозможное событие никогда не наступит.

<u>Определение 8</u>. Будем говорить: «Случайные события $\boldsymbol{A}$ и $\boldsymbol{B}$ - ***несовместны***», если у них нет общих благоприятствующих им элементарных исходов.

Значит, несовместные случайные события не могут наступить одновременно. В теории множеств мы про такие множества говорим, что они не пересекаются.

<u>Определение 9</u>. Случайное событие $\boldsymbol{E}$ называется ***разностью*** событий $\boldsymbol{A}$ и $\boldsymbol{B}$, если ему благоприятствуют элементарные исходы, благоприятствующие событию $\boldsymbol{A}$, которые при этом не благоприятствуют событию $\boldsymbol{B}$.

Используя обозначения теории множеств, можем записать: $\boldsymbol{E} = \boldsymbol{A} \setminus \boldsymbol{B} = \{\boldsymbol{\omega}_i \in \boldsymbol{\Omega} : \boldsymbol{\omega}_i \in \boldsymbol{A}, \wedge \boldsymbol{\omega}_i \notin \boldsymbol{B}\}$.

<u>Определение 10</u>. Случайное событие $\overline{\boldsymbol{A}}$ называется ***противоположным*** случайному событию $\boldsymbol{A}$, если ему благоприятствуют все элементарные исходы множества элементарных исходов $\boldsymbol{\Omega}$, которые не благоприятствуют случайному событию $\boldsymbol{A}$.

Кратко это запишем так: $\overline{\boldsymbol{A}} = \{\boldsymbol{\omega}_i \in \boldsymbol{\Omega} : \boldsymbol{\omega}_i \notin \boldsymbol{A}\}$.

Из определения противоположного случайного события $\overline{\boldsymbol{A}}$ следует, что всегда будут выполняться равенства: $\boldsymbol{A} \bigcup \overline{\boldsymbol{A}} = \boldsymbol{\Omega}$ и $\boldsymbol{A} \bigcap \overline{\boldsymbol{A}} = \boldsymbol{\varnothing}$.

***Комментарий***

Ясно, что сформулировав, назвав некоторые случайные события, мы можем сформулировать, назвать новые случайные события, которые будут суммами или произведениями этих случайных событий. По этим новым случайным событиям мы можем сформулировать названия случайных событий, которые будут произведениями и суммами этих новых событий. Такой процесс формулирования в теории вероятностей новых событий эквивалентен процессу образования системы подмножеств рассматриваемого множества. Получающаяся система подмножеств в теории множеств называется *алгеброй* подмножеств рассматриваемого множества. Дадим точное определение *алгебры* подмножеств в терминах теории вероятностей.

**Теория**

Определение 11. Совокупность случайных событий называется ***алгеброй случайных событий*** $\mathcal{A}$, если её элементы (случайные события) удовлетворяют следующим требованиям:

*1)* $\boldsymbol{\Omega}\in\mathcal{A}$; *2) если* $\boldsymbol{A}\in\mathcal{A}$ *и* $\boldsymbol{B}\in\mathcal{A}$, *то* $\boldsymbol{A}\bigcup\boldsymbol{B}\in\mathcal{A}$ *и* $\boldsymbol{A}\bigcap\boldsymbol{B}\in\mathcal{A}$; *3) если* $\boldsymbol{A}\in\mathcal{A}$, *то* $\overline{\boldsymbol{A}}\in\mathcal{A}$.

Элементами множества $\boldsymbol{\Omega}$ являются элементарные исходы $\boldsymbol{\omega}_i$. Любое подмножество множества $\boldsymbol{\Omega}$ является случайным событием. Эти подмножества, включая и само множество $\boldsymbol{\Omega}$, являются элементами множества $\mathcal{A}$. То есть множество $\mathcal{A}$ (алгебра случайных событий) – «конструкция» более высокого иерархического уровня, чем множество $\boldsymbol{\Omega}$ (множество элементарных исходов).

Очевидно, что *минимальной* алгеброй событий некоторого испытания будет алгебра $\mathcal{A}_*$, состоящая из двух элементов $\boldsymbol{\Omega}$ и $\boldsymbol{\varnothing}$, то есть $\mathcal{A}_*=\{\boldsymbol{\Omega};\boldsymbol{\varnothing}\}$. Алгебра $\mathcal{A}_A=\{\boldsymbol{\Omega};\boldsymbol{\varnothing};\boldsymbol{A};\overline{\boldsymbol{A}}\}$ называется алгеброй порождённой случайным событием $\boldsymbol{A}$. *Максимальной* алгеброй будет алгебра $\mathcal{A}^*$, состоящая из всех подмножеств множества $\boldsymbol{\Omega}$. Эта алгебра состоит из $2^{|\Omega|}$ элементов.

Определение 12. Пара объектов: множество элементарных исходов $\boldsymbol{\Omega}$ и алгебра случайных событий $\mathcal{A}$ называется ***измеримым пространством*** и обозначается $<\boldsymbol{\Omega},\mathcal{A}>$.

## §2.I.Введение вероятностной функции

***Комментарий***

Термин *измеримое пространство* означает для нас, что объекты - случайные события (в разных ролях входящие в $\boldsymbol{\Omega}$ и в $\mathcal{A}$) мы будем **измерять,** то есть каждому случайному событию мы будем ставить в соответствие неотрицательное число, которое будем понимать как **меру** этого события.

Мы, изучая или просто рассматривая объект, стараемся как-то измерить в некотором смысле этот объект. Далекий наш предок, например, старался подсчитать количество овец в своём стаде и количество «врагов» в соседнем враждебном племени. В построенной по аксиомам Евклида геометрии греки измеряли длины отрезков, площади треугольников и трапеций, объёмы призм и конусов. И.Кеплер нашёл объёмы многих тел вращения. Введя понятие функции, и построив в системе координат Декарта графики различных функций, математики стали измерять скорость изменения функции, длины интервалов на которых значения функции сохраняют постоянный знак, вычислять площадь криволинейной трапеции и объёмы произвольных тел.

Анализируя демографические изменения численности населения, развивая страхование от несчастных случаев и перевозки товаров парусными кораблями, наблюдая устойчивость результатов азартных игр, математики стали делать попытки измерить наблюдаемую устойчивость частоты случайных событий при проведении большого числа испытаний. Эти попытки привели к введению понятия **вероятности наступления случайного события** как *меры уверенности* в его наступлении при проведении испытания, опыта.

Математики XVII – XVIII века рассматривали испытания или опыты, в которых множество элементарных исходов $\boldsymbol{\Omega}$ имело конечное число элементов $|\boldsymbol{\Omega}| = \boldsymbol{n}$, то есть все возможные элементарные исходы $\boldsymbol{\omega}_i$ можно было выписать в виде конечной последовательности: $\{\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_n\}$. Любое подмножество этой последовательности является случайным событием $\boldsymbol{A}$ – элементом алгебры случайных событий $\mathcal{A}^*$, содержащей $|\mathcal{A}^*| = 2^n$ элементов.

Понятие вероятности случайного события в теории вероятностей, как и понятие площади в геометрии, вводится аксиоматически.

**Теория**

Рассматривая конкретное испытание или опыт, мы строим измеримое пространство $<\boldsymbol{\Omega}, \mathcal{A}>$.

*Каждому элементарному исходу* $\boldsymbol{\omega}_i$ *некоторым разумным способом поставим в соответствие положительное число* $\boldsymbol{p}(\boldsymbol{\omega}_i)$, *кратко:* $\boldsymbol{\omega}_i \Rightarrow \boldsymbol{p}(\boldsymbol{\omega}_i) > 0$, *так чтобы сумма всех этих положительных чисел, когда мы переберём все элементарные исходы множества* $\boldsymbol{\Omega}$, *была равна единице, то есть:* $\sum_{\omega_i \in \Omega} \boldsymbol{p}(\boldsymbol{\omega}_i) = 1$. *Тогда* ***вероятность*** $\boldsymbol{P}(\boldsymbol{A})$ *наступления случайного события* $\mathbf{A}$, *являющегося элементом алгебры событий* $\mathcal{A}$, *будет равна сумме тех положительных чисел* $\boldsymbol{p}(\boldsymbol{\omega}_i)$, *для которых элементарный исход* $\boldsymbol{\omega}_i$ *является благоприятствующим наступлению этого случайного события, или, кратко: для* $\forall \boldsymbol{A} \in \mathcal{A}$ *вероятность его наступления равна* $\boldsymbol{P}(\boldsymbol{A}) = \sum_{\omega_i \in A} \boldsymbol{p}(\boldsymbol{\omega}_i)$.

Мы ввели понятие вероятности наступления случайного события $\boldsymbol{A}$, как меру подмножества $\boldsymbol{A}$ множества $\boldsymbol{\Omega}$, являющегося элементом алгебры $\mathcal{A}$. Каждому элементу $\boldsymbol{A}$ алгебры $\mathcal{A}$ вероятностная функция $\boldsymbol{P}$ ставит в соответствие число $\boldsymbol{P}(\boldsymbol{A})$, называемое вероятностью наступления при проведении испытания случайного события $\boldsymbol{A}$.

<u>Определение.</u> Тройку объектов $<\boldsymbol{\Omega}, \mathcal{A}, \boldsymbol{P}>$, где $\boldsymbol{\Omega}$ - множество элементарных исходов, $\mathcal{A}$ - алгебра случайных событий, $\boldsymbol{P}$ – вероятностная функция, называется ***вероятностным пространством***.

Решение задач определения вероятностей наступлений случайных событий, предусматривает рассмотрение и описание этих трёх объектов для каждого конкретного испытания, опыта.

***Комментарий***

Практическое применение сделанного аксиоматического определения вероятности $\boldsymbol{P}(\boldsymbol{A})$ наступления случайного события $\boldsymbol{A}$ при проведении испытания затрудняется неопределённостью выбора набора положительных чисел, который должен подбираться *некоторым разумным способом*. Ясно, что перечислить все разумные способы выбора набора $\{\boldsymbol{p}(\boldsymbol{\omega}_i)\}$ для различных испытаний мы не сможем. Да и как определить: *разумно*, или *неразумно* для рассматриваемого испытания экспериментатором установлено соответствие $\boldsymbol{\omega}_i \Rightarrow \boldsymbol{p}(\boldsymbol{\omega}_i) > 0$?

Единственным теоретически обоснованным способом проверки правильности выбора набора $\{\boldsymbol{p}(\boldsymbol{\omega}_i)\}$ является применение теоремы Бернулли из «Закона больших чисел». Но до этой теоремы нам надо ещё «дойти», - у нас на данном этапе знакомства с теорией вероятностей ещё слишком маленький запас теоретических знаний.

Однако можно рассмотреть наиболее часто встречающиеся в практической жизни испытания, разбить их на группы, в которых аналогично формулируются элементарные исходы, и, сообразуясь со здравым смыслом, описать метод определения *разумным способом* набора $\{\boldsymbol{p}(\boldsymbol{\omega}_i)\}$ в каждой из выделенных групп.

Условие $\sum\limits_{\omega_i \in \Omega} \boldsymbol{p}(\boldsymbol{\omega}_i) = 1$ будем называть условием нормировки.

Естественно, что первый разумный способ установления соответствия $\boldsymbol{\omega}_i \Rightarrow \boldsymbol{p}(\boldsymbol{\omega}_i)$ был описан при рассмотрении испытаний, которые встречаются при описании многих азартных игр. Он называется ***классическим.***

## §3.I Классическое определение вероятности

### Теория

Рассматриваются испытания, в которых выполняются два условия:

1) множество элементарных исходов имеет ***конечное число элементов***, то есть $|\boldsymbol{\Omega}| = \boldsymbol{n} < \infty$;

2) все элементарные исходы $\{\boldsymbol{\omega}_i\}$, которые являются простейшими возможными результатами испытания, – ***равновозможные***.

***Комментарий***

Равновозможность реализации элементарных исходов при проведении испытания – одно из основных понятий теории вероятностей, которому мы не можем дать определение. Но мы хорошо понимаем, что термин равновозможность элементарных исходов означает, что все исходы имеют равные шансы на реализацию при проведении испытания. Нет ни одного исхода, который имел бы преимущество на реализацию по сравнению с остальными элементарными исходами.

*Замечание.* Ясно, что объяснение понятия «равновозможность» как «равенство шансов на реализацию» аналогично ответу на вопрос: «Что такое масло?» - «Это то, что масляное!». Использованный вариант объяснения понятия «равновозможность» ещё раз иллюстрирует утверждение, о том, что основным понятиям нельзя дать точное определение. Но мы уверенны, что всегда можно договориться о том, что мы будем понимать под вводимым термином.

**Теория**

Если перечисленные два условия выполнены, то будет *разумно* каждому элементарному исходу $\boldsymbol{\omega}_i$ поставить в соответствие положительное число $\frac{1}{|\boldsymbol{\Omega}|}$, то есть: $\boldsymbol{\omega}_i \Rightarrow \boldsymbol{p}(\boldsymbol{\omega}_i) = \frac{1}{\boldsymbol{n}}$.

Ясно, что условие нормировки выполняется: $\sum\limits_{\omega_i \in \Omega} \boldsymbol{p}(\boldsymbol{\omega}_i) = \sum\limits_{i=1}^{n} \frac{1}{\boldsymbol{n}} = 1$. Тогда для $\forall \boldsymbol{A} \in \mathcal{A}$ вероятность его наступления будет равна: $\boldsymbol{P}(\boldsymbol{A}) = \sum\limits_{\omega_i \in A} \boldsymbol{p}(\boldsymbol{\omega}_i) = \sum\limits_{\omega_i \in A} \frac{1}{\boldsymbol{n}} = \frac{\boldsymbol{m}}{\boldsymbol{n}}$. То есть, если наступлению случайного события $\boldsymbol{A}$ благоприятствуют $\boldsymbol{m}$ элементарных исходов из множества $\boldsymbol{\Omega}$, состоящему из $\boldsymbol{n}$ элементов, то вероятность наступления события случайного события $\boldsymbol{A}$ будет равна отношению числа благоприятствующих ему элементарных исходов к общему числу элементарных исходов: $\boldsymbol{P}(\boldsymbol{A}) = \frac{\boldsymbol{m}}{\boldsymbol{n}}$.

***Комментарий***

Разделить число $\boldsymbol{m}$ на число $\boldsymbol{n}$ – нетрудно, но для определения вероятности наступления случайного события мы должны проделать некоторые действия. Последовательность этих действий при решении задачи с помощью применения классического определения вероятности можно представить так.

*а)* Сначала мы должны мысленно смоделировать комплекс условий испытания и представить действия экспериментатора, направленные на достижение цели испытания.

*б)* Даём чёткое название элементарного исхода, не допускющее возможности различных толкований его.

*в)* Определяем количество элементарных исходов, которые могут реализоваться при проведении испытания.

*г)* Проверяем выполнение условий, при которых можно будет применить классическое определение вероятности.

*д)* Даём название случайного события, наступление которого интересует экспериментатора, то есть - название цели испытания.

*е)* Подсчитываем количество элементарных исходов, благоприятствующих наступлению этого случайного события.

*ж)* Вычисляем значение дроби $\frac{\boldsymbol{m}}{\boldsymbol{n}}$, которое будет искомой вероятностью наступления случайного события $\boldsymbol{A}$.

Последовательно рассматривая приведённые ранее примеры испытаний, определим вероятности наступления некоторых случайных событий, применяя, там, где это возможно классическое определение вероятности.

**Практика**

Пример 1. Множество элементарных исходов состоит из двух элементов: $\boldsymbol{\Omega}=\{\omega_1,\omega_2\}$. Первое условие применения классического определения вероятности выполняется: $|\boldsymbol{\Omega}|=\boldsymbol{n}=2$. Если выбранная монета «не гнутая» и она подбрасывается на ровную, горизонтальную поверхность, то у нас нет сомнений в равновозможности названных элементарных исходов на реализацию при проведении опыта. Второе условие применения классического определения вероятности выполняется. Следовательно, будет разумно каждому элементарному исходу: $\omega_1$ и $\omega_2$ поставить в соответствие число: $p(\omega_i)=\frac{1}{2}$. Дадим название случайному событию $\boldsymbol{A}$ - *на верхней грани монеты* - «*герб*». Этому случайному событию благоприятствует один элементарный исход: $\boldsymbol{A}=\{\omega_1\}$, значит: $|\boldsymbol{A}|=\boldsymbol{m}=1$. Следовательно, применяя классическое определение вероятности, получаем, что вероятность наступления случайного события $\boldsymbol{A}$ равна $\boldsymbol{P}(\boldsymbol{A})=\frac{\boldsymbol{m}}{\boldsymbol{n}}=\frac{1}{2}$.

В примере 2 мы записали два варианта выделения элементарных исходов – простейших результатов опыта. В первом варианте множество $\boldsymbol{\Omega}$ состоит из восьми элементов: $|\boldsymbol{\Omega}|=\boldsymbol{n}=8$, а во втором - множество $\boldsymbol{\Omega}$ состоит из четырёх элементов: $|\boldsymbol{\Omega}|=\boldsymbol{n}=4$.

Равновозможность на реализацию при проведении испытания у всех элементарных исходов первого варианта примера не вызывает сомнения.

Определим случайное событие $\boldsymbol{A}$ – *в результате трёх подбрасываний монеты «герб» появился один раз*. Наступлению этого случайного события благоприятствуют три элементарных исхода $\boldsymbol{A}=\{\omega_2,\omega_3,\omega_4\}$, то есть: $|\boldsymbol{A}|=\boldsymbol{m}=3$. Следовательно, применяя классическое определение вероятности, получаем, что вероятность наступления случайного события $\boldsymbol{A}$ равна $\boldsymbol{P}(\boldsymbol{A})=\frac{\boldsymbol{m}}{\boldsymbol{n}}=\frac{3}{8}$.

Во втором варианте наступлению этого случайного события благоприятствует один элементарный исход $\boldsymbol{A}=\{\omega_1\}$, то есть: $|\boldsymbol{A}|=\boldsymbol{m}=1$. Получается, что вероятность наступления случайного события $\boldsymbol{A}$ равна $\boldsymbol{P}(\boldsymbol{A})=\frac{\boldsymbol{m}}{\boldsymbol{n}}=\frac{1}{4}$.

Ясно, что вероятность наступления события $\boldsymbol{A}$ не должна зависеть от варианта выделения элементарного исхода. Но так как $\frac{3}{8}\neq\frac{1}{4}$, то в одном из вариантов выделения элементарных исходов мы не имели права применять классическое определение вероятности.

Первое условие применения этого определения выполнено: оба множества элементарных исходов имеют конечное число элементов.

Второе условие предусматривает равновозможную реализацию каждого элементарного исхода. Равновозможность на реализацию при проведении испытания элементарных исходов в первом варианте выделения не вызывает сомнения. Во втором варианте условие равновозможности не выполняется. Например, исход $\boldsymbol{\omega}_1 - (\textit{один} \quad \textit{герб})$ имеет при трёх подбрасываниях монеты в три раза больше «шансов» на реализацию, чем элементарный исход $\boldsymbol{\omega}_0 - (\textit{ноль} \quad \textit{гербов})$. Значит, применять классическое определение вероятности в этом варианте мы не имели права.

В примере 3*а)* множество $\boldsymbol{\Omega}$ состоит из $|\boldsymbol{\Omega}| = \boldsymbol{n} = 36$ элементов. Если колода карт тщательно перемешана и карта извлекается «наудачу», то все элементарные исходы имеют равные возможности на реализацию. А поэтому, считая, что $\boldsymbol{p}(\boldsymbol{\omega}_i) = \frac{1}{36}$ для любого из 36 возможных элементарных исходов, мы, подсчитав количество благоприятствующих исходов каждому случайному событию, можем быстро определить вероятности наступлений этих случайных событий:

случайному событию $\boldsymbol{A}$ – *появился «король»* - благоприятствуют $|\boldsymbol{A}| = \boldsymbol{m} = 4$ элементарных исхода и $\boldsymbol{P}(\boldsymbol{A}) = \frac{\boldsymbol{m}}{\boldsymbol{n}} = \frac{4}{36} = \frac{1}{9}$;

случайному событию $\boldsymbol{B}$ – *появилась «картинка»* - благоприятствуют $|\boldsymbol{B}| = \boldsymbol{m} = 12$ элементарных исходов и, следовательно, $\boldsymbol{P}(\boldsymbol{B}) = \frac{\boldsymbol{m}}{\boldsymbol{n}} = \frac{12}{36} = \frac{1}{3}$;

случайному событию $\boldsymbol{C}$ – *извлечена карта «бубновой масти»* - благоприятствуют $|\boldsymbol{C}| = \boldsymbol{m} = 9$ элементарных исходов и $\boldsymbol{P}(\boldsymbol{C}) = \frac{\boldsymbol{m}}{\boldsymbol{n}} = \frac{9}{36} = \frac{1}{4}$;

случайному событию $\boldsymbol{D}$ – *извлечён «туз»* - благоприятствуют $|\boldsymbol{D}| = \boldsymbol{m} = 4$ элементарных исхода и $\boldsymbol{P}(\boldsymbol{D}) = \frac{\boldsymbol{m}}{\boldsymbol{n}} = \frac{4}{36} = \frac{1}{9}$.

В примере 3*б)* множество элементарных исходов – конечное, названные элементарные исходы являются равновозможными. Но мы можем определить вероятность наступления только одного из четырёх рассмотренных выше событий. Вероятность наступления случайного события $\boldsymbol{C}$ будет равна $\boldsymbol{P}(\boldsymbol{C}) = \frac{\boldsymbol{m}}{\boldsymbol{n}} = \frac{1}{4}$. Естественно, что этот результат совпадает с результатом примера 3*а)*. Вероятности наступления остальных трёх случайных событий мы при таком варианте записи элементарных исходов мы не сможем определить. В таких случаях будем говорить, что для определения вероятностей наступления этих событий элементарные исходы выделены неудачно.

Ясно, что в примере 3*в)* элементарные исходы не будут равновозможными, а поэтому попытка определить вероятность наступления случайного события $\boldsymbol{D}$ приводит к неверному результату $\boldsymbol{P}(\boldsymbol{D}) = \frac{\boldsymbol{m}}{\boldsymbol{n}} = \frac{1}{3}$, который противоречит результату, полученному в примере 3*а)*.

Пример 5. Здесь множество элементарных исходов $\mathbf{\Omega}$ - счётное множество или $|\mathbf{\Omega}| = a$. Классическое определение вероятности применять мы не можем.

## §4.I Простейшие свойства вероятностной функции

### Теория

Введя аксиоматически вероятностную функцию для испытаний, в которых множество элементарных исходов не более чем счётное, мы можем сформулировать теоремы, в которых рассматриваются простейшие свойства вероятностной функции $\boldsymbol{P}$. Доказательство этих теорем мы пока сможем провести, рассматривая только конечные множества $\mathbf{\Omega}$, в которых все элементарные исходы – равновозможные. Будем рассматриваемые теоремы называть свойствами вероятностной функции.

*Свойство 1.* Вероятность наступления любого случайного события $\boldsymbol{A}$ является числом, заключённым между нулём и единицей, то есть:

$$\textit{Для } \forall \boldsymbol{A} \in \boldsymbol{\mathcal{A}} \quad \textit{справедливо } 0 \le \boldsymbol{P}(\boldsymbol{A}) \le 1.$$

Так как любое случайное событие является подмножеством множества элементарных исходов $\mathbf{\Omega}$, то число элементарных исходов $\boldsymbol{m}$, ему благоприятствующих, не может быть отрицательным и не может быть больше чем $\boldsymbol{n}$ - число элементов множества $\mathbf{\Omega}$, то есть всегда справедливо: $0 \le \boldsymbol{m} \le \boldsymbol{n}$. Разделив члены этого двойного неравенства на число $\boldsymbol{n}$ и, учитывая равновозможность реализации всех элементарных исходов, убеждаемся в справедливости утверждения: $0 \le \boldsymbol{P}(\boldsymbol{A}) \le 1$.

*Свойство 2.* Если случайные события $\boldsymbol{A}$ и $\boldsymbol{B}$ – несовместные, то вероятность наступления суммы этих событий равна сумме вероятностей наступлений этих событий, то есть:

$$\textit{Если } \boldsymbol{A} \cap \boldsymbol{B} = \boldsymbol{\varnothing}\textit{, то будет справедливо равенство } \boldsymbol{P}(\boldsymbol{A} \cup \boldsymbol{B}) = \boldsymbol{P}(\boldsymbol{A}) + \boldsymbol{P}(\boldsymbol{B}).$$

Пусть наступлению случайных событий $\boldsymbol{A}$ и $\boldsymbol{B}$ благоприятствуют, соответственно $\boldsymbol{m}$ и $\boldsymbol{k}$ элементарных исходов. Так как эти события – несовместные, то наступлению случайного события $\boldsymbol{A} \cup \boldsymbol{B}$ будут благоприятствовать $\boldsymbol{m} + \boldsymbol{k}$ элементарных исходов.

Все исходы – равновозможные. Следовательно, применяя классическое определение вероятности, запишем $\boldsymbol{P}(\boldsymbol{A} \cup \boldsymbol{B}) = \dfrac{\boldsymbol{m} + \boldsymbol{k}}{\boldsymbol{n}}$. Так как $\dfrac{\boldsymbol{m} + \boldsymbol{k}}{\boldsymbol{n}} = \dfrac{\boldsymbol{m}}{\boldsymbol{n}} + \dfrac{\boldsymbol{k}}{\boldsymbol{n}}$, то, заменив в правой части полученные дроби на вероятности событий $\boldsymbol{A}$ и $\boldsymbol{B}$, получим требуемое равенство.

*Следствие.* Для любого конечного числа попарно несовместных событий вероятность их суммы равна сумме их вероятностей: $\boldsymbol{P}\left(\bigcup\limits_{\boldsymbol{i}=1}^{\boldsymbol{l}} \boldsymbol{A}_{\boldsymbol{i}}\right) = \sum\limits_{\boldsymbol{i}=1}^{\boldsymbol{l}} \boldsymbol{P}(\boldsymbol{A}_{\boldsymbol{i}})$.

Доказательство легко провести, используя метод математической индукции.

*Свойство 3.* Если наступление случайного события $\boldsymbol{A}$ *влечёт за собой* наступление случайного события $\boldsymbol{B}$, то вероятность наступления события $\boldsymbol{A}$

будет не больше, чем вероятность наступления события $\boldsymbol{B}$, то есть:

$$\textit{Если } \boldsymbol{A} \subset \boldsymbol{B}\textit{, то } \boldsymbol{P}(\boldsymbol{A}) \leq \boldsymbol{P}(\boldsymbol{B}).$$

Случайное событие $\boldsymbol{B}$, согласно определению термина «*влечёт за собой*», можно представить в виде суммы двух событий: $\boldsymbol{B} = \boldsymbol{A} \bigcup (\boldsymbol{B} \setminus \boldsymbol{A})$. Причём события - слагаемые будут несовместными событиями. Применяя свойство 2, запишем:

$$\boldsymbol{P}(\boldsymbol{B}) = \boldsymbol{P}(\boldsymbol{A} \bigcup (\boldsymbol{B} \setminus \boldsymbol{A})) = \boldsymbol{P}(\boldsymbol{A}) + \boldsymbol{P}(\boldsymbol{B} \setminus \boldsymbol{A}).$$

А так как по первому свойству вероятность наступления любого события – неотрицательное число, $\boldsymbol{P}(\boldsymbol{B} \setminus \boldsymbol{A}) \geq 0$, то приходим к выводу, что $\boldsymbol{P}(\boldsymbol{B}) \geq \boldsymbol{P}(\boldsymbol{A})$

*Свойство 4.* Вероятность достоверного события равна единице, вероятность невозможного события равна нулю:

$$\boldsymbol{P}(\boldsymbol{U}) = 1 \textit{ и } \boldsymbol{P}(\boldsymbol{V}) = 0.$$

Доказательство этого свойства не вызывает затруднений, если учесть, что, согласно определениям достоверного и невозможного событий: $|\boldsymbol{U}| = |\boldsymbol{\Omega}| = \boldsymbol{n}$ и $|\boldsymbol{V}| = |\varnothing| = 0$.

*Свойство 5.* Вероятность наступления суммы двух произвольных случайных событий равна сумме вероятностей наступлений этих событий, уменьшенной на вероятность наступления их произведения:

$$\boldsymbol{P}(\boldsymbol{A} \bigcup \boldsymbol{B}) = \boldsymbol{P}(\boldsymbol{A}) + \boldsymbol{P}(\boldsymbol{B}) - \boldsymbol{P}(\boldsymbol{A} \bigcap \boldsymbol{B}).$$

Как и при доказательстве свойства 2 предположим, что наступлению случайных событий $\boldsymbol{A}$ и $\boldsymbol{B}$ благоприятствуют, соответственно $\boldsymbol{m}$ и $\boldsymbol{k}$ элементарных исходов, и пусть наступлению случайного события $\boldsymbol{A} \bigcap \boldsymbol{B}$ благоприятствуют $\boldsymbol{l}$ элементарных исходов, которые одновременно благоприятствуют наступлению и события $\boldsymbol{A}$, и события $\boldsymbol{B}$. Ясно, что $\boldsymbol{l} \leq \boldsymbol{m}$ и $\boldsymbol{l} \leq \boldsymbol{k}$.

Подсчитывая количество элементарных исходов благоприятствующих случайному событию $\boldsymbol{A} \bigcup \boldsymbol{B}$, мы сначала переберём все исходы, которые благоприятствуют событию $\boldsymbol{A}$, а потом переберём все исходы, которые благоприятствуют событию $\boldsymbol{B}$. Но при этом $\boldsymbol{l}$ элементарных исходов попадут под наше внимание дважды. Следовательно, случайному событию $\boldsymbol{A} \bigcup \boldsymbol{B}$ будут благоприятствовать $\boldsymbol{m} + \boldsymbol{k} - \boldsymbol{l}$ элементарных исходов. Все исходы - равновозможные, следовательно, применяя классическое определение вероятности, запишем $\boldsymbol{P}(\boldsymbol{A} \bigcup \boldsymbol{B}) = \dfrac{\boldsymbol{m} + \boldsymbol{k} - \boldsymbol{l}}{\boldsymbol{n}}$. А отсюда легко получаем:

$$\boldsymbol{P}(\boldsymbol{A} \bigcup \boldsymbol{B}) = \frac{\boldsymbol{m} + \boldsymbol{k} - \boldsymbol{l}}{\boldsymbol{n}} = \frac{\boldsymbol{m}}{\boldsymbol{n}} + \frac{\boldsymbol{k}}{\boldsymbol{n}} - \frac{\boldsymbol{l}}{\boldsymbol{n}} = \boldsymbol{P}(\boldsymbol{A}) + \boldsymbol{P}(\boldsymbol{B}) - \boldsymbol{P}(\boldsymbol{A} \bigcap \boldsymbol{B}).$$

*Замечание.* Если случайные события $\boldsymbol{A}$ и $\boldsymbol{B}$ несовместные, то общее число благоприятствующих им элементарных исходов будет равно нулю ($\boldsymbol{l}$=0). Следовательно, свойство 2 является частным случаем свойства 3.

*Следствие.* Применяя метод математической индукции, мы можем обобщить свойство 3 для любого конечного числа случайных событий - слагаемых. В частности, вероятность наступлении суммы трёх случайных событий будет равна:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

*Свойство 6.* Если случайное событие $\overline{A}$ - противоположное событие случайному событию $A$, то: $P(A) + P(\overline{A}) = 1$.

Доказательство этого свойства основывается на выполнении следующих равенств $A \cup \overline{A} = U$ и $A \cap \overline{A} = V$ и на использовании свойств 2 и 4:

$$1 = P(U) = P(A \cup \overline{A}) = P(A) + P(\overline{A}).$$

**Практика**

Предположим, что имеется совокупность некоторых элементов, количество которых равно $N$. Каждый элемент совокупности обладает одним из двух контролируемых признаков. Такие ситуации встречаются, например: *а)* при контроле качества изделий, каждое из которых является или «хорошим», или - «бракованным»; *б)* при проверке приобретённых лотерейных билетов, каждый из которых является либо «выигрышным», либо «невыигрышным»; *в)* при анализе результатов экзаменационной сессии в студенческой группе, каждый из студентов имеет либо только отличные и хорошие оценки, либо у него есть ещё и другие оценки. Все подобные ситуации описываются следующей схемой, называемой «урновой схемой».

Рассмотрим пример, объясняющий термин «урновая схема».

В урне находится $N$ шаров, из них - $M$ штук белого цвета и $N$-$M$ штук чёрного цвета. Наудачу из урны извлекается $k$ шаров. Определить вероятность того, что среди извлечённых шаров будет $l$ шаров белого цвета.

Испытание, как было записано нами ранее, это комплекс <u>условий</u> и <u>действий</u>, направленный на достижение какой-то цели.

В рассматриваемом примере комплекс <u>условий</u> описан достаточно полно. Разве что можно добавить, что экспериментатор не видит и не ощущает цвет шара, который он взял в урне для извлечения.

Комплекс <u>действий</u> описан не полностью, что не позволяет чётко определить элементарный исход и подсчитать количество элементарных исходов.

Во-первых, в условии проведения опыта должно быть указано: каждый извлечённый шар перед следующим извлечением *возвращается обратно в урну после фиксации его цвета*, или - извлечённый шар перед следующим извлечением *обратно в урну не возвращаются.* Будем считать, что шары извлекаются *без возвращения*, то есть все извлекаемые шары выкладываются рядом с урной, из которой они извлекались.

Во-вторых, так как нас интересует только количество извлечённых шаров белого цвета и нам неважно, в каком порядке они появлялись при извлечениях, то получающиеся наборы $k$ штук шаров будем называть *неупорядоченными.*

После сделанных уточнений мы можем дать чёткое название элементарного исхода: $\omega_i$ - *набор, состоящей из $k$ штук шаров*.

При этом нас не интересует: количество шаров каждого цвета в этом наборе, и в какой последовательности появлялись шары белого и чёрного цвета.

При сделанных уточнениях количество возможных элементарных исходов, то есть – число элементов, составляющих множество элементарных исходов, будет равно: $|\boldsymbol{\Omega}| = \boldsymbol{n} = \boldsymbol{C}_N^k$.

Ясно, что это число удовлетворяет первому условию применения классического определения вероятности. Так как любой из наборов ***k*** шаров, организовывается произвольно, без каких либо предпочтений на число белых или чёрных шаров в этом наборе, то все элементарные исходы будут *равновозможными*, то есть все элементарные исходы имеют равные шансы на реализацию при однократном проведении опыта или испытания.

Следовательно и второе условие применения классического определения вероятности – выполняется.

Запишем теперь название случайного события, наступление которого интересует авторов задачи:

Случайное событие ***A*** – в получившемся наборе ***k*** шаров имеется ***l*** шаров белого цвета и ***k-l*** шаров чёрного цвета.

Подсчитаем теперь количество элементарных исходов из общего числа возможных исходов, реализация которых благоприятствует наступлению случайного событие ***A***, то есть количество элементов, составляющих подмножество ***A*** во множестве **Ω**:

Организуя набор, содержащий ***l*** шаров белого цвета и ***k-l*** шаров чёрного цвета, мы можем: $\boldsymbol{C}_M^l$ способами из ***M*** шаров белого цвета без возвращения и в произвольном порядке взять ***l*** шаров белого цвета и дополнить этот набор ***k-l*** шарами чёрного цвета из ***N-M*** шаров чёрного цвета. Это сделать мы можем $\boldsymbol{C}_{N-M}^{k-l}$ способами. Таким образом, количество элементарных исходов, благоприятствующих наступлению случайного события ***A***, будет равно: $|\boldsymbol{A}| = \boldsymbol{m} = \boldsymbol{C}_M^l \cdot \boldsymbol{C}_{N-M}^{k-l}$.

Теперь можно определить вероятность наступления случайного события ***A***:

$$\boldsymbol{P}(\boldsymbol{A}) = \frac{\boldsymbol{m}}{\boldsymbol{n}} = \frac{\boldsymbol{C}_M^l \cdot \boldsymbol{C}_{N-M}^{k-l}}{\boldsymbol{C}_N^k}.$$

Ясно, что числа ***k***, ***l*** и ***k-l*** должны быть не больше, чем числа ***N***, ***M*** и ***N-M***, соответственно.

Набор вероятностей $\left\{ \boldsymbol{P}(\boldsymbol{A}_l) = \dfrac{\boldsymbol{C}_M^l \cdot \boldsymbol{C}_{N-M}^{k-l}}{\boldsymbol{C}_N^k} \right\}$, вычисленный для всех возможных значений ***l***, называется *гипергеометрическим распределением вероятностей.*

## §5.I Условная вероятность. Зависимые и независимые события

## Теория

В примере 3*а)* мы записали случайные события: событие ***A*** – *появился «король»* и событие ***B*** – *появилась «картинка»*. Была вычислена $\boldsymbol{P}(\boldsymbol{A}) = \dfrac{1}{9}$ - вероятность наступления события ***A***. Пусть мы узнали каким-то способом, что извлекаемая карта – *«картинка»*, то есть наступило случайное событие ***B***. Но нам неизвестно: какая именно «*картинка*» извлекается из колоды карт.

Нас интересует вероятность $\boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{B}\right)$ - вероятность наступления события $\boldsymbol{A}$, при *условии*, что событие $\boldsymbol{B}$ *уже наступило*. Такую вероятность мы будем называть ***условной вероятностью***.

Так как событие $\boldsymbol{B}$ уже наступило, то это означает, что при извлечении одной карты уже реализовался один из исходов благоприятствующих событию $\boldsymbol{B}$. Значит, для наступления случайного события $\boldsymbol{A}$ надо, чтобы из всех исходов, благоприятствующих событию $\boldsymbol{B}$, реализовался исход, который благоприятствует его наступлению. Так как среди 12 «*картинок*» имеется 4 «*короля*», то $\boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{B}\right)=\frac{4}{12}=\frac{1}{3}$.

Информация о наступлении события $\boldsymbol{B}$ в три раза «увеличила» вероятность наступления события $\boldsymbol{A}$.

<u>Определение.</u> ***Условной вероятностью*** наступления случайного события $\boldsymbol{A}$ при условии, что случайное событие $\boldsymbol{B}$ уже наступило, называется вероятность $\boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{B}\right)$, вычисляемая по формуле: $\boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{B}\right)\overset{def}{=}\frac{\boldsymbol{P}(\boldsymbol{A}\cap\boldsymbol{B})}{\boldsymbol{P}(\boldsymbol{B})}$.

***Комментарий***

Формула, по которой определяется значение условной вероятности, вводится по определению. Это не теорема, которую можно доказать, зная конкретный разумный способ введения вероятностной функции.

Покажем логичность этого определения в случае, когда вероятностная функция вводится на множестве, имеющем конечное число элементарных исходов, и которые обладают свойством равновозможности на реализацию при проведении испытания.

Пусть наступлению случайных событий $\boldsymbol{A}$, $\boldsymbol{B}$ и $\boldsymbol{A}\cap\boldsymbol{B}$ благоприятствуют, соответственно, $\boldsymbol{m}$, $\boldsymbol{k}$ и $\boldsymbol{l}$ элементарных исходов. Случайное событие $\boldsymbol{B}$ - наступило, то есть реализовался один из $\boldsymbol{k}$ элементарных исходов. Среди этих исходов имеется $\boldsymbol{l}$ исходов, которые благоприятствуют наступлению случайного события $\boldsymbol{A}$. Так как все исходы равновозможные, то вероятность наступления события $\boldsymbol{A}$, при условии, что событие $\boldsymbol{B}$ уже наступило, будет равна: $\boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{B}\right)=\frac{\boldsymbol{l}}{\boldsymbol{k}}$. Дальнейшие преобразования этой вероятности – очевидны:

$$\boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{B}\right)=\frac{\boldsymbol{l}}{\boldsymbol{k}}:\frac{\boldsymbol{n}}{\boldsymbol{n}}=\frac{\boldsymbol{l}}{\boldsymbol{n}}:\frac{\boldsymbol{k}}{\boldsymbol{n}}=\frac{\boldsymbol{P}(\boldsymbol{A}\cap\boldsymbol{B})}{\boldsymbol{P}(\boldsymbol{B})}.$$

Из определения условной вероятности следует, что после наступления события $\boldsymbol{B}$, множество исходов ему благоприятствующих становится множеством возможных исходов, при реализации которых анализируется возможность наступления других событий. Обозначим $\boldsymbol{B}=\boldsymbol{\Omega}_{\boldsymbol{B}}$, тогда алгебра $\mathcal{A}_{\boldsymbol{B}}$ - алгебра случайных событий - подмножеств множества $\boldsymbol{\Omega}$ вида $\boldsymbol{A}\cap\boldsymbol{B}$, где $\boldsymbol{A}$ –элементы алгебры $\mathcal{A}$. Следовательно, условная вероятность это вероятностная функция $\boldsymbol{P_B}$ определяемая на измеримом пространстве $<\Omega_B, \mathcal{A}_B>$.

Так как операция пересечения множеств – коммутативна, то условная вероятность наступления случайного события $\boldsymbol{B}$ при условии, что случайное событие $\boldsymbol{A}$ уже наступило, называется вероятность, вычисляемая по формуле: $P\left(B/A\right) = \dfrac{P(A \cap B)}{P(A)}$.

Из определения условной вероятности получаем два варианта формулы для вычисления *вероятности наступления произведения случайных событий*:

$$P(A \cap B) = P(A) \cdot P\left(B/A\right) = P(B) \cdot P\left(A/B\right).$$

## Теория

<u>Определение.</u> Случайные события $\boldsymbol{A}$ и $\boldsymbol{B}$ называются ***независимыми***, если условная вероятность наступления любого из них равна безусловной вероятности его наступления: $P\left(A/B\right) = P(A)$ или $P\left(B/A\right) = P(B)$.

Очевидно, что если выполняется одно из этих условий, то выполняется и другое условие, а так же будут справедливы равенства: $P\left(\overline{A}/B\right) = P(\overline{A})$ и $P\left(\overline{B}/A\right) = P(\overline{B})$.

Значит, если события $\boldsymbol{A}$ и $\boldsymbol{B}$ независимые, то вероятность наступления произведения этих событий равна произведению вероятностей наступления этих событий: $P(A \cap B) = P(A) \cdot P(B)$.

### *Комментарий*

Случайные события $\boldsymbol{A}$ и $\boldsymbol{B}$ могут быть совместными или несовместными, зависимыми или независимыми.

Совместность или несовместность событий мы проверяем при определении вероятности суммы событий $P(A \cup B)$ и принимаем решение, проанализировав, могут, или нет, события наступить одновременно, то есть: есть, или нет, у них общие элементарные исходы.

Зависимость или независимость событий при определении вероятности произведения событий мы проверяем *только* путём сравнения чисел $P\left(A/B\right)$ и $P(A)$, или $P\left(B/A\right)$ и $P(B)$.

## Практика

<u>Пример</u>. Из множества натуральных чисел наудачу выбирается число. Обозначим случайные события: $\boldsymbol{A}$ - *выбранное число делится на два*; $\boldsymbol{B}$ - *выбранное число делится на три*; $\boldsymbol{C}$ - *выбранное число делится на четыре.* Определим вероятности следующих случайных событий $A \cap B$ - *выбранное число делится на два и на три*; $A \cap C$ - *выбранное число делится на два и на четыре*; $B \cap C$ - *выбранное число делится на три и на четыре.*

Сначала надо определить вероятности случайных событий $\boldsymbol{A}$, $\boldsymbol{B}$ и $\boldsymbol{C}$.

К настоящему моменту мы знаем только классическое определение вероятности, основанное на конечности числа и равновозможности элементов множества элементарных исходов. Определим элементарный исход $\omega_i$ - *выбранное натуральное число*. Так как и множество натуральных чисел, и множество чётных натуральных чисел - бесконечные и эквивалентные множества (они – счётные множества), то $|\Omega| = \infty$ и $|A| = \infty$. Классическое определение вероятности при таком варианте выделения элементарных исходах - неприменимо.

Рассмотрим такой вариант. При определении: чётное или нет выбранное число, мы делим его на два и анализируем получившийся от деления остаток. Поэтому при определении вероятности наступления события $\boldsymbol{A}$ определим элементарные исходы как возможные остатки от деления выбранного числа на два: $\omega_1$- *остаток равен нулю* и $\omega_2$- *остаток равен единице*. Здесь $\Omega = \{\omega_1, \omega_2\}$ и $|\Omega| = n = 2$, элементарные исходы – равновозможные. Наступлению случайного события $\boldsymbol{A}$ благоприятствует элементарный исход $\omega_1$, то есть $|A| = m = 1$. Применяя классическое определение вероятности, получаем: $P(A) = \dfrac{m}{n} = \dfrac{1}{2}$. Аналогично, для определения вероятности наступления события $\boldsymbol{B}$ определим элементарные исходы как возможные остатки от деления выбранного числа на три: $\omega_1$- *остаток равен нулю*, $\omega_2$- *остаток равен единице*, $\omega_3$- *остаток равен двум*. Теперь $\Omega = \{\omega_1, \omega_2, \omega_3\}$ и $|\Omega| = n = 3$, элементарные исходы – равновозможные. Наступлению случайного события $\boldsymbol{B}$ благоприятствует элементарный исход $\omega_1$, то есть $|A| = m = 1$. Применяя классическое определение вероятности, получаем: $P(B) = \dfrac{m}{n} = \dfrac{1}{3}$. Ясно, что вероятность наступления случайного события $\boldsymbol{C}$ будет равна $P(C) = \dfrac{1}{4}$.

Определим вероятность наступления произведения случайных событий $\boldsymbol{A}$ и $\boldsymbol{B}$, то есть вероятность наступления события $A \cap B$. Сначала определим условную вероятность $P\left(B/A\right)$ - вероятность того, что *выбранное чётное число делится на три*. Ясно, что при делении произвольного чётного числа на три возможны остатки равные: нулю, единице и трём. Эти остатки, то есть элементарные исходы, – равновозможные, следовательно $P\left(B/A\right) = \dfrac{1}{3}$. Так как $P\left(B/A\right) = P(B)$, то события $\boldsymbol{A}$ и $\boldsymbol{B}$ – независимые, а поэтому $P(A \cap B) = P(A) \cdot P(B) = \dfrac{1}{2} \cdot \dfrac{1}{3} = \dfrac{1}{6}$.

Определим условную вероятность $P\left(C/A\right)$ - вероятность того, что *выбранное чётное число делится на четыре*. При делении произвольного чётного числа на четыре остатки могут быть равными нулю или двум. Эти остатки, то есть элементарные исходы, – равновозможные, следовательно $P\left(C/_{A}\right)=\frac{1}{2}$.

Так как $P\left(C/_{A}\right)\neq P(C)$, то события $A$ и $C$ –зависимые, а поэтому $P(A\cap C)=P(A)\cdot P\left(C/_{A}\right)=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$.

Аналогично рассуждая можно убедиться, что события $B$ и $C$ – независимые, следовательно $P(B\cap C)=P(B)\cdot P(C)=\frac{1}{3}\cdot\frac{1}{4}=\frac{1}{12}$.

***Комментарий***

Наступление события $A\cap B$ означает, что наудачу выбранное натуральное число делится на шесть, наступление события $A\cap C$ означает, что четное число делится на четыре, а наступления события $B\cap C$ означает, что число делится на двенадцать. Ясно, что вероятности наступления этих событий равны соответственно: $\frac{1}{6}$, $\frac{1}{4}$ и $\frac{1}{12}$. Эти вероятности можно легко определить, не применяя формулы вероятности произведения событий и не проверяя каждый раз наличие зависимости или независимости случайных событий - сомножителей. Столь подробное решение примера сделано для иллюстрации необходимых действий при определении вероятности произведения случайных событий.

Прав был китайский мыслитель Конфуций, сказав: «Когда бросаешь камешек в воду, всегда попадаешь точно в центр круга».

**Теория-**

В случае, когда рассматриваются совокупность случайных событий, состоящая более чем из двух событий $\{A_1,A_2,...,A_n\}$ ($n$>*2*), надо различать ***попарную независимость*** событий и ***независимость в совокупности*** событий.

Определение. Случайные события $\{A_1,A_2,...,A_n\}$ называются ***попарно независимыми***, если для любой пары индексов $i$ и $k$ будет выполняться: $P(A_i\cap A_k)=P(A_i)\cdot P(A_k)$.

Определение. Случайные события $\{A_1,A_2,...,A_n\}$ называются ***независимыми в совокупности***, если для любого числа $(k\leq n)$ и любого набора индексов $\{i_1,i_2,...,i_k\}$ будет выполняться: $P(A_{i_1}\cap A_{i_2}\cap...\cap A_{i_k})=P(A_{i_1})\cdot P(A_{i_2})\cdot...\cdot P(A_{i_k})$.

Ясно, что из независимости в совокупности следует парная независимость. Для этого во втором определении достаточно положить, что $k$=2.

Обратное не всегда верно, что иллюстрирует следующий пример.

Пример. Пусть множество $\Omega$ состоит из четырёх равновозможных элементарных исходов $\Omega=\{\omega_0,\omega_1,\omega_2,\omega_3\}$. Определим три случайных события: $A=\{\omega_0,\omega_1\}$, $B=\{\omega_0,\omega_2\}$ и $C=\{\omega_0,\omega_3\}$.

Применяя классическое определение вероятности, получаем:

$$P(A)=P(B)=P(C)=\frac{1}{2},\;\; P(A\cap B)=P(A\cap C)=P(B\cap C)=\frac{1}{4},\;\; P(A\cap B\cap C)=\frac{1}{4}.$$

События $A$, $B$ и $C$ будут попарно независимыми, так как для всех возможных пар событий выполняется определение парной независимости. Например, для пары событий $A$ и $B$ имеем: $P(A\cap B)=\frac{1}{4}=\left(\frac{1}{2}\right)^2=P(A)\cdot P(B)$. В то же время определение независимости в совокупности не выполняется:

$$P(A\cap B\cap C)=\frac{1}{4}\neq\frac{1}{8}=\left(\frac{1}{2}\right)^3=P(A)\cdot P(B)\cdot P(C).$$

## §6.I Повторные независимые испытания. Схема Бернулли. Биномиальная схема

### *Комментарий*

В этом параграфе мы рассмотрим вторую, наиболее часто встречающуюся модель испытаний, которая имеет три варианта названий, написанных в заглавии этого параграфа. Пример 2, записанный нами при знакомстве с основными понятиями теории вероятностей, является примером такой модели испытаний. Познакомившись подробней с этой моделью, мы реализуем второй *разумный* способ установления соответствия $\omega_i \Rightarrow p(\omega_i)$, то есть введения вероятностной функции $P$.

### Теория

Проводится серия из $n$ одинаковых испытаний. В каждом испытании может наступить либо случайное событие $A$, либо случайное событие $\overline{A}$. Причём результат любого испытания не зависит от исходов предыдущих испытаний, и не влияет на исходы последующих испытаний.

Определение. Такие серии из $n$ одинаковых испытаний будем называть ***повторными независимыми испытаниями***.

Я.Бернулли изучал такие испытания. Так как мы предполагаем в каждом испытании только два возможных результата, то они называются биномиальными испытаниями.

Записывая результаты каждого из $n$ испытаний, мы получим некоторую последовательность, состоящую из $n$ букв $A$ и $\overline{A}$. Естественно, что каждая такая последовательность является одной из возможных реализаций проведённой серии испытаний, поэтому мы можем назвать её элементарным исходом $\omega_k$, то есть $\omega_k=\{A,A,\overline{A},A,\overline{A},\overline{A},A,A,A,\overline{A}\}$. В этом примере длина серии испытаний $n$=10. В этой серии испытаний событие $A$ наступило равно $k$ раз, а событие $\overline{A}$ наступило $n$-$k$ раз. Очевидно, что мы можем записать ещё несколько элементарных исходов - последовательностей $\omega_k$, в которых буква $A$ встречается, а буква $\overline{A}$ встречается $n$-$k$ раз.

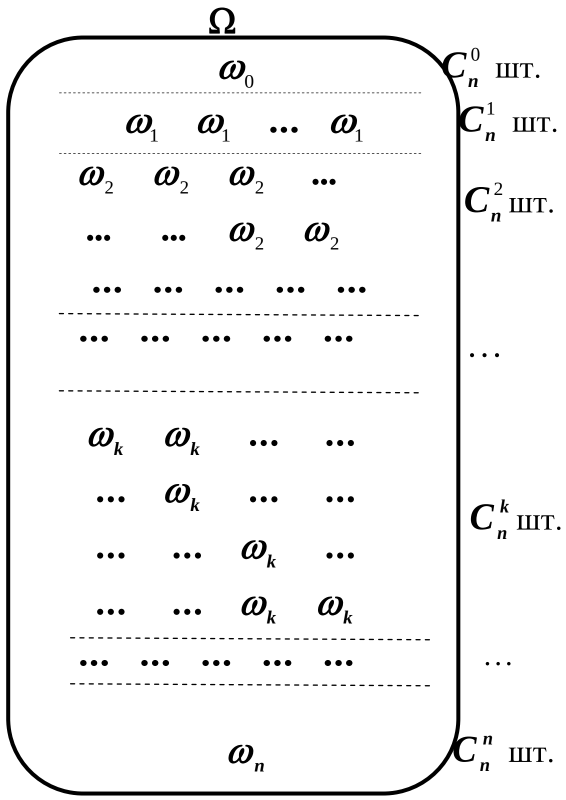
Для введения разумным способом вероятностной функции нам надо построить множество элементарных исходов $\Omega$, ознакомиться с его структурой и определить его мощность $|\Omega|$, то есть количество элементов в нём.

Упрощая форму записи каждого элементарного исхода, вместо буквы $\boldsymbol{A}$ будем писать 1, а вместо буквы $\overline{\boldsymbol{A}}$ будем писать 0, то есть записанный выше элементарный исход перепишем так: $\omega_k = \{1,1,0,1,0,0,1,1,1,0\}$. Здесь 1 встречается $\boldsymbol{k}$ раз, а 0 - $\boldsymbol{n}$-$\boldsymbol{k}$ раз.

Теперь можем сказать, что множество элементарных исходов $\boldsymbol{\Omega}$ можно рассматривать как множество двоичных последовательностей, состоящих из $\boldsymbol{n}$ элементов.

Будем последовательно, «шагами» перебирать элементарные исходы, увеличивая в записи двоичных последовательностей количество единиц от 0 до $\boldsymbol{n}$. Таких «шагов» у нас будет $\boldsymbol{n}$+1.

Сначала, «нулевой шаг», выпишем единственный элементарный исход, в котором единица встречается ноль раз, то есть $\omega_0 = \{0,0,0,.....0\}$. Затем, «первый шаг», выпишем все элементарные исходы, в которых единица встречается один раз: $\omega_1 = \{1,0,0,.....0\}$, $\omega_1 = \{0,1,.....0\}$,….., $\omega_1 = \{0,0,0,.....1\}$. Таких исходов будет $\boldsymbol{n} = \boldsymbol{C}_{\boldsymbol{n}}^1$ штук. На следующем, на «втором шаге» выпишем все элементарные исходы, в которых единица встречается два раза: $\omega_2 = \{1,1,0,.....0\}$, $\omega_2 = \{1,0,1,.....0\}$,….., $\omega_2 = \{0,0,.....1,1\}$. Таких исходов будет $\dfrac{\boldsymbol{n}\cdot(\boldsymbol{n}-1)}{2!} = \boldsymbol{C}_{\boldsymbol{n}}^2$ штук. На «$\boldsymbol{k}$-ом шаге» выпишем все элементарные исходы, в которых единица встречается $\boldsymbol{k}$ раз. Ясно, что таких исходов будет $\boldsymbol{C}_{\boldsymbol{n}}^{k}$ штук. И, наконец, на последнем, « $\boldsymbol{n}$-ом шаге» выпишем единственный элементарный исход $\omega_{\boldsymbol{n}}$, в котором единица встречается $\boldsymbol{n}$ раз.




Все элементарные исходы, составляющие множество $\boldsymbol{\Omega}$ мы разделили на $\boldsymbol{n}$+1 группу, каждую из которых составляют исходы, имеющие одинаковые количества единиц и нулей. Количество всех элементарных исходов, то есть мощность множества $\Omega$, равно: $|\Omega| = \sum_{k=0}^{n} C_n^k = 2^n$. Алгебра $\mathcal{A}^*$ всех случайных событий состоит из всевозможных подмножеств множества $\Omega$. Её мощность равна: $|\mathcal{A}^*| = 2^{|\Omega|}$.

Введём вероятностную функцию $\boldsymbol{P}$ на измеримом пространстве $<\Omega, \mathcal{A}>$.

Каждый элементарный исход $\omega_k$, который реализуется при проведении $\boldsymbol{n}$ одинаковых испытаний, мы можем рассматривать как произведение $\boldsymbol{n}$ событий:

$$\omega_k = A \cap A \cap \overline{A} \cap A \cap \overline{A} \cap \overline{A} \cap A \cap A \cap A \cap \overline{A}.$$

Из определения повторных испытаний следует, случайные события - сомножители $A$ и $\overline{A}$ - независимые события.

Согласно определению вероятности произведения независимых событий, мы можем записать вероятность реализации элементарного исхода $\omega_k$ как произведение вероятностей $\boldsymbol{n}$ событий:

$$P(\omega_k) = P(A)\cdot P(A)\cdot P(\overline{A})\cdot P(A)\cdot P(\overline{A})\cdot P(\overline{A})\cdot P(A)\cdot P(A)\cdot P(A)\cdot P(\overline{A}).$$

Пусть вероятность наступления события $A$ в каждом из $\boldsymbol{n}$ испытаний равна $P(A) = p$, тогда $P(\overline{A}) = 1 - p = q$. Следовательно, вероятность реализации выписанного элементарного исхода $\omega_k$ будет равна: $P(\omega_k) = p^k \cdot q^{n-k}$.

Теперь очевидно, что будет *разумно* каждому элементарному исходу поставить в соответствие эту вероятность, то есть: $\omega_k \Rightarrow p(\omega_k) = p^k \cdot q^{n-k}$.

Выполнение условия нормировки проверим, используя ассоциативный закон арифметики и разложение бинома Ньютона:

$$\sum_{\omega_k \in \Omega} p(\omega_k) = \sum_{k=0}^{n} C_n^k \left(p^k \cdot q^{n-k}\right) = (p+q)^n = 1.$$

### *Комментарий*

Алгебра случайных событий, которые могут наступить при проведении повторных испытаний, состоит из $|\mathcal{A}^*| = 2^{|\Omega|}$ элементов. Если мы монету подбрасываем три раза (**n**=3), и у нас элементарный исход – последовательность, состоящая из трёх букв «*г*» и «*р*», то при таких повторных независимых испытаниях множество элементарных исходов $\Omega$ состоит из $2^3$ =8 элементов, а алгебра случайных событий состоит из $2^8$ =256 элементов. При длине серии испытаний равной **n**=10 элементарных исходов будет $2^{10}$=1024. Число случайных событий, которые могут наступить при десяти подбрасываниях монеты равно $2^{1024}$.

Ясно, что при проведении десяти подбрасываний монеты экспериментатора, скорее всего, будет интересовать только количество выпадений «герба», а не какие-то конкретные комбинации единиц и нулей в последовательности, состоящей из десяти единиц и нулей.

### Теория

Обозначим $A_k$ - случайное событие: «*в серии из* ***n*** *испытаний случайное событие* ***A*** *наступило ровно* ***k*** *раз*». Нас при этом не интересует: в какой последовательности чередовались события $\boldsymbol{A}$ и $\overline{\boldsymbol{A}}$.

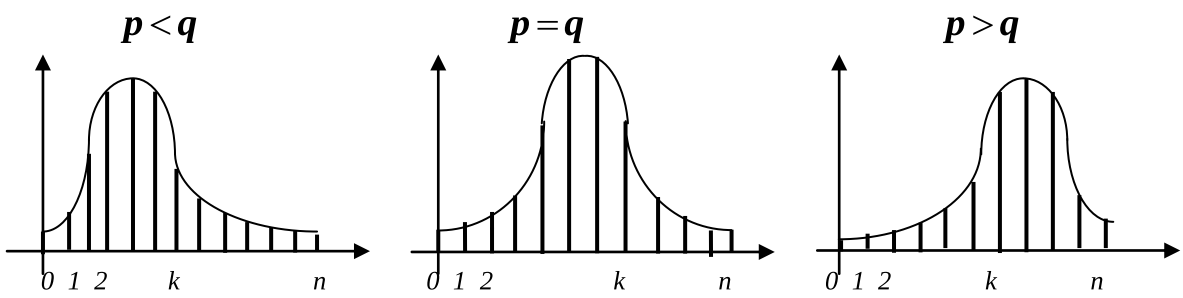
Событию $\boldsymbol{A}_k$ будут благоприятствовать все элементарные исходы $\boldsymbol{\omega}_k$, составляющие $\boldsymbol{k}$-тую группу множества элементарных исходов $\boldsymbol{\Omega}$. Так как для $\forall \boldsymbol{A} \in \mathcal{A}$ его вероятность равна $\boldsymbol{P}(\boldsymbol{A}) = \sum\limits_{\omega_i \in A} \boldsymbol{p}(\boldsymbol{\omega}_i)$, то для случайного события $\boldsymbol{A}_k$ получаем: $\boldsymbol{P}(\boldsymbol{A}_k) = \sum\limits_{\omega_k \in A_k} \boldsymbol{p}^k \cdot \boldsymbol{q}^{n-k} = \boldsymbol{C}_n^k \boldsymbol{p}^k \boldsymbol{q}^{n-k}$.

Формулу $\boldsymbol{P}(\boldsymbol{A}_k) = \boldsymbol{C}_n^k \boldsymbol{p}^k \boldsymbol{q}^{n-k}$ называют ***формулой Бернулли***.

Определение. Набор вероятностей $\{\boldsymbol{P}(\boldsymbol{A}_k)\}$, где $\boldsymbol{k} = 0,1,\ldots,\boldsymbol{n}$, называется ***биномиальным распределением вероятностей***.

Ясно, что $\sum\limits_{k=0}^{n} \boldsymbol{P}(\boldsymbol{A}_k) = 1$, то есть в результате проведения $\boldsymbol{n}$ повторных испытаний обязательно наступит одно, только одно из случайных событий $\boldsymbol{A}_k$.

Биномиальному распределению, для разных отношений между значениями $\boldsymbol{p}$ и $\boldsymbol{q}$, можно дать геометрическую интерпретацию.




## Практика

Решим два примера, в которых рассматриваются повторные независимые испытания.

Пример 1. Проводится серия из $\boldsymbol{n}$ повторных испытаний. В каждом из них случайное событие $\boldsymbol{A}$ может наступить с вероятностью $\boldsymbol{p}$. Определить вероятность того, что событие $\boldsymbol{A}$ наступит в этой серии испытаний хотя бы один раз.

Пусть случайное событие $\boldsymbol{B}$ – *при проведении* $\boldsymbol{n}$ *испытаний событие* $\boldsymbol{A}$ *наступило хотя бы один раз.* Ясно, что случайное событие $\boldsymbol{B}$ будет суммой $\boldsymbol{n}$ событий $\boldsymbol{A}_k$, где *k=1,2,...,n*: $\boldsymbol{B} = \bigcup\limits_{k=1}^{n} \boldsymbol{A}_k$. А так как события-слагаемые несовместные события, то $\boldsymbol{P}(\boldsymbol{B}) = \sum\limits_{k=1}^{n} \boldsymbol{P}(\boldsymbol{A}_k)$, где значение каждого слагаемого вычисляют по формуле Бернулли $\boldsymbol{P}(\boldsymbol{A}_k) = \boldsymbol{C}_n^k \boldsymbol{p}^k \boldsymbol{q}^{n-k}$.

Арифметические вычисления каждого слагаемого, а потом и их суммы, при больших значениях $\boldsymbol{n}$ могут потребовать больших затрат времени. Рассмотрим другой вариант решения этого примера. Сформулируем противоположное случайному событию $\boldsymbol{B}$ случайное событие $\overline{\boldsymbol{B}}$ – *при проведении* $\boldsymbol{n}$ *испытаний* *событие* $\boldsymbol{A}$ *ни разу не наступит.* Ясно, что случайное событие $\overline{\boldsymbol{B}}$ наступит только в том случае, если все $\boldsymbol{n}$ раз наступит событие $\overline{\boldsymbol{A}}$, то есть $\overline{\boldsymbol{B}} = \boldsymbol{A}_0$. А тогда $\boldsymbol{P}(\boldsymbol{B}) = 1 - \boldsymbol{P}(\overline{\boldsymbol{B}}) = 1 - \boldsymbol{q}^n$, то есть $\boldsymbol{P}(\boldsymbol{B}) = 1 - (1 - \boldsymbol{p})^n$.

Пример 2. Сколько нужно провести повторных независимых испытаний, чтобы с уверенностью не меньшей чем $\boldsymbol{P}$ утверждать, что случайное событие $\boldsymbol{A}$ наступит хотя бы один раз?

В первом примере мы, зная длину серии повторных испытаний $\boldsymbol{n}$, определили степень уверенности в том, что случайное событие $\boldsymbol{A}$ наступит хотя бы один раз. Теперь, задавшись конкретным значением (числом $\boldsymbol{P}$) степени уверенности в том, что случайное событие $\boldsymbol{A}$ наступит хотя бы один раз, мы должны определить необходимую длину серии повторных испытаний $\boldsymbol{n}$.

Согласно условию задачи должно выполняться неравенство:

$$\boldsymbol{P}(\boldsymbol{B}) = 1 - (1 - \boldsymbol{p})^n \ge \boldsymbol{P}.$$

Решая его относительно числа $\boldsymbol{n}$, получим: $\boldsymbol{n} \ge \dfrac{\boldsymbol{ln}(1 - \boldsymbol{P})}{\boldsymbol{ln}(1 - \boldsymbol{p})}$.

Если подбрасывается монета, то при $\boldsymbol{n} \ge 4$ подбрасываниях мы с уверенностью $\boldsymbol{P}$=0,9 можем утверждать, что хотя бы один раз наступит случайное событие $\boldsymbol{A}$ – *появился* «*герб*». То есть в 9 из 10 серий по 4 подбрасывания мы хотя бы один раз увидим выпадение «*герба*». Если мы увеличим степень уверенности до $\boldsymbol{P}$=0,99, длина серии повторных испытаний должна быть увеличена до $\boldsymbol{n} \ge 7$.

### ***Комментарий***

Дадим сначала следующее определение.

Определение. Случайные события $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_m$ образуют ***полную группу событий*** (ПГС), если они попарно несовместны, то есть для любой пары индексов $\boldsymbol{i}$ и $\boldsymbol{j}$ справедливо $\boldsymbol{A}_i \cap \boldsymbol{A}_j = \varnothing$, а их сумма их является достоверным событием: $\bigcup_{i=1}^{m} \boldsymbol{A}_i = \boldsymbol{\Omega}$.

Рассмотренную модель повторных независимых испытаний с двумя возможными исходами каждого испытания можно обобщить, предполагая, что в результате каждого из $\boldsymbol{n}$ испытаний могут наступить не одно из двух, а одно из трёх (или из большего числа) случайных событий, образующих полную группу событий. Вероятности наступлений этих событий в любом испытании постоянны и не зависят от номера испытания. Такая модель повторных испытаний называется *полиномиальной схемой.*

Например. Пусть в результате одного испытания может наступить только одно из трёх попарно несовместных событий $\boldsymbol{A}$, $\boldsymbol{B}$ или $\boldsymbol{C}$. Вероятности наступлений этих событий равны, соответственно: $\boldsymbol{p}$, $\boldsymbol{q}$ и $\boldsymbol{r}$. Ясно, что $\boldsymbol{p} + \boldsymbol{q} + \boldsymbol{r} = 1$. В этом случае элементарный исход можно представить как троичную последовательность $\boldsymbol{n}$ штук 0,1 и 2, считая, что $\boldsymbol{A} \leftrightarrow 0$, $\boldsymbol{B} \leftrightarrow 1$, $\boldsymbol{C} \leftrightarrow 2$,: $\boldsymbol{\omega}_{k,m} = \{0,2,2,1,0,0,1,2,0,1\}$.

Здесь $k$=4 – количество нулей в этой последовательности, $m$=3 –количество единиц, тогда $n$-($k$+$m$)=3 - количество двоек. Аналогично тому, как это сделано в биномиальной модели, строится множество элементарных исходов $\Omega$, которое состоит из $3^n$ элементов.

Вероятностная функция вводится назначением соответствия $\omega_{k,m} \Rightarrow p(\omega_{k,m}) = p^k q^m r^{n-(k+m)}$. Рассматриваются случайные события $A_{k,m}$, где $k$ и $m$ $\geq 0$, а $k + m \leq n$. Каждому случайному событию $A_{k,m}$ благоприятствуют $C_n^k \cdot C_{n-k}^m$ равновозможных между собой элементарных исходов.

Вероятность наступления случайного события $A_{k,m}$ будет равна: $P(A_{k,m}) = \dfrac{n!}{k! \cdot m! \cdot (n-(k+m))!} p^k q^m r^{n-(k+m)}$.

## §6.I Испытания до первого положительного исхода

### *Комментарий*

Испытание, описанное в примере 5 перечня простейших испытаний, отличается от других тем, что множество элементарных исходов $\Omega$ содержит бесконечное число элементов. Покажем, что здесь тоже можно ввести вероятностную функцию $P$. Это будет третий пример *введения вероятностной функции разумным способом.*

### Теория

Проводятся одинаковые независимые испытания, в каждом из которых может наступить или случайное событие $A$, или случайное событие $\overline{A}$. Вероятность наступления события $A$ в любом испытании равна $P(A) = p$. Испытания прекращаются после первого наступления случайного события $A$. Благодаря накопленному опыту из предыдущих примеров, мы можем записать элементарный исход в виде последовательности $\omega_k = \{0,0,\ldots,0,1\}$, в которой будет $k$-1 нулей, что соответствуют случаю, когда было проведено $k$ испытаний, в которых $k$-1 раз наступало событие $\overline{A}$ и $k$-тое испытание закончилось наступлением события $A$. Так как число проведённых испытаний может быть любым натуральным числом, то множество элементарных исходов будет счётным множеством, мощность которого равна $|\Omega| = a$.

Любой элементарный исход $\omega_k$ можно рассматривать как событие, являющееся произведением $k$ случайных событий: $\omega_k = \overline{A} \cap \overline{A} \cap \ldots \cap \overline{A} \cap A$. Так как вероятность наступления события $A$ в любом испытании не зависит от номера испытания, то события-сомножители будут независимыми событиями, а поэтому вероятность $P(\omega_k)$ будет равна произведению вероятностей событий-сомножителей: $P(\omega_k) = P(\overline{A}) \cdot P(\overline{A}) \cdot \ldots \cdot P(\overline{A}) \cdot P(A) = q^{k-1} p$. Очевидно, что будет *разумно* поставить в соответствие каждому элементарному исходу $\omega_k$ получившуюся вероятность $P(\omega_k)$, то есть - положительное число $q^{k-1} p$, или: $\omega_k \Rightarrow p(\omega_k) = q^{k-1} p$.

Проверим выполнение условия нормировки: $\sum\limits_{\omega_k\in\Omega} p(\omega_k)=\sum\limits_{k=1}^{\infty} q^{k-1}p=\frac{p}{1-q}=1.$

Обозначим случайное событие $A_k$ - *в первый раз случайное событие* $A$ *наступило при проведении* $k$*-того испытания*. Ясно, что вероятность наступления такого события равна $P(A_k)=q^{k-1}p$. Так как эти вероятности являются членами бесконечно убывающей геометрической прогрессии, то набор вероятностей $\{P(A_k)\}$, где $k$=*1,2,3,.....*, называется *геометрическим распределением вероятностей.*

***Комментарий***

После изучения биномиальной схемы испытаний мы рассмотрели возможность обобщения этого типа испытаний и ввели понятие полиномиальной схемы испытаний. Аналогично можно обобщить понятие «испытания до первого положительного исхода» и рассматривать «испытания до $k$-того положительного исхода». То есть, мы проводим одинаковые независимые испытания до тех пока не зафиксируем появление события $A$ ровно $k$ раз.

Элементарный исход записывается в виде последовательности $\omega_{m+k}=\{0,0,1,0,...,1,1,0,...,0,1\}$, в которой, до записи последней $k$-той единицы, записывается последовательность, состоящая из $k$-1-ной единицы и $m$ нулей. Число $m$ может быть любым от 0 до $\infty$, то есть множество элементарных исходов $\Omega=\{\omega_{m+k}\}$ - счётное множество. Ясно, что число последовательностей, в которых будет $k$-1 единица и $m$ нулей, будет равно $C^m_{m+k-1}$, а поэтому, определяя вероятностную функцию, будет *разумно* установить соответствие: $\omega_{m+k}\Rightarrow p(\omega_{m+k})=C^m_{m+k-1}q^m p^k$.

Обозначим $A_m$ - случайное событие - $k$*-тое наступление события* $A$ *наступило при проведении* $(m+k)$*-того испытания*. Вероятность наступления этого события равна $P(A_m)=C^m_{m+k-1}q^m p^k$. Набор вероятностей $\{P(A_m)\}$, где $m$=*0,1,2,3,.....*, называется *отрицательно-биномиальным распределением вероятностей.*

## §8.I Формула полной вероятности. Формула Байеса

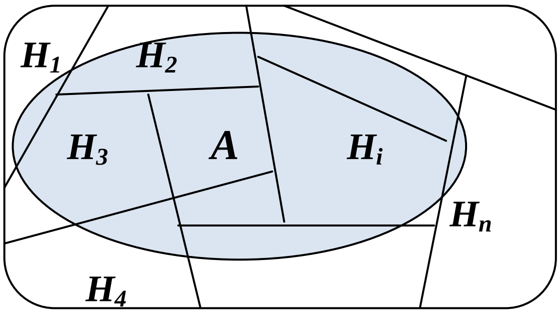
***Комментарий***

Две формулы, которые мы сейчас рассмотрим, являются следствиями свойств вероятностей суммы и произведения случайных событий. Применение этих формул при решении задач определения вероятности наступления случайного события, являющегося суммой произведений нескольких событий, позволяет быстрее получить числовой результат.

**Теория**

Пусть $\{H_i\}$, $i$=1,2,…,$n$, полная группа событий, то есть эти события попарно несовместные и сумма их - достоверное событие. События $H_i$ будем называть *гипотезами*. Случайное событие $A$ может наступить, или не наступить, с одной из этих гипотез. Определим вероятность наступления этого случайного события, то есть $P(A)$.

Событие $\boldsymbol{A}$ можно представить так: $\boldsymbol{A}=\bigcup_{i=1}^{n}\left(\boldsymbol{A}\cap\boldsymbol{H}_i\right)$. События- слагаемые в этом выражении - несовместные события, а поэтому $\boldsymbol{P}(\boldsymbol{A})=\sum_{i=1}^{n}\boldsymbol{P}\left(\boldsymbol{A}\cap\boldsymbol{H}_i\right)$.




Воспользуемся коммутативностью операции пересечения множеств и формулой вероятности произведения событий $\boldsymbol{P}(\boldsymbol{A})=\sum_{i=1}^{n}\boldsymbol{P}(\boldsymbol{H}_i)\cdot\boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{H}_i\right)$.

Полученная формула называется *формулой полной вероятности*.

## Практика

<u>Пример 1.</u> Среди $\boldsymbol{n}$ экзаменационных билетов находится $\boldsymbol{m}$ «хороших» для студента. Студент пытается решить, когда ему лучше идти на экзамен: первым, вторым, третьим или – последним, чтобы вероятность взять «хороший» билет было наибольшей.

*<u>Замечание.</u> Если бы такая наибольшая вероятность существовала, то студенты уже давно пользовались бы таким способом взять «хороший» билет.*

Последовательно рассмотрим варианты, когда студент идёт на экзамен первым, вторым, третьим. Нам здесь гораздо важнее понять: как перед проведением испытания формулируются гипотезы, и какова их роль в определении значений условных вероятностей.

Вариант **I**. Студент пошёл на экзамен первым. Случайное событие $\boldsymbol{A}$ – *студент взял «хороший» билет*. Применяя классическое определение вероятности, получим $\boldsymbol{P}(\boldsymbol{A})=\dfrac{\boldsymbol{m}}{\boldsymbol{n}}$.

Вариант **II**. Студент пошёл на экзамен вторым. Студент не знает - какой билет взял его предшественник. Поэтому он предполагает две гипотезы:

$\boldsymbol{H}_1$ *– предшественник взял «хороший» билет и* $\boldsymbol{H}_2$ *– предшественник взял «плохой» билет.* Вероятности этих гипотез соответственно равны: $\boldsymbol{P}(\boldsymbol{H}_1)=\dfrac{\boldsymbol{m}}{\boldsymbol{n}}$ и $\boldsymbol{P}(\boldsymbol{H}_2)=\dfrac{\boldsymbol{n}-\boldsymbol{m}}{\boldsymbol{n}}$. В первом случае возможность шансов, по сравнению с числом $\dfrac{\boldsymbol{m}}{\boldsymbol{n}}$, взять «хороший» билет уменьшается, а во втором – увеличивается. Поэтому: $\boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{H}_1\right)=\dfrac{\boldsymbol{m}-1}{\boldsymbol{n}-1}$ и $\boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{H}_2\right)=\dfrac{\boldsymbol{m}}{\boldsymbol{n}-1}$. Применяя формулу полной вероятности, получаем: $\boldsymbol{P}(\boldsymbol{A})=\boldsymbol{P}(\boldsymbol{H}_1)\cdot\boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{H}_1\right)+\boldsymbol{P}(\boldsymbol{H}_2)\cdot\boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{H}_2\right)$.

Вычисляем: $\boldsymbol{P}(\boldsymbol{A})=\frac{\boldsymbol{m}}{\boldsymbol{n}}\cdot\frac{\boldsymbol{m}-1}{\boldsymbol{n}-1}+\frac{\boldsymbol{n}-\boldsymbol{m}}{\boldsymbol{n}}\cdot\frac{\boldsymbol{m}}{\boldsymbol{n}-1}$. Окончательно получаем $\boldsymbol{P}(\boldsymbol{A})=\frac{\boldsymbol{m}}{\boldsymbol{n}}$, что и следовало ожидать.

Вариант **III**. Студент пошёл на экзамен третьим. Не зная, какие билеты взяли его предшественники, студент предполагает три гипотезы:

$\boldsymbol{H}_1$ – *предшественники взяли два «хороших» билета;*

$\boldsymbol{H}_2$ – *предшественники взяли один «хороший» билет и один «плохой билет»;*

$\boldsymbol{H}_3$ – *предшественники взяли два «плохих» билета*.

Определим вероятности выдвинутых гипотез и соответствующие условные вероятности:

$$\boldsymbol{P}(\boldsymbol{H}_1)=\frac{\boldsymbol{C}_{\boldsymbol{m}}^2\cdot\boldsymbol{C}_{\boldsymbol{n}-\boldsymbol{m}}^0}{\boldsymbol{C}_{\boldsymbol{n}}^2},\quad \boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{H}_1\right)=\frac{\boldsymbol{m}-2}{\boldsymbol{n}-2};$$

$$\boldsymbol{P}(\boldsymbol{H}_2)=\frac{\boldsymbol{C}_{\boldsymbol{m}}^1\cdot\boldsymbol{C}_{\boldsymbol{n}-\boldsymbol{m}}^1}{\boldsymbol{C}_{\boldsymbol{n}}^2},\quad \boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{H}_2\right)=\frac{\boldsymbol{m}-1}{\boldsymbol{n}-2};$$

$$\boldsymbol{P}(\boldsymbol{H}_3)=\frac{\boldsymbol{C}_{\boldsymbol{m}}^0\cdot\boldsymbol{C}_{\boldsymbol{n}-\boldsymbol{m}}^2}{\boldsymbol{C}_{\boldsymbol{n}}^2},\quad \boldsymbol{P}\left(\boldsymbol{A}/\boldsymbol{H}_3\right)=\frac{\boldsymbol{m}}{\boldsymbol{n}}.$$

По формуле полной вероятности получаем:

$$\boldsymbol{P}(\boldsymbol{A})=\frac{\boldsymbol{m}\cdot(\boldsymbol{m}-1)\cdot(\boldsymbol{m}-2)+2\cdot\boldsymbol{m}\cdot(\boldsymbol{n}-\boldsymbol{m})\cdot(\boldsymbol{m}-1)+(\boldsymbol{n}-\boldsymbol{m})\cdot(\boldsymbol{n}-\boldsymbol{m}-1)\cdot\boldsymbol{m}}{\boldsymbol{n}\cdot(\boldsymbol{n}-1)\cdot(\boldsymbol{n}-2)}=\frac{\boldsymbol{m}}{\boldsymbol{n}}$$

.

Следовательно, для того, чтобы вероятность $\boldsymbol{P}(\boldsymbol{A})$ взять «хороший» билет была большой, студент должен при подготовке к экзамену постараться, чтобы число $\boldsymbol{m}$ как можно меньше отличалось от числа $\boldsymbol{n}$.

### *Комментарий*

Гипотезы $\{\boldsymbol{H}_{\boldsymbol{i}}\}$, $\boldsymbol{i}$=1,2,…,$\boldsymbol{n}$, попарно несовместные события и сумма их является достоверным событием. Поэтому, если мы выписали все возможные гипотезы и определили их вероятности, то всегда должно выполняться: $\sum_{\boldsymbol{i}=1}^{\boldsymbol{n}}\boldsymbol{P}(\boldsymbol{H}_{\boldsymbol{i}})=1$. Выполнение этого равенства может быть использовано как промежуточный контроль правильности сделанных вычислений перед определением условных вероятностей и вероятности $\boldsymbol{P}(\boldsymbol{A})$ наступления случайного события $\boldsymbol{A}$.

### Теория

Для вывода формулы Байеса считаем, что мы находимся в тех же условиях, что и при выводе формулы полной вероятности, то есть $\{\boldsymbol{H}_{\boldsymbol{i}}\}$ - гипотезы и событие $\boldsymbol{A}$ может наступить, или не наступить, с одной из этих гипотез.

Так как операция пересечения множеств коммутативна, то для любого номера гипотезы $\boldsymbol{k}$ будет справедливо: $\boldsymbol{H}_{\boldsymbol{k}}\cap\boldsymbol{A}=\boldsymbol{A}\cap\boldsymbol{H}_{\boldsymbol{k}}$. Следовательно: $\boldsymbol{P}(\boldsymbol{H}_{\boldsymbol{k}}\cap\boldsymbol{A})=\boldsymbol{P}(\boldsymbol{A}\cap\boldsymbol{H}_{\boldsymbol{k}})$. Применяя формулу для вычисления вероятности произведения событий, запишем: $P(H_k)\cdot P\left(A/H_k\right)=P(A)\cdot P\left(H_k/A\right)$. Отсюда получаем: $P\left(H_k/A\right)=\dfrac{P(H_k)\cdot P\left(A/H_k\right)}{P(A)}$ или, если воспользоваться формулой полной вероятности,: $P\left(H_k/A\right)=\dfrac{P(H_k)\cdot P\left(A/H_k\right)}{\sum\limits_{i=1}^{n}P(H_i)\cdot P\left(A/H_i\right)}$ . Полученная дробь называется *формулой Байеса* или *формулой после опытной переоценки вероятностей гипотез.*

### *Комментарий*

Первым признаком необходимости применения формулы полной вероятности или формулы Байеса является понимание того, что рассматриваемое испытание проводится в два этапа. На первом этапе проведения испытания возможные предварительные действия или предположения экспериментатора мы формулируем в виде гипотез. Рассматривая второй этап проведения испытания, мы рассматриваем возможности наступления случайного события $A$ при справедливости или при реализации той, или иной, гипотезы.

Для выяснения смысла формулы Байеса проанализируем последовательность наших действий при решении задачи.

Ознакомившись с содержанием испытания как комплекса условий и действий, мы выдвигаем гипотезы, при которых может наступить, или не наступить случайное событие.

1) Гипотезы $\{H_i\}$ составляются, формулируются до проведения испытания.

2) Вероятности $\{P(H_i)\}$ называются *доопытными (априорными) вероятностями гипотез*.

3) Проводится испытание, в результате которого наступило случайное событие $A$. Но мы не знаем, с какой из гипотез оно наступило.

4) Применяя формулу Байеса, то есть, вычисляя $P\left(H_k/A\right)$, мы переоцениваем доопытную вероятность гипотезы $H_k$ с учётом информации о том, что событие $A$ уже наступило.

5) Вероятности $\left\{P\left(H_k/A_i\right)\right\}$ называются *послеопытными (апостериорными) вероятностями гипотез.*

6) Ясно, что всегда будет справедливо: $\sum\limits_{k=1}^{n}P\left(H_k/A\right)=1$.

# ГЛАВА ВТОРАЯ
# Аксиоматическое построение вероятностного пространства.
## §1.II Аксиомы А.Н.Колмогорова
### *Комментарий*

Предложенный в первой главе вариант аксиоматического введения вероятностной функции как установление для каждого элементарного исхода $\omega_i$ разумным способом соответствия $\omega_i \Rightarrow p(\omega_i)$ и вычисление вероятности наступления события $A$ как $P(A)=\sum_{\omega_i \in A} p(\omega_i)$, оказался «тупиковым». Такой вариант введения вероятностной функции позволяет рассматривать только такие вероятностные пространства $<\Omega, \mathcal{A}, P>$, в которых множество элементарных исходов $\Omega$ или имеет конечное число элементов, или оно – счётное, то есть, если его элементы можно записать в виде бесконечной последовательности.

Наивные представления XVI века, предложенные Дж.Кардано, о конечном множестве элементарных исходов, и представления XVII века о случайном событии и его вероятности уже перестали удовлетворять запросам науки.

Теория формулировалась на основании эксперимента, в котором можно было выписать все возможные элементарные исходы в виде последовательности.

Но можно придумать и рассмотреть много испытаний, в которых элементы множества $\Omega$ нельзя выписать в виде конечной или бесконечной последовательности.

Например, наудачу выбирается точка из множества точек отрезка [0;1]. Ясно, что в таком испытании любая точка $x$, принадлежащая этому отрезку, является элементарным исходом $\omega$, а множество точек отрезка [0;1] является множеством элементарных исходов: $\Omega$=[0;1]. Из курса математического анализа известно, что все точки отрезка [0;1] нельзя выписать в виде последовательности, то есть мощность множества [0;1] равна $c$ - мощности континуум. Значит, мы не сможем разумным способом каждому элементарному исходу $\omega$ поставить в соответствие положительное число $p(\omega)$. Однако, если мы обозначим случайное событие $A$ – *наудачу выбранная точка принадлежит отрезку* [0,25;0,5], то, так как все элементарные исходы равновозможные, интуитивно мы понимаем, что вероятность наступления этого события будет равна $P(A)=\frac{1}{4}$.

Паскаль потребовал, чтобы вероятность достоверного события, как степень уверенности в его наступлении, всегда была равна единице, то есть $P(\Omega)=1$. Теперь, интерпретируя распределение вероятностей как распределение на множестве элементарных исходов $\Omega$ некоторым разумным способом единичной массы, мы понимаем вероятность наступления случайного события $P(A)$ как долю единичной массы, оказавшуюся над соответствующим этому событию подмножеством $A$ множества $\Omega$.

Если $\Omega$=[0;1], то множество всех подмножеств множества $\Omega$ имеет мощность, равную $|\mathcal{A}^*|=2^{|\Omega|}=2^{a}=c$. Возникает два вопроса. Во-первых, какие именно подмножества множества $\boldsymbol{\Omega}$ можно понимать как случайные события и как их конструировать? Во-вторых, подмножество $\boldsymbol{A}$, которое мы назвали случайным событием, может иметь такую сложную конструкцию, что не представляется возможным применить определённый интеграл Римана для вычисления доли единичной массы, оказавшейся над ним. Как тогда определить вероятность $\boldsymbol{P}(\boldsymbol{A})$ наступления события $\boldsymbol{A}$?

На рубеже XIX-XX веков широта применения теории вероятностей, запросы естествознания, практики настойчиво требовали создания аксиоматики теории вероятностей. Если ранее теория формулировалась на основании эксперимента, то стало ясно, что должна быть разработана общая теория, независящая от особенностей эксперимента, а каждый конкретный эксперимент должен иллюстрировать теорию.

Предложенная в 1933 году А.Н.Колмогоровым аксиоматика теории вероятностей существенно опирается на аппарат теории множеств и теории меры.

Создание аксиоматики теории вероятностей стало возможным благодаря работам:

*а)* Георга Кантора (1845-1918), в которых доказана несчётность множества всех действительных чисел, сформулировано общее понятие мощности множества, установлено существование бесконечных множеств, имеющих разные мощности;

*б)* Эмиля Бореля (1871-1956), в которых введены понятия борелевского множества, как некоторого подмножества множества действительных чисел, и алгебры борелевских множеств;

*в)* Анри Лебега (1875-1941), в которых создана теории меры (мера Лебега) и сделано обобщение понятия интеграла (интеграл Лебега), основанное на теории меры и позволяющее интегрировать чрезвычайно широкий класс функций.

**Теория**

Следуя А.Н.Колмогорову, проведём построение вероятностного пространства, последовательно формулируя необходимые определения и аксиомы.

1) Пусть $\boldsymbol{\Omega}$ - некоторое множество элементов. Мощность $|\boldsymbol{\Omega}|$ нас не интересует.

2) Определение. Совокупность $\boldsymbol{\mathcal{A}}$ подмножеств множества $\boldsymbol{\Omega}$ называется ***σ-алгеброй***, если:

*a)* $\boldsymbol{\Omega} \in \boldsymbol{\mathcal{A}}$ , то есть множество $\boldsymbol{\Omega}$ является элементом совокупности $\boldsymbol{\mathcal{A}}$,

*b)* для любой счётной последовательности подмножеств $\{\boldsymbol{A}_i\}$, *i=1,2,3,….,* множества $\boldsymbol{\Omega}$, принадлежащей совокупности $\boldsymbol{\mathcal{A}}$, следует что и счётное объединение, и счётное пересечение этих подмножеств также являются элементами этой совокупности: $\bigcup_{i=1}^{\infty} \boldsymbol{A}_i \in \boldsymbol{\mathcal{A}}$ и $\bigcap_{i=1}^{\infty} \boldsymbol{A}_i \in \boldsymbol{\mathcal{A}}$,

*c)* если $\forall \boldsymbol{A} \in \boldsymbol{\mathcal{A}}$, то и $\overline{\boldsymbol{A}} = \boldsymbol{\Omega} \setminus \boldsymbol{A} \in \boldsymbol{\mathcal{A}}$.

Из определения следует, что $\sigma$-алгебра $\mathcal{A}$ есть класс множеств, замкнутый относительно счётного числа операций объединения, пересечения и дополнения.

3) <u>Определение</u>. Пару объектов $<\Omega, \mathcal{A}>$ будем называть ***измеримым пространством.***

4) Каждому элементу $A$ из $\sigma$-алгебры $\mathcal{A}$ поставим в соответствие неотрицательное действительное число $\mu(A)$, то есть на элементах $\sigma$-алгебры определим неотрицательную числовую функцию $\mu : \mathcal{A} \to [0;+\infty)$.

5) <u>Аксиома 1</u>. Числовая функция $\mu$ будет ***нормированной***, если $\mu(\Omega) = 1$.

6) <u>Аксиома 2</u>. Числовая функция $\mu$ будет ***конечно-аддитивной***, то есть для любых $A$ и $B$, таких, что $A \cap B = \varnothing$, следует: $\mu(A \cup B) = \mu(A) + \mu(B)$.

7) Нормированную, конечно-аддитивную функцию $\mu$ будем называть *конечной вероятностной функцией*, или конечной вероятностью.

8) <u>Аксиома 3</u>. Числовая функция $\mu$ будет ***$\sigma$-аддитивной***, если для любой счётной последовательности подмножеств $\{A_i\}$, *i=1,2,3,….,* таких, что $A_i \cap A_k = \varnothing$, $(i \neq k)$ следует: $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$.

9) Нормированную, $\sigma$-аддитивную функцию $\mu$ подмножеств $A$ будем называть ***вероятностной функцией*** $P$ и записывать $\mu(A) = P(A)$.

<u>Определение</u>. Тройка объектов $<\Omega, \mathcal{A}, P>$, где

множество $\Omega$ - множество элементарных исходов,

$\sigma$-алгебра $\mathcal{A}$ – алгебра случайных событий,

функция $P$ – вероятностная функция,

называется ***вероятностным пространством***.

Сформулировав определения и записав аксиомы, мы должны изучить свойства вероятностной функции, которые по существу являются теоремами. Эти теоремы позволят в дальнейшем решать различные выдвигаемые практикой задачи теории вероятностей.

<u>Свойство 1.</u> Вероятность невозможного события равна нулю: $P(V) = 0$.

Действительно, событие $V$ мы понимаем как пустое подмножество $\varnothing$ множества $\Omega$. При этом справедливо: $\Omega \cup \varnothing = \Omega$ и $\Omega \cap \varnothing = \varnothing$. Для доказательства $P(V) = 0$ остаётся использовать аддитивность и нормированность вероятностной функции.

<u>Свойство 2.</u> Если $A$ и $B \in \mathcal{A}$, то $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Для доказательства этого свойства представим сумму событий $A$ и $B$ как сумму трёх попарно несовместных событий:

$$A \cup B = (A \setminus (A \cap B)) \cup (A \cap B) \cup (B \setminus (A \cap B)).$$

Воспользуемся аддитивностью вероятностной функции:

$P(A \cup B) = P(A \setminus (A \cap B)) + P(A \cap B) + P(B \setminus (A \cap B))$. Значения первого и третьего слагаемых определяем из очевидных равенств:

$$P(A) = P(A \setminus (A \cap B)) + P(A \cap B) \quad \text{и} \quad P(B) = P(B \setminus (A \cap B)) + P(A \cap B).$$

Следовательно: $\boldsymbol{P}(\boldsymbol{A}\bigcup\boldsymbol{B})=\boldsymbol{P}(\boldsymbol{A})-\boldsymbol{P}(\boldsymbol{A}\bigcap\boldsymbol{B})+\boldsymbol{P}(\boldsymbol{A}\bigcap\boldsymbol{B})+\boldsymbol{P}(\boldsymbol{B})-\boldsymbol{P}(\boldsymbol{A}\bigcap\boldsymbol{B})$.

То есть: $\boldsymbol{P}(\boldsymbol{A}\bigcup\boldsymbol{B})=\boldsymbol{P}(\boldsymbol{A})+\boldsymbol{P}(\boldsymbol{B})-\boldsymbol{P}(\boldsymbol{A}\bigcap\boldsymbol{B})$.

*Следствие.* Применяя метод математической индукции, мы можем обобщить это свойство для любого конечного числа случайных событий-слагаемых:

$$\boldsymbol{P}(\boldsymbol{A})=\sum_{i}\boldsymbol{P}(\boldsymbol{A}_i)-\sum_{1=i<j=n}\boldsymbol{P}\left(\boldsymbol{A}_i\bigcap\boldsymbol{A}_j\right)+\sum_{1=i<j<k=n}\boldsymbol{P}\left(\boldsymbol{A}_i\bigcap\boldsymbol{A}_j\bigcap\boldsymbol{A}_k\right)-\ldots+(-1)^{n+1}\boldsymbol{P}\left(\bigcap_{i=1}^{n}\boldsymbol{A}_i\right).$$

<u>Свойство 3.</u> Если $\boldsymbol{A},\boldsymbol{B}\in\mathcal{A}$, и $\boldsymbol{A}\subseteq\boldsymbol{B}$, то $\boldsymbol{P}(\boldsymbol{A})\le\boldsymbol{P}(\boldsymbol{B})$.

Доказательство очевидно, для этого достаточно представить $\boldsymbol{B}$ как сумму двух несовместных событий: $\boldsymbol{B}=\boldsymbol{A}\bigcap(\boldsymbol{B}\setminus\boldsymbol{A})$.

<u>Свойство 4.</u> Свойство $\boldsymbol{\sigma}$-полуаддитивности вероятностной функции.

Если $\boldsymbol{A}_i\in\mathcal{A}$, где *i=1,2,3,…* и $\bigcup\limits_{i=1}^{\infty}\boldsymbol{A}_i=\boldsymbol{A}\in\mathcal{A}$, то $\boldsymbol{P}\left(\bigcup\limits_{i=1}^{\infty}\boldsymbol{A}_i\right)\le\sum\limits_{i=1}^{\infty}\boldsymbol{P}(\boldsymbol{A}_i)$.

Доказательство. Обозначим $\boldsymbol{B}_1=\boldsymbol{A}_1,\quad \boldsymbol{B}_2=\boldsymbol{A}_2\setminus\boldsymbol{A}_1,\quad \ldots\quad \boldsymbol{B}_n=\boldsymbol{A}_n\setminus\left(\bigcup\limits_{k=1}^{n-1}\boldsymbol{A}_k\right),\quad \ldots$.

Очевидно, что: 1) $\bigcup\limits_{i=1}^{\infty}\boldsymbol{A}_i=\bigcup\limits_{i=1}^{\infty}\boldsymbol{B}_i$ , 2) $\boldsymbol{B}_i\bigcap\boldsymbol{B}_j=\varnothing$, если $i\neq j$, и 3) $\boldsymbol{B}_i\subseteq\boldsymbol{A}_i$.

Значит: $\boldsymbol{P}\left(\bigcup\limits_{i=1}^{\infty}\boldsymbol{A}_i\right)\overset{1)}{=}\boldsymbol{P}\left(\bigcup\limits_{i=1}^{\infty}\boldsymbol{B}_i\right)\overset{2)}{=}\sum\limits_{i=1}^{\infty}\boldsymbol{P}(\boldsymbol{B}_i)\overset{3)}{\le}\sum\limits_{i}^{\infty}\boldsymbol{P}(\boldsymbol{A}_i)$, что и требовалось доказать.

<u>Свойство 5.</u> Теорема ***о непрерывности*** вероятностной функции.

Пусть $\boldsymbol{P}$ – конечно-аддитивная функция множеств, заданная на $\boldsymbol{\sigma}$-алгебре подмножеств $\mathcal{A}$ множества $\Omega$, и $\boldsymbol{P}(\Omega)=1$. Тогда следующие четыре утверждения эквивалентны:

1) функция $\boldsymbol{P}$ - $\boldsymbol{\sigma}$*-аддитивная* функция;

2) функция $\boldsymbol{P}$ – *непрерывная «сверху»* функция, то есть, если $\{\boldsymbol{A}_n\}$, *n=1,2,3,…* счётная последовательность множеств, принадлежащих $\mathcal{A}$, таких что $\boldsymbol{A}_n\subseteq\boldsymbol{A}_{n+1}$ и пусть $\boldsymbol{A}=\bigcup\limits_{n=1}^{\infty}\boldsymbol{A}_n$ , тогда $\boldsymbol{P}(\boldsymbol{A})=\lim\limits_{n\to\infty}\boldsymbol{P}(\boldsymbol{A}_n)$;

3) функция $\boldsymbol{P}$ – *непрерывная «снизу»* функция, то есть, если $\{\boldsymbol{A}_n\}$, *n=1,2,3,…* счётная последовательность множеств, принадлежащих $\mathcal{A}$, таких что $\boldsymbol{A}_n\supseteq\boldsymbol{A}_{n+1}$ и пусть $\boldsymbol{A}=\bigcap\limits_{m=1}^{\infty}\boldsymbol{A}_n$ , тогда $\boldsymbol{P}(\boldsymbol{A})=\lim\limits_{n\to\infty}\boldsymbol{P}(\boldsymbol{A}_n)$;

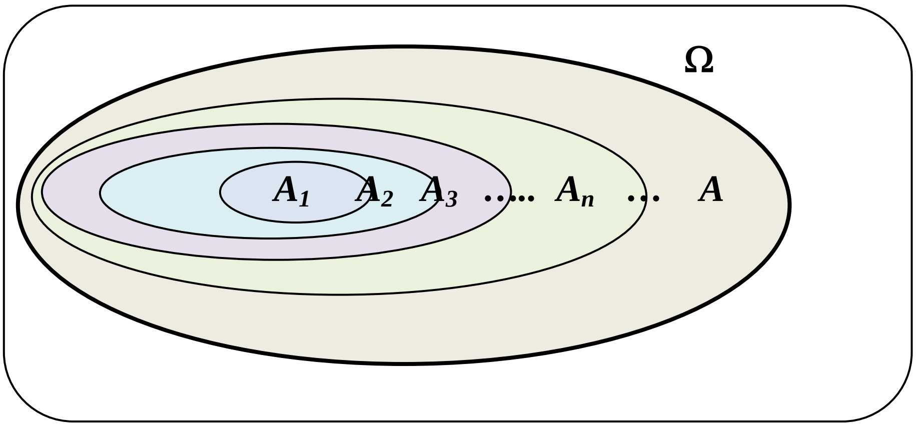
4) функция $\boldsymbol{P}$ – *непрерывная «в нуле»* функция, то есть, если $\{\boldsymbol{A}_n\}$, *n=1,2,3,…* счётная последовательность множеств, принадлежащих $\mathcal{A}$, таких что $\boldsymbol{A}_n\supseteq\boldsymbol{A}_{n+1}$ и пусть $\bigcap\limits_{n=1}^{\infty}\boldsymbol{A}_n=\varnothing$, тогда $\lim\limits_{n\to\infty}\boldsymbol{P}(\boldsymbol{A}_n)=0$.

Так как теорема заявляет, что все утверждения эквивалентны, то, доказывая теорему, последовательно рассмотрим импликации $\langle1\rangle\Rightarrow\langle2\rangle$; $\langle2\rangle\Rightarrow\langle3\rangle$; $\langle3\rangle\Rightarrow\langle4\rangle$ и завершим доказательство рассмотрением импликации $\langle4\rangle\Rightarrow\langle1\rangle$.

$\langle1\rangle\Rightarrow\langle2\rangle$. Дано: функция $\boldsymbol{P}$ - $\boldsymbol{\sigma}$*-аддитивная* функция. Доказать, что функция $\boldsymbol{P}$ – *непрерывная «сверху»* функция.

Пусть $\{A_n\}$, $n=1,2,3,\dots$, счётная последовательность расширяющихся множеств, то есть $A_n \subseteq A_{n+1}$, принадлежащих $\sigma$-алгебре $\mathcal{A}$.

Определим $A=\bigcup\limits_{n=1}^{\infty} A_n$. Ясно, что $A\in\mathcal{A}$. Нам надо доказать, что $P(A)=\lim\limits_{n\to\infty} P(A_n)$.




Очевидно, что множество $A$ можно представить так: $A=A_1\bigcup\left[\bigcup\limits_{n=2}^{\infty}(A_n \setminus A_{n-1})\right]$. Причём в этом выражении множества-слагаемые будут попарно несовместными. Так как по условию функция $P$ $\sigma$-аддитивная функция и $A\in\mathcal{A}$, то можем записать $P(A)=P(A_1)+\sum\limits_{n=2}^{\infty} P(A_n \setminus A_{n-1})$. В правой части этого равенства записана сумма членов числового ряда. Так как всегда $P(A)\le 1$, то числовой ряд $P(A_1), P(A_2 \setminus A_1), P(A_3 \setminus A_2),\dots, P(A_n \setminus A_{n-1}),\dots$ имеет конечную сумму своих членов, то есть он - сходится. В курсе математического анализа сумма членов сходящегося ряда определяется как предел последовательности частичных сумм членов этого ряда: $P(A)=\lim\limits_{n\to\infty}\left[P(A_1)+\sum\limits_{k=1}^{n-1}\left(P(A_{k+1})-P(A_k)\right)\right]$. То есть:

$$P(A)=\lim_{n\to\infty}\left[P(A_1)+\left(P(A_2)-P(A_1)\right)+\left(P(A_3)-P(A_2)\right)+\dots+\left(P(A_n)-P(A_{n-1})\right)\right].$$

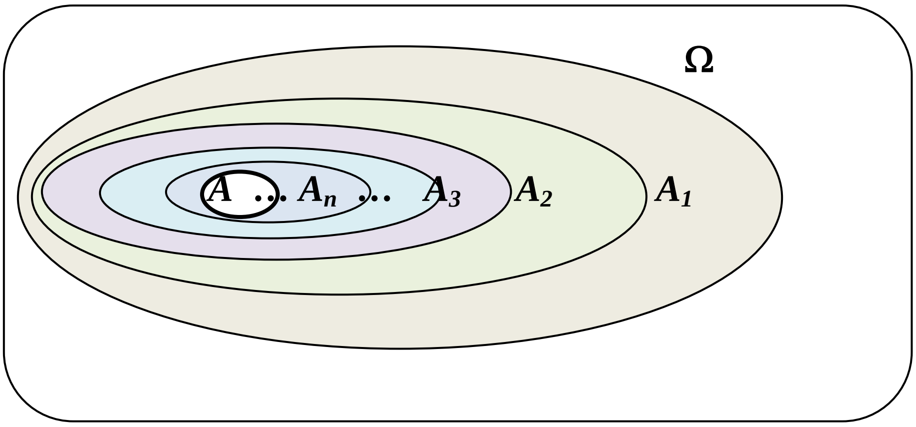
приведя подобные члены, получаем: $P(A)=\lim\limits_{n\to\infty} P(A_n)$.

$\langle 2\rangle \Rightarrow \langle 3\rangle$. Дано: Функция $P$ – *непрерывная «сверху»* функция. Доказать, что функция $P$ – *непрерывная «снизу»* функция.

Пусть $\{A_n\}$, $n=1,2,3,\dots$, счётная последовательность сужающихся множеств, то есть $A_n \supseteq A_{n+1}$, принадлежащих $\mathcal{A}$. Определим $A=\bigcap\limits_{n=1}^{\infty} A_n$. Ясно, что $A\in\mathcal{A}$. Нам надо доказать, что $P(A)=\lim\limits_{n\to\infty} P(A_n)$.

Из курса алгебры известно, что $\Omega \setminus A=\Omega \setminus \bigcap\limits_{n=1}^{\infty} A_n=\bigcup\limits_{n=1}^{\infty}(\Omega \setminus A_n)$.

Рассмотрим последовательность $\{\Omega \setminus A_n\}$, $n=1,2,3,\dots$ . Ясно, что $\Omega \setminus A_n \subseteq \Omega \setminus A_{n+1}$, то есть полученная последовательность – расширяющаяся. Так как функция $\boldsymbol{P}$ – непрерывная «сверху» функция, то $\boldsymbol{P}(\boldsymbol{\Omega} \setminus \boldsymbol{A}) = \lim\limits_{n\to\infty} \boldsymbol{P}(\boldsymbol{\Omega} \setminus \boldsymbol{A}_n)$ или




$\boldsymbol{P}(\boldsymbol{\Omega}) - \boldsymbol{P}(\boldsymbol{A}) = \boldsymbol{P}(\boldsymbol{\Omega}) - \lim\limits_{n\to\infty} \boldsymbol{P}(\boldsymbol{A}_n)$. Следовательно, $\boldsymbol{P}(\boldsymbol{A}) = \lim\limits_{n\to\infty} \boldsymbol{P}(\boldsymbol{A}_n)$.

$\langle 3\rangle \Rightarrow \langle 4\rangle$. Дано: Функция $\boldsymbol{P}$ – *непрерывная «снизу»* функция. Доказать, что функция $\boldsymbol{P}$ – *непрерывная «в нуле»* функция.

Пусть $\{\boldsymbol{A}_n\}$, *n=1,2,3,…,* счётная последовательность сужающихся множеств, то есть $\boldsymbol{A}_n \supseteq \boldsymbol{A}_{n+1}$, принадлежащих $\mathcal{A}$, и $\bigcap\limits_{n=1}^{\infty} \boldsymbol{A}_n = \varnothing$.

Надо доказать, что $\lim\limits_{n\to\infty} \boldsymbol{P}(\boldsymbol{A}_n) = 0$.

Так как $\{\boldsymbol{A}_n\}$, *n=1,2,3,…,* последовательность сужающихся множеств, то из непрерывности «снизу» функции $\boldsymbol{P}$ следует, что $\lim\limits_{n\to\infty} \boldsymbol{P}(\boldsymbol{A}_n) = \boldsymbol{P}(\varnothing)$.

А так как $\boldsymbol{P}(\varnothing)=0$, то $\lim\limits_{n\to\infty} \boldsymbol{P}(\boldsymbol{A}_n) = 0$.

$\langle 3\rangle \Rightarrow \langle 4\rangle$. Дано: Функция $\boldsymbol{P}$ – *непрерывная «в нуле»* функция. Доказать, что функция $\boldsymbol{P}$ – $\boldsymbol{\sigma}$*-аддитивная* функция.

Пусть $\{\boldsymbol{A}_i\}$, *i=1,2,3,…,* счётная последовательность попарно непересекающихся множеств, принадлежащих $\mathcal{A}$, то есть $\boldsymbol{A}_i \cap \boldsymbol{A}_j = \varnothing$ для любой пары индексов $i \neq j$. Ясно, что $\bigcup\limits_{i=1}^{\infty} \boldsymbol{A}_i \in \mathcal{A}$.

Надо доказать, что $\boldsymbol{P}\left(\bigcup\limits_{i=1}^{\infty} \boldsymbol{A}_i\right) = \sum\limits_{i=1}^{\infty} \boldsymbol{P}(\boldsymbol{A}_i)$.

Так как по условию теоремы $\boldsymbol{P}$ – конечная вероятностная функция, то:

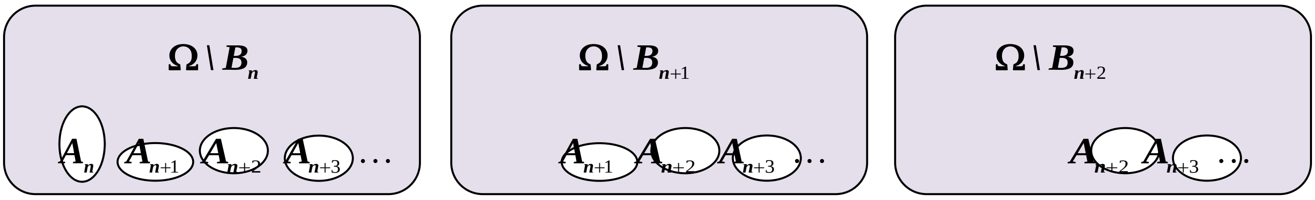
$$\boldsymbol{P}\left(\bigcup_{i=1}^{\infty} \boldsymbol{A}_i\right) = \sum_{i=1}^{n-1} \boldsymbol{P}(\boldsymbol{A}_i) + \boldsymbol{P}\left(\bigcup_{i=n}^{\infty} \boldsymbol{A}_i\right). \qquad (*)$$

Обозначим $\boldsymbol{B}_n = \bigcup\limits_{i=n}^{\infty} \boldsymbol{A}_i$ и рассмотрим последовательность множеств $\{\boldsymbol{B}_n\}$, *n=1,2,3,…,*. Ясно, что $\boldsymbol{B}_n \supseteq \boldsymbol{B}_{n+1}$, то есть $\{\boldsymbol{B}_n\}$ - последовательность сужающихся множеств. Чтобы воспользоваться непрерывностью «в нуле» вероятностной функции, надо показать, что $\bigcap\limits_{n=1}^{\infty} \boldsymbol{B}_n = \varnothing$.

Рассмотрим последовательность $\{\boldsymbol{\Omega} \setminus \boldsymbol{B}_n\}$, *n=1,2,3,…,*. Ясно, что $\boldsymbol{\Omega} \setminus \boldsymbol{B}_n \subseteq \boldsymbol{\Omega} \setminus \boldsymbol{B}_{n+1} \subseteq \boldsymbol{\Omega} \setminus \boldsymbol{B}_{n+2}$, то есть эта последовательность расширяющаяся и

$\boldsymbol{\Omega} \setminus \boldsymbol{B}_n \uparrow \boldsymbol{\Omega}$ при $\boldsymbol{n} \to \infty$.

Следовательно: $\boldsymbol{\Omega} \setminus \bigcap\limits_{n=1}^{\infty} \boldsymbol{B}_n = \bigcup\limits_{n=1}^{\infty} \left( \boldsymbol{\Omega} \setminus \boldsymbol{B}_n \right) \underset{n\to\infty}{\longrightarrow} \boldsymbol{\Omega}$, значит: $\bigcap\limits_{n=1}^{\infty} \boldsymbol{B}_n = \varnothing$.




Итак, последовательность множеств $\{\boldsymbol{B}_n\}$, *n=1,2,3,…*, - сужающаяся и $\bigcap\limits_{n=1}^{\infty} \boldsymbol{B}_n = \varnothing$, а так как функция $\boldsymbol{P}$ непрерывная «в нуле», то $\lim\limits_{n\to\infty} \boldsymbol{P}(\boldsymbol{B}_n) = 0$ или $\lim\limits_{n\to\infty} \boldsymbol{P}\left( \bigcup\limits_{i=n}^{\infty} \boldsymbol{A}_i \right) = 0$. В равенстве $(*)$ перейдём к пределу при $\boldsymbol{n} \to \infty$

$$\boldsymbol{P}\left( \bigcup_{i=1}^{\infty} \boldsymbol{A}_i \right) = \lim_{n\to\infty} \Big( \sum_{i=1}^{n-1} \boldsymbol{P}(\boldsymbol{A}_i) + \boldsymbol{P}\left( \bigcup_{i=n}^{\infty} \boldsymbol{A}_i \right) \Big) = \sum_{i=1}^{\infty} \boldsymbol{P}(\boldsymbol{A}_i) + \lim_{n\to\infty} \boldsymbol{P}\left( \bigcup_{i=n}^{\infty} \boldsymbol{A}_i \right) = \sum_{i=1}^{\infty} \boldsymbol{P}(\boldsymbol{A}_i) + 0,$$

то есть: $\boldsymbol{P}\left( \bigcup\limits_{i=1}^{\infty} \boldsymbol{A}_i \right) = \sum\limits_{i=1}^{\infty} \boldsymbol{P}(\boldsymbol{A}_i)$.

Теорема доказана.

***Комментарий***

Аксиоматический метод построения математической теории, которому мы следуем, предполагает введение с самого начала описаний основных понятий (исходных терминов) и определений объектов, изучаемых в теории. Сформулированная после этого система аксиом будет непротиворечивой и полной, если она позволяет путём логической дедукции получать и доказывать теоремы данной теории, описывающие свойства изучаемых объектов.

Непротиворечивость системы аксиом означает, что не существует такого утверждения $\mathscr{F}$, что это и противоположное ему утверждение можно одновременно доказать, используя принятую систему аксиом.

Полнота системы аксиом означает, что нельзя какую-либо аксиому из принятой системы доказать как теорему, применяя другие аксиомы.

Пятый постулат Евклида заявляет: «Через точку $\boldsymbol{A}$, не лежащую на прямой $\boldsymbol{a}$, всегда можно провести прямую $\boldsymbol{b}$, и при том только одну, параллельную этой прямой $\boldsymbol{a}$». Математики две тысячи лет, считая, что пятый постулат - это теорема, пытались доказать этот постулат, применяя остальные постулаты-аксиомы системы.

Приняв, предложенную Евклидом систему аксиом, мы доказываем различные теоремы о свойствах геометрических фигур. В частности, применяя пятый постулат Евклида, мы доказываем теорему: «Сумма углов любого треугольника равна $\boldsymbol{\pi}$». Но если принять в качестве пятого постулата утверждение: «Сумма углов любого треугольника равна $\boldsymbol{\pi}$», то утверждение: «Через точку $\boldsymbol{A}$, не лежащую на прямой $\boldsymbol{a}$, всегда можно провести прямую $\boldsymbol{b}$, и при том только одну, параллельную этой прямой $\boldsymbol{a}$» можно доказать как теорему.

Если требование $\sigma$-*аддитивности* числовой функции $\mu$ мы принимаем как аксиому 3, то утверждения о непрерывности «*сверху»*, «*снизу»* или «*в нуле*» становятся свойствами числовой функции $\mu$. И они доказываются как теоремы.

Теорема о *непрерывности* вероятностной функции позволяет сказать, что при аксиоматическом построении вероятностного пространства $<\Omega, \mathcal{A}, P>$ в аксиоме 3 требование $\sigma$-*аддитивности* числовой функции $\mu$ можно заменить любым другим требованием: непрерывности «*сверху»*, или непрерывности «*снизу»*, или непрерывности «*в нуле*».

В книге А.Н.Колмогорова «Основные понятия теории вероятностей» (1933г.,1936г.) формулируется *аксиома непрерывности*, требующая от числовой функции $\mu$ непрерывность «*в нуле*». Требование $\sigma$-*аддитивности* формулируется как свойство числовой функции $\mu$ и доказывается как теорема.

В геометрии числовая функция множеств $\mu$, как значение меры фигур изучаемых в геометрии, вводится *аксиоматически* как площадь прямоугольника, длины сторон которого являются рациональными числами. То есть, аксиоматически вводится понятие конечно-аддитивной функции множеств $\mu$. Аксиомы Евклида и аксиома конечной аддитивности функции $\mu$ позволяют нам определить площадь параллелограмма, треугольника, трапеции и любого другого множества, которое можно представить в виде конечного объединения непересекающихся прямоугольников, треугольников, параллелограммов, трапеций.

Если длины сторон прямоугольника будут иррациональными числами, то система аксиом Евклида и аксиома конечной аддитивности становится неполной. Поэтому мы формулируем аксиому $\sigma$-аддитивности числовой функции $\mu$ и, используя её непрерывность «*сверху»* и «*снизу»*, определяем площадь любого прямоугольника как произведение величин длин его основания и высоты, которые являются действительными числами.

Теорема о *непрерывности* числовой функции множеств $\mu$ используется и при определении площади круга.

В курсе геометрии средней школы рассматриваются последовательности вписанных в круг и описанных около круга правильных многоугольников, у которых на каждом шаге удваивается число сторон. Получаются, соответственно, расширяющаяся $\{\breve{S}_k\}$ и сужающаяся $\{\widehat{S}_k\}$ последовательности множеств ($S_1$ – треугольник, $S_2$ – шестиугольник, $S_3$ – двенадцатиугольник и т.д.). Ясно, что круг (множество – $S$) можно представить так: $S = \bigcup_{k=1}^{\infty} \breve{S}_k$ или

$$S = \bigcap_{k=1}^{\infty} \widehat{S}_k .$$

Величина площади круга это значение функции $\mu$, аргументом которой является круг - элемент алгебры геометрических фигур, изучаемых в средней школе. В силу непрерывности «*сверху»* и «*снизу»* числовой функции множеств

$\boldsymbol{\mu}$, площадь круга определяется как $\mu(S)=\lim\limits_{k\to\infty}\mu(\breve{S}_k)=\lim\limits_{k\to\infty}\mu(\widehat{S}_k)$. То есть, площадь круга $\mu(S)$ – *определяется* как предел последовательности площадей вписанных $\{\mu(\breve{S}_k)\}$, или описанных $\{\mu(\widehat{S}_k)\}$, многоугольников, получающейся при неограниченном удвоении числа сторон многоугольников. Площадь круга *равна* $\mu(S)=\pi\cdot r^2$

Аналогично в курсе математического анализа мы, применяя теорему о *непрерывности* числовой функции множеств $\mu$, вводим понятие определённого интеграла Римана.

Рассматривается класс функций $\{f(x)\}$, ограниченных требованием непрерывности на отрезке $[a;b]$. Определяется $\sigma-$алгебра множеств: всевозможных прямоугольников, трапеций, основаниями которых будут отрезки прямых $x=x_i$, $i=1,2,...$, $a\le x_i\le b$. Элементом этой алгебры также будет криволинейная трапеция $S$, сторонами которой будут: параллельные отрезки прямых $x=a$ и $x=b$, отрезок $[a;b]$ и часть графика непрерывной функции $f(x)$, построенного для $a\le x\le b$. Числовой функцией $\mu$, определённой на этой $\sigma-$алгебре будет функция, значениями которой будут площади элементов $\sigma-$ алгебры.

Для определения значения числовой функцией $\mu(S)$ строятся сужающиеся последовательности $\{\widehat{S}_k\}$ и расширяющиеся последовательности множеств $\{\breve{S}_k\}$, состоящих из прямоугольников (для любого $k$: $S\subset\widehat{S}_k$ и $S\supset\breve{S}_k$), называемые «верхними» и «нижними» суммами Дарбу. Криволинейная трапеция $S$ является как объединением $S=\bigcup\limits_{k=1}^{\infty}\breve{S}_k$, так и пересечением $S=\bigcap\limits_{k=1}^{\infty}\widehat{S}_k$ этих последовательностей множеств.

Числовые последовательности значений сумм площадей прямоугольников $\{\mu(\widehat{S}_k)\}$ и $\{\mu(\breve{S}_k)\}$, образованные «верхними» и «нижними» суммами Дарбу, сходятся к одному пределу, который *определяется* как значение определённого интеграла. То есть:

$$\lim_{k\to\infty}\mu(\widehat{S}_k)=\lim_{k\to\infty}\mu(\breve{S}_k)=\mu(S)=\int_a^b f(x)dx.$$

## §2.II Измеримые пространства и вероятностные меры на них

***Комментарий***

Если множество элементарных исходов $\Omega$ не более чем счётное, то вероятностная функция $P$ аксиоматически определяется на множестве $\Omega$ заданием конечного или счётного набора положительных чисел $\{p(\omega_i)\}$, $i=1,2,3,...,n$ или $i=1,2,3,...,\infty$. То есть каждому элементарному исходу $\omega_i$, как событию $A=\{\omega_i\}$, мы можем приписать положительную вероятность $P(A)$ равную $p(\omega_i)$.

Если мощность множества элементарных исходов $\Omega$ равна мощности континуума, то есть $|\Omega|=c$, то нельзя каждому элементу $\Omega$ приписать вероятность больше нуля.

Согласно аксиоматическому построению вероятностного пространства $<\Omega, \mathcal{A}, P>$, областью определения вероятностной функции $\boldsymbol{P}$ будет $\sigma$-алгебра $\mathcal{A}$ подмножеств множества $\Omega$.

Если множеством элементарных исходов $\Omega$ будет множество мощности континуум, то $\sigma$-алгеброй всех подмножеств множества $\Omega$ будет алгебра содержащая $\left|\mathcal{A}^*\right| = 2^{|\Omega|} = 2^c$ элементов – случайных событий. На такой алгебре задать *согласованным способом* вероятностную функцию $\boldsymbol{P}$ - невозможно.

Но измеримое пространство $<\Omega, \mathcal{A}>$ можно построить следующим образом. Выбирается некоторая система измеримых подмножеств $\boldsymbol{A}$ множества $\Omega$. Очевидно, что существуют $\sigma$-алгебры, содержащие эту систему, например, $\sigma$-алгебра $\mathcal{A}^*$ всех подмножеств множества $\Omega$. Пересечение всех таких $\sigma$-алгебр, называется *минимальной $\sigma$-алгеброй,* которая содержит систему $\boldsymbol{A}$. Минимальная $\sigma$-алгебра, содержащая систему $\boldsymbol{A}$, будет замкнута относительно операций объединения, пересечения и дополнения, то есть её элементы получаются путём применения к элементам системы $\boldsymbol{A}$ не более, чем счётного числа операций объединения, пересечения и дополнения. На этой *минимальной $\sigma$-алгебре* $\mathcal{A}$ можно определить вероятностную функцию $\boldsymbol{P}$.

**Теория**

**Измеримое пространство $<\boldsymbol{R}, \mathcal{B}(\boldsymbol{R})>$**

Пусть $\Omega$ - множество всех действительных чисел $\boldsymbol{R}$. В качестве системы подмножеств $\boldsymbol{A}$ возьмём систему всех полуинтервалов $\{[\boldsymbol{a},\boldsymbol{b})\}$, где $\boldsymbol{a}$ и $\boldsymbol{b}$ – действительные числа, включая $\boldsymbol{a}=-\infty$ и $\boldsymbol{b}=+\infty$ (под интервалом $[-\infty,\boldsymbol{b})$ будем понимать интервал $(-\infty,\boldsymbol{b})$). Применяя к элементам этой системы не более чем счётное число операций объединения, пересечения и дополнения, будем получать различные множества, являющиеся подмножествами множества $\boldsymbol{R}$. Эти множества называются борелевскими, а совокупность всех таких множеств будет минимальной $\sigma$-алгеброй, содержащей систему $\boldsymbol{A}$. Эту $\sigma$-алгебру будем обозначать $\mathcal{B}(\boldsymbol{R})$ и называть алгеброй борелевских множеств.

Привычные нам множества $[\boldsymbol{a},\boldsymbol{b}]$, $(\boldsymbol{a},\boldsymbol{b})$, $(\boldsymbol{a},\boldsymbol{b}]$ и одноточечное множество $\{\boldsymbol{a}\}$ являются борелевскими множествами, так как они могут быть «построены» из элементов системы $\boldsymbol{A}$ применением счётного числа операций объединения и пересечения:

$$[\boldsymbol{a},\boldsymbol{b}]=\bigcap_{\boldsymbol{n}=1}^{\infty}\left[\boldsymbol{a},\boldsymbol{b}+\frac{1}{\boldsymbol{n}}\right); \qquad (\boldsymbol{a},\boldsymbol{b})=\bigcup_{\boldsymbol{n}=1}^{\infty}\left[\boldsymbol{a}+\frac{1}{\boldsymbol{n}},\boldsymbol{b}\right);$$

$$(\boldsymbol{a},\boldsymbol{b}]=\left(\bigcup_{\boldsymbol{n}=1}^{\infty}\left[\boldsymbol{a}+\frac{1}{\boldsymbol{n}},\frac{\boldsymbol{a}+\boldsymbol{b}}{2}\right)\right)\cup\left(\bigcap_{\boldsymbol{n}=1}^{\infty}\left[\frac{\boldsymbol{a}+\boldsymbol{b}}{2},\boldsymbol{b}+\frac{1}{\boldsymbol{n}}\right)\right); \quad \{\boldsymbol{a}\}=\bigcap_{\boldsymbol{n}=1}^{\infty}\left[\boldsymbol{a},\boldsymbol{a}+\frac{1}{\boldsymbol{n}}\right).$$

Все элементы борелевской алгебры $\mathcal{B}(\boldsymbol{R})$ мы будем называть случайными событиями. Запас элементов борелевской алгебры $\mathcal{B}(\boldsymbol{R})$ позволяет решать большой круг практических задач, в которых надо определять вероятности наступления случайных событий. При таком способе определения алгебры случайных событий мы сможем осуществить практическое построение на алгебре $\mathcal{B}(\boldsymbol{R})$ нормированной и $\sigma$-аддитивной числовой функции множеств $\mu(A)$, то есть вероятностной функции $\boldsymbol{P}$.

***Комментарий***

Выбор вида подмножеств, образующих систему подмножеств $\mathcal{A}$ множества $\boldsymbol{R}$, не ограничивается какими-либо условиями. В некоторых учебниках и монографиях по теории вероятностей рассматривается система подмножеств $\mathcal{A}$ состоящая из полуинтервалов $\{(a,b]\}$. Из этих полуинтервалов можно «построить» привычные нам множества $[a,b]$, $(a,b)$, $[a,b)$ и одноточечное множество $\{a\}$. Ясно, что минимальной $\sigma$-алгеброй, содержащей такую систему подмножеств $\mathcal{A}$, будет уже знакомая нам $\sigma$-алгебра борелевских множеств $\mathcal{B}(\boldsymbol{R})$.

**Теория**

**Измеримое пространство $<\boldsymbol{R}^n, \mathcal{B}(\boldsymbol{R}^n)>$**

Множество всех действительных чисел $\boldsymbol{R}$ мы можем рассматривать как множество точек числовой прямой $\boldsymbol{R} = (-\infty, +\infty)$. Тогда множество $\boldsymbol{R}^n$ мы будем рассматривать как декартово произведение $\boldsymbol{n}$ экземпляров числовых прямых - координатных осей: $\boldsymbol{R}^n = \boldsymbol{R}_1 \times \boldsymbol{R}_2 \times \ldots \times \boldsymbol{R}_n$. Любой элемент $\boldsymbol{x} \in \boldsymbol{R}^n$ - это упорядоченный набор $\boldsymbol{x} = (x_1, \ldots, x_k, \ldots, x_n)$, где $x_k \in \boldsymbol{R}_k$ - $\boldsymbol{k}$-ая координата элемента $\boldsymbol{x}$, рассматриваемого как точка $\boldsymbol{n}$-мерного пространства $\boldsymbol{R}^n$.

Аналогично одномерному случаю, в качестве $\Omega$ возьмём множество всех точек $\boldsymbol{n}$-мерного пространства $\boldsymbol{R}^n$, а в качестве системы подмножеств $\mathcal{A}$ возьмём систему всех «прямоугольников» $\{[a,b)\}$, где $[a,b) = [a_1,b_1) \times [a_2,b_2) \times \ldots \times [a_n,b_n)$. Применяя к элементам системы подмножеств $\mathcal{A}$ не более чем счётное число операций объединения, пересечения и дополнения, будем получать различные множества, являющиеся подмножествами множества $\boldsymbol{R}^n$. Эти множества называются борелевскими, а совокупность всех таких множеств будет минимальной $\sigma$-алгеброй, содержащей систему $\mathcal{A}$. Эту минимальную $\sigma$-алгебру будем обозначать $\mathcal{B}(\boldsymbol{R}^n)$ и называть алгеброй борелевских множеств - алгеброй случайных событий.

***Комментарий***

Процедура построения измеримых пространств $<\boldsymbol{R}, \mathcal{B}(\boldsymbol{R})>$ и $<\boldsymbol{R}^n, \mathcal{B}(\boldsymbol{R}^n)>$ путём использования системы подмножеств $\mathcal{A}$ позволяет «проследить» процесс образования каждого борелевского множества и, впоследствии, осуществить логичное введение нормированной, $\sigma$-аддитивной числовой функции множеств на элементах $\sigma$-алгебры борелевских множеств.

При построении алгебры борелевских множеств $\mathcal{B}(\boldsymbol{R}^n)$ наряду с системой прямоугольников $\mathcal{A}$ можно рассматривать систему прямоугольников $\mathcal{A}^n$, получающихся как декартово произведение элементов систем $n$ подмножеств $\mathcal{A}_k=\{[a_k,b_k)\}$, где $\mathcal{A}_k \subset \mathcal{B}(R_k)$, то есть, если $[a,b)=[a_1,b_1)\times[a_2,b_2)\times\ldots\times[a_n,b_n)$, то $[a,b)\in\mathcal{A}^n$. Наименьшая $\sigma$-алгебра, содержащая такие прямоугольники, обозначается $\mathcal{B}(R_1)\times\mathcal{B}(R_2)\times\ldots\times\mathcal{B}(R_n)$ и называется *прямым произведением* $n$ экземпляров $\sigma$-алгебр $\mathcal{B}(R_k)$. Можно доказать, хотя это и так видно из процессов построений этих алгебр, что $\mathcal{B}(R^n)=\mathcal{B}(R_1)\times\mathcal{B}(R_2)\times\ldots\times\mathcal{B}(R_n)$.

## §2.II Типы и примеры вероятностных функций

### Измеримое пространство $<R,\mathcal{B}(R)>$

### *Комментарий*

В первой главе мы рассматривали измеримые пространства $<\Omega, \mathcal{A}>$, в которых множество элементарных исходов каждого конкретного примера имело или конечное, или счётное число элементов, то есть $|\Omega|=n$ или $|\Omega|=a$. Мера Лебега таких множеств равна нулю: $mes\Omega=0$. Случайные события, составляющие алгебру $\mathcal{A}$, в рассмотренных примерах имели мощность не более чем мощность счётного множества.

В аксиоматическом построении вероятностного пространства А.Н.Колмогорова множество $\Omega$ может иметь мощность континуума $|\Omega|=c$. Мера Лебега множеств такой мощности может быть или равной положительному числу $mes\Omega>0$, или быть равной нулю $mes\Omega=0$. Соответственно, случайные события - борелевские множества, составляющие алгебру $\mathcal{B}(R)$, могут иметь мощность не более чем мощность счётного множества, или иметь мощность континуума. Мера Лебега случайных событий может быть или положительной, или быть равной нулю.

В соответствии с этим вероятностная функция $P$, определяемая на элементах борелевской алгебры $\mathcal{B}(R)$, может быть трёх типов: дискретного, непрерывного или сингулярного.

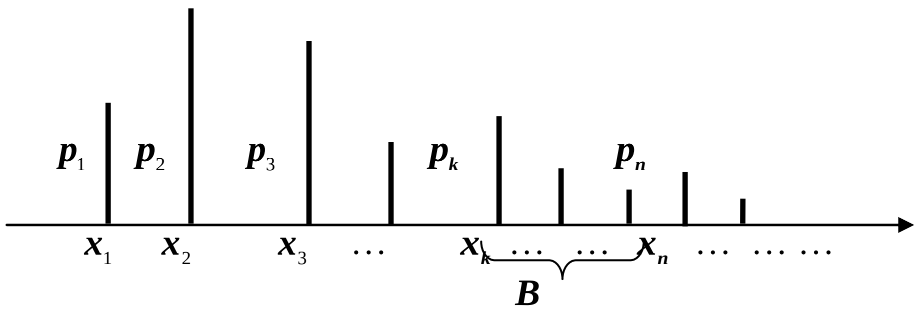
### Теория

### Дискретные вероятностные функции

Определение 1. Вероятностная функция $P$, называется ***вероятностной функцией дискретного типа***, если областью её определения является не более чем счётное множество точек числовой прямой $\{x_k\}$, $k=1,2,\ldots,n,\ldots$. Значения функции $P$ задаются набором положительных чисел $\{p_k\}$, $k=1,2,\ldots,n,\ldots$, то есть $P(x_k)=p_k$. При этом всегда сумма возможных значений функции $P$ равна единице: $\sum\limits_{k=1}^{n,\infty} p_k=1$.

Множество точек числовой прямой $G=\{x_k\}$, $k=1,2,\ldots,n,\ldots$ называется *носителем* вероятностной меры $P$, а набор положительных чисел $\{p_k\}$, $k=1,2,\ldots,n,\ldots$, называются *дискретным распределением вероятностей*.

Графически функцию дискретного типа $\boldsymbol{P}$ можно интерпретировать как размещение на точках $\{\boldsymbol{x}_k\}$ вертикальных отрезков прямых, длины которых равны $\{\boldsymbol{p}_k\}$.




Вероятность $\boldsymbol{P}(\boldsymbol{B})$ наступления случайного события $\boldsymbol{B}$, являющегося элементом борелевской алгебры событий $\mathcal{B}(\boldsymbol{R})$, будет равна $\boldsymbol{P}(\boldsymbol{B}) = \sum\limits_{x_k \in B} \boldsymbol{p_k}$ . То есть вероятность наступления события $\boldsymbol{B}$ равна сумме длин вертикалей $\boldsymbol{p}_k$, основания которых – точки $\boldsymbol{x}_k$ принадлежат множеству $\boldsymbol{B}$.

Ясно, что дискретных распределений вероятностей можно записать много, но есть распределения, наиболее часто встречающиеся в практических задачах, а поэтому имеющие соответствующие наименования.

1. <u>Дискретное равномерное распределение вероятностей</u>.

Это распределение встречается при проведении испытаний, в которых множество элементарных исходов имеет конечное число элементов, а все элементарные исходы равновозможные.

Элементарные исходы интерпретируются как конечное множество точек числовой прямой $\{\boldsymbol{x}_k\}$, $\boldsymbol{k} = 1,2,\ldots,\boldsymbol{n}$ . Все положительные числа $\boldsymbol{p}_k$ - одинаковые, то есть для любого номера $\boldsymbol{k}$ вероятность $\boldsymbol{P}(\boldsymbol{x}_k) = \boldsymbol{p}_k = \dfrac{1}{\boldsymbol{n}}$.

2. <u>Гипергеометрическое распределение вероятностей</u>.

Испытания, в которых возникает гипергеометрическое распределение, кратко можно описать так. Имеется $\boldsymbol{M}_1$ элементов одного вида и $\boldsymbol{M}_2$ элементов другого вида, всего $\boldsymbol{M} = \boldsymbol{M}_1 + \boldsymbol{M}_2$ элементов. Производится выбор $\boldsymbol{m}$ элементов без возвращения. Получающийся неупорядоченный набор содержит $\boldsymbol{k}$ элементов первого вида и $\boldsymbol{m}$-$\boldsymbol{k}$ элементов второго вида.

Элементарными исходами такого испытания будут все возможные наборы, содержащие $\boldsymbol{k}$ элементов первого вида, где $\boldsymbol{k}$ может принимать такие значения, при которых будут выполняться неравенства: $0 \le \boldsymbol{k} \le \boldsymbol{M}_1$ и $0 \le \boldsymbol{m} - \boldsymbol{k} \le \boldsymbol{M}_2$. Всего таких наборов будет $\boldsymbol{C}_{\boldsymbol{M}}^{\boldsymbol{m}}$.

Определим область определения вероятностной функции $\boldsymbol{P}$ как множество точек $\{\boldsymbol{x}_k\}$, где $\boldsymbol{x}_k = \boldsymbol{k}$ - возможные значения количества элементов первого вида, то есть количество элементарных исходов, в записи которых встречается $\boldsymbol{k}$ элементов первого типа. Положительные числа $\{\boldsymbol{p}_k\}$ определим так:

$\boldsymbol{p_k}=\boldsymbol{P}(\boldsymbol{k})=\dfrac{\boldsymbol{C}_{M_1}^{k}\cdot\boldsymbol{C}_{M_2}^{m-k}}{\boldsymbol{C}_{M}^{m}}$. Набор вероятностей $\left\{\dfrac{\boldsymbol{C}_{M_1}^{k}\cdot\boldsymbol{C}_{M_2}^{m-k}}{\boldsymbol{C}_{M}^{m}}\right\}$, где $0\le\boldsymbol{k}\le\boldsymbol{M}_1$ и $0\le\boldsymbol{m}-\boldsymbol{k}\le\boldsymbol{M}_2$, называется *гипергеометрическим распределением вероятностей.*

Так как всего элементарных исходов будет $\boldsymbol{C}_M^m$ штук, а все исходы будут равновозможными, то можно сказать, что гипергеометрическое распределение вероятностей есть частный случай равномерного распределения, элементарные исходы которого сгруппированы в $\boldsymbol{m}$ групп.

3. Бернуллиевское распределение вероятностей.

При однократном проведении испытания некоторое случайное событие $\boldsymbol{A}$ может наступить с вероятностью $\boldsymbol{P}(\boldsymbol{A})=\boldsymbol{p}$ и не наступить с вероятностью $\boldsymbol{P}(\overline{\boldsymbol{A}})=1-\boldsymbol{p}=\boldsymbol{q}$. Область определения бернуллиевской вероятностной функции $\boldsymbol{P}$ состоит из двух точек: $\boldsymbol{x}_1=0$ и $\boldsymbol{x}_2=1$. Значениями вероятностной функции будут числа $\boldsymbol{p}_1=\boldsymbol{P}(0)=\boldsymbol{q}$ и $\boldsymbol{p}_2=\boldsymbol{P}(1)=\boldsymbol{p}$. Кратко бернуллиевское распределение будем обозначать $\boldsymbol{B}_1(\boldsymbol{p})$.

4. Биномиальное распределение вероятностей.

И область определения $\{\boldsymbol{x_k}\}$, и множество возможных значений $\{\boldsymbol{p_k}\}$ вероятностной функции $\boldsymbol{P}$ определяются проведением $\boldsymbol{n}$ раз повторных независимых испытаний, или - испытаний по схеме Бернулли.

Здесь $\boldsymbol{x_k}=\boldsymbol{k}$, где $\boldsymbol{k}=0,1,\dots,\boldsymbol{n}$, а $\boldsymbol{p_k}=\boldsymbol{P}(\boldsymbol{k})=\boldsymbol{C}_n^k\boldsymbol{p}^k\boldsymbol{q}^{n-k}$.

Кратко биномиальное распределение вероятностей $\{\boldsymbol{C}_n^k\boldsymbol{p}^k\boldsymbol{q}^{n-k}\}$, $\boldsymbol{k}=0,1,\dots,\boldsymbol{n}$, будем обозначать $\boldsymbol{B_n}(\boldsymbol{p})$.

5. Пуассоновское распределение распределение.

Это распределение можно рассматривать как предельный случай биномиального распределения. Встречается при проведении «очень большого» числа повторных независимых испытаний ($\boldsymbol{n}$ велико), в каждом из которых вероятность $\boldsymbol{p}$ наступления события $\boldsymbol{A}$ «очень мала». То есть, $\boldsymbol{x_k}=\boldsymbol{k}$, где $\boldsymbol{k}=0,1,2,\dots$, а $\boldsymbol{p_k}=\boldsymbol{P}(\boldsymbol{k})=\dfrac{\boldsymbol{\lambda}^k\cdot\boldsymbol{e}^{-\lambda}}{\boldsymbol{k}!}$. Здесь положительное число $\boldsymbol{\lambda}$ - параметр распределения, смысл этого параметра узнаем позже. Проверим условие нормировки:

$$\sum_{k=0}^{\infty}\boldsymbol{p_k}=\sum_{k=0}^{\infty}\frac{\boldsymbol{\lambda}^k\cdot\boldsymbol{e}^{-\lambda}}{\boldsymbol{k}!}=\boldsymbol{e}^{-\lambda}\sum_{k=0}^{\infty}\frac{\boldsymbol{\lambda}^k}{\boldsymbol{k}!}=\boldsymbol{e}^{-\lambda}\cdot\boldsymbol{e}^{\lambda}=1.$$

Набор вероятностей $\left\{\dfrac{\boldsymbol{\lambda}^k\cdot\boldsymbol{e}^{-\lambda}}{\boldsymbol{k}!}\right\}$, $\boldsymbol{k}=0,1,2,\dots$ называется пуассоновским распределением вероятностей и будем его обозначать $\Pi(\lambda)$.

### ***Комментарий***

Термины «вероятность $\boldsymbol{P}(\boldsymbol{A})$ наступления события $\boldsymbol{A}$ – мала» и «число проводимых испытаний $\boldsymbol{n}$ – очень велико» имеют весьма относительное значение. Если нам говорят: «Вероятность $\boldsymbol{P(A)}$ наступления некоторого события $\boldsymbol{A}$ равна 0,02», то возникает вопрос: как оценивать это значение вероятности $\boldsymbol{P(A)}$=0,02? Это - большая, или - маленькая вероятность? Если мы не знаем в чём состоят последствия наступления этого события, то ничего определённого о величине значения вероятности мы сказать не сможем.

Изменим немного текст сообщения о значении вероятности наступления события. Рассмотрим два варианта.

Вариант первый: « Вероятность $\boldsymbol{P(A)}$ того, что при прыжке из самолёта парашют этой модели не раскроется, равна 0,02».

Вариант второй: « Вероятность $\boldsymbol{P(A)}$ того, что с первой попытки я не сдам экзамен по теории вероятностей, равна 0,02».

Если мы знаем чем мы «платим» за наступление, или ненаступление события $\boldsymbol{A}$, то мы сразу определим в каком случае вероятность 0,02 – большая, а в каком случае – маленькая.

Аналогичная неопределённость возникает, когда мы рассматриваем повторные независимые испытания и пытаемся определить: «Если число повторных независимых испытаний равно $\boldsymbol{n}$=150. Это - большое, или – маленькое число?».

Например. К середине лета на взрослом, хорошо развитом вишнёвом дереве созревает «большое» количество вишен. Пусть это число равно $\boldsymbol{n}$. Каждая косточка, находящаяся внутри ягоды, следующей весной может прорасти и стать началом нового вишнёвого дерева, или - не прорасти (событие $\boldsymbol{A}$ наступит, или не наступит, вероятность $\boldsymbol{P(A)}$ - мала). Если нас интересует число $\boldsymbol{k}$ – число косточек, которые проросли следующей весной, то количество вишен $\boldsymbol{n}$ – велико по сравнению с числом $\boldsymbol{k}$, и мы можем считать, что $\boldsymbol{k} = 0,1,2,\dots$. То есть число проросших косточек $\boldsymbol{k}$ мало по сравнению с числом $\boldsymbol{n}$, поэтому закон Пуассона называют «законом редких событий».

**Теория**

5. <u>Геометрическое распределение вероятностей</u>.

Это распределение вероятностей встречается при проведении повторных независимых испытаний до первого положительного исхода.

Областью определения вероятностной функции $\boldsymbol{P}$ будет множество натуральных чисел: $\boldsymbol{x_k} = \boldsymbol{k}$, где $\boldsymbol{k} = 1,2,\dots$, а положительные числа $\{\boldsymbol{p_k}\}$ являются членами бесконечно убывающей геометрической прогрессии: $\boldsymbol{p_k} = \boldsymbol{q}^{k-1} \cdot \boldsymbol{p}$, где $\boldsymbol{p}$ – вероятность положительного исхода при проведении одного испытания. Поэтому набор вероятностей $\{\boldsymbol{q}^{k-1} \cdot \boldsymbol{p}\}$, $\boldsymbol{k} = 1,2,\dots$, называется *геометрическим распределением вероятностей.*

6. <u>Отрицательно-биномиальное распределение вероятностей</u>.

Это распределение вероятностей встречается при проведении повторных независимых испытаний до $\boldsymbol{m}$-го положительного исхода. Очевидно, что геометрическое распределение будет частным случаем этого распределения если $\boldsymbol{m}$=1.

Областью определения вероятностной функции $\boldsymbol{P}$ будут натуральные числа: $\boldsymbol{x_k} = \boldsymbol{k}$, где $\boldsymbol{k} = \boldsymbol{m}, \boldsymbol{m}+1, \boldsymbol{m}+2, \ldots$, а положительные числа $\{\boldsymbol{p_k}\}$ будут равны: $\boldsymbol{p_k} = \boldsymbol{C_{k-1}^{m-1} q^{k-m}} \cdot \boldsymbol{p^m}$. Читателю предлагается самостоятельно проверить, что $\sum\limits_{k=m}^{\infty} \boldsymbol{C_{k-1}^{m-1} q^{k-m}} \cdot \boldsymbol{p^m} = 1$.

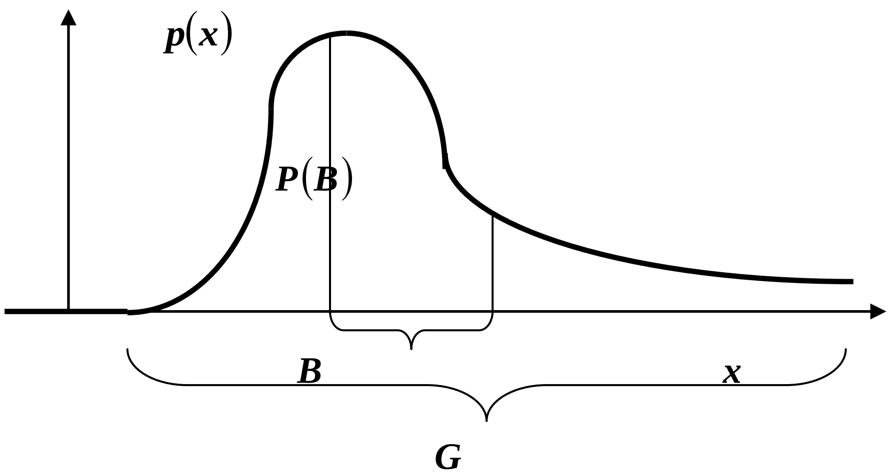
Набор вероятностей $\left\{\boldsymbol{C_{k-1}^{m-1} q^{k-m}} \cdot \boldsymbol{p^m}\right\}$, $\boldsymbol{k} = \boldsymbol{m}, \boldsymbol{m}+1, \boldsymbol{m}+2, \ldots$, называется *отрицательным биномиальным распределением вероятностей.*

## Абсолютно-непрерывные вероятностные функции

Определение 2. Вероятностная функция $\boldsymbol{P}$, называется ***вероятностной функцией абсолютно-непрерывного типа***, если областью её определения является множество $\boldsymbol{G} \subseteq \boldsymbol{R}$, у которого мощность равна мощности континуум: $|\boldsymbol{G}| = \boldsymbol{c}$, а мера Лебега – положительное число: $\boldsymbol{mes G} > 0$.

Определяется вероятностная функция $\boldsymbol{P}$ путём задания интегрируемой по Риману кусочно-непрерывной, неотрицательной функции $\boldsymbol{p}(\boldsymbol{x})$, такой что $\int\limits_{R} \boldsymbol{p}(\boldsymbol{x})\boldsymbol{dx} = 1$. Функция $\boldsymbol{p}(\boldsymbol{x})$ называется ***плотностью*** вероятностной функции $\boldsymbol{P}$

Значение функции $\boldsymbol{P}$ для любого случайного события $\boldsymbol{B}$, являющегося элементом борелевской алгебры событий $\mathcal{B}(\boldsymbol{R})$, определяется как значение интеграла $\int\limits_{B} \boldsymbol{p}(\boldsymbol{x})\boldsymbol{dx} = \boldsymbol{P}(\boldsymbol{B})$.




Если $\boldsymbol{x} \notin \boldsymbol{G}$, то $\boldsymbol{p}(\boldsymbol{x})=0$; если $\boldsymbol{x} \in \boldsymbol{G}$, то $\boldsymbol{p}(\boldsymbol{x}) \geq 0$. Ясно, что $\int\limits_{R} \boldsymbol{p}(\boldsymbol{x})\boldsymbol{dx} = \int\limits_{G} \boldsymbol{p}(\boldsymbol{x})\boldsymbol{dx} - \int\limits_{R \setminus G} \boldsymbol{p}(\boldsymbol{x})\boldsymbol{dx} = 1$. Множество $\boldsymbol{G}$ называется *носителем* вероятностной меры $\boldsymbol{P}$.

Приведём наиболее часто встречающиеся в практических задачах теории вероятностей плотности вероятности $\boldsymbol{p}(\boldsymbol{x})$.

1. Непрерывное равномерное распределение вероятностей.

$$\boldsymbol{p}(\boldsymbol{x}) = \begin{cases} 0, & \text{если} \quad \boldsymbol{x} \notin [\boldsymbol{a}, \boldsymbol{b}] \\ \dfrac{1}{\boldsymbol{b} - \boldsymbol{a}}, & \text{если} \quad \boldsymbol{x} \in [\boldsymbol{a}, \boldsymbol{b}] \end{cases}.$$

2. <u>Нормальное (Гаусса) распределение вероятностей</u>.

$$\boldsymbol{p}(\boldsymbol{x}) = \frac{1}{\sqrt{2\pi}\boldsymbol{\sigma}} \cdot \boldsymbol{e}^{-\frac{(x-m)^2}{2\sigma^2}}, \quad \boldsymbol{x} \in \boldsymbol{R}.$$

## Практика

<u>Пример</u>. Определить вероятность наступления случайного события $\boldsymbol{B} = [\boldsymbol{a}; \boldsymbol{b})$, если вероятностная функция $\boldsymbol{P}$ определяется плотностью $\boldsymbol{p}(\boldsymbol{x})$ - плотностью вероятности нормального закона (закона Гаусса) с параметрами $\boldsymbol{m}$ и $\boldsymbol{\sigma}$ (кратко $\mathcal{N}(\boldsymbol{m}, \boldsymbol{\sigma})$). Смысл параметров $\boldsymbol{m}$ и $\boldsymbol{\sigma}$ мы установим позже. Событие $\boldsymbol{B} = [\boldsymbol{a}; \boldsymbol{b})$ - борелевское множество и $\boldsymbol{P}(\boldsymbol{B}) = \frac{1}{\sqrt{2\pi}\boldsymbol{\sigma}} \int_a^b \boldsymbol{e}^{-\frac{(x-m)^2}{2\sigma^2}} \boldsymbol{dx}$. Сделаем замену переменной интегрирования, обозначив: $\frac{\boldsymbol{x} - \boldsymbol{m}}{\boldsymbol{\sigma}} = \boldsymbol{z}$. Тогда $\frac{\boldsymbol{dx}}{\boldsymbol{\sigma}} = \boldsymbol{dz}$ и, пересчитав пределы интегрирования, получаем: $\boldsymbol{P}(\boldsymbol{B}) = \frac{1}{\sqrt{2\pi}} \int_{\frac{a-m}{\sigma}}^{\frac{b-m}{\sigma}} \boldsymbol{e}^{-\frac{z^2}{2}} \boldsymbol{dz}$. То есть в любой задаче, где вероятностная функция $\boldsymbol{P}$ имеет плотность вероятности – плотность нормально закона, мы всегда придём к интегралу от функции $\boldsymbol{\varphi}(\boldsymbol{z}) = \frac{1}{\sqrt{2\pi}} \boldsymbol{e}^{-\frac{z^2}{2}}$, которая, как доказывается в курсе математического анализа, не имеет первообразной функции.

Как будет показано дальше, нормальный закон является исключительным, наиболее часто встречающимся в практических задачах, законом распределения вероятностей. Поэтому нам важно уметь вычислять вероятности наступления случайных событий $\boldsymbol{P}(\boldsymbol{B})$.

Рассматривается функция $\boldsymbol{\Phi}(\boldsymbol{x}) = \frac{1}{\sqrt{2\pi}} \int_0^x \boldsymbol{e}^{-\frac{z^2}{2}} \boldsymbol{dz}$, которая называется *функцией Лапласа.* Ясно, что $\boldsymbol{P}(\boldsymbol{B}) = \boldsymbol{\Phi}\left(\frac{\boldsymbol{b} - \boldsymbol{m}}{\boldsymbol{\sigma}}\right) - \boldsymbol{\Phi}\left(\frac{\boldsymbol{a} - \boldsymbol{m}}{\boldsymbol{\sigma}}\right)$. Методами численного интегрирования вычисляются значения функции $\boldsymbol{\Phi}(\boldsymbol{x})$ и составляются таблицы её значений:

| $\boldsymbol{x}$ | 0,00 | 0,01 | 0,02 | … | 1,00 | 1,01 | … |
|---|---|---|---|---|---|---|---|
| $\boldsymbol{\Phi}(\boldsymbol{x})$ | 0,0000 | 0,0040 | 0,0080 | … | 0,3413 | 0,3438 | ... |

Значение аргумента $\boldsymbol{x}$ может быть любым действительным числом от $-\infty$ до $+\infty$. Ясно, что составить таблицы для всех возможных значений $\boldsymbol{x}$ невозможно. И при определении значений функции Лапласа кроме таблиц используются свойства функции Лапласа:

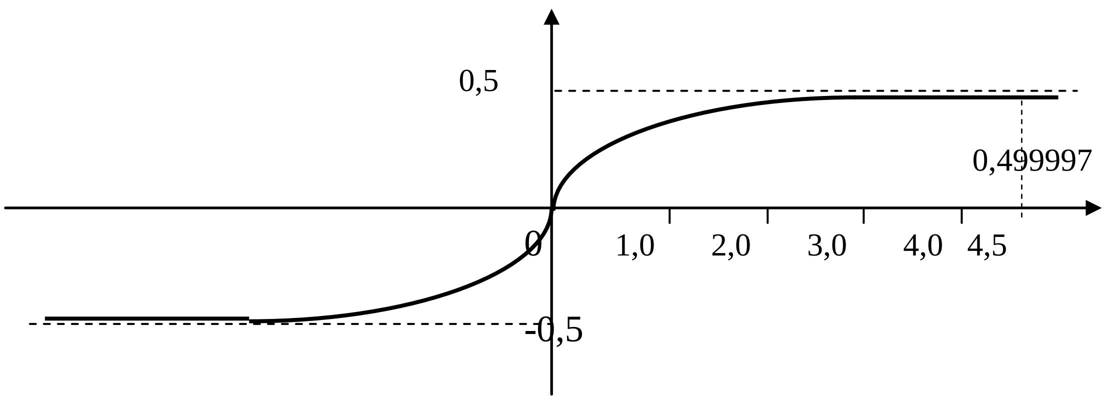
1) $\boldsymbol{\Phi}(0)=0$;

2) Функция Лапласа – возрастающая функция;

3) Функция Лапласа – нечётная функция, то есть $\boldsymbol{\Phi}(-\boldsymbol{x})=-\boldsymbol{\Phi}(\boldsymbol{x})$;

4) $\boldsymbol{\Phi}(+\infty)=0,5$.

Значения функции Лапласа при изменении значений аргумента $\boldsymbol{x}$ от 0 до 4,5 «довольно быстро» увеличиваются от 0 до 0,499997, и для любого значения аргумента $\boldsymbol{x}$, которое больше чем 4,0, можно считать, что $\boldsymbol{\Phi}(\boldsymbol{x})$=0,5.




Следовательно, при вычислении вероятности наступления случайного события $\boldsymbol{B}=[\boldsymbol{a};\boldsymbol{b})$ необходимо:

*а)* пересчитать пределы интегрирования $\boldsymbol{\alpha}=\dfrac{\boldsymbol{a}-\boldsymbol{m}}{\boldsymbol{\sigma}}$ и $\boldsymbol{\beta}=\dfrac{\boldsymbol{b}-\boldsymbol{m}}{\boldsymbol{\sigma}}$;

*б)* по таблицам значений функции Лапласа определить значения $\boldsymbol{\Phi}(\boldsymbol{\alpha})$ и $\boldsymbol{\Phi}(\boldsymbol{\beta})$ (если значения аргументов $\boldsymbol{\alpha}$ и $\boldsymbol{\beta}$ не принадлежат интервалу (0;4,0), то для определения значения функции $\boldsymbol{\Phi}(\boldsymbol{x})$ надо использовать её свойства);

*в)* Применить формулу Ньютона-Лейбница: $\boldsymbol{P}(\boldsymbol{B})=\boldsymbol{\Phi}(\boldsymbol{\beta})-\boldsymbol{\Phi}(\boldsymbol{\alpha})$.

**Теория**

3. Экспоненциальное распределение вероятностей.

$$\boldsymbol{p}(\boldsymbol{x})=\begin{cases} 0, & \text{если } \boldsymbol{x}\in(-\infty;0) \\ \dfrac{1}{\boldsymbol{\mu}}\cdot \boldsymbol{e}^{-\frac{x}{\mu}}, & \text{если } \boldsymbol{x}\in[0;+\infty) \end{cases}.$$

4. Гамма распределение вероятностей.

$$\boldsymbol{p}(\boldsymbol{x})=\begin{cases} 0, & \textbf{\textit{если}} \quad \boldsymbol{x}\in(-\infty;0) \\ \dfrac{\boldsymbol{x}^{\alpha-1}\cdot \boldsymbol{e}^{-\frac{x}{\beta}}}{\Gamma(\alpha)\boldsymbol{\beta}^{\alpha}}, & \textbf{\textit{если}} \quad \boldsymbol{x}\in[0;+\infty) \end{cases}.$$

5. Двустороннее экспоненциальное (Лапласа) распределение вероятностей.

$$p(x)=\frac{1}{2\mu}\cdot e^{-\frac{|x|}{\mu}},\ x\in R.$$

6. Распределение вероятностей Коши.

$$p(x)=\frac{\vartheta}{\pi\cdot\left(\vartheta^2+x^2\right)},\ x\in R.$$

В записях плотностей вероятностей $p(x)$ распределений $2\div 6$ участвуют числовые параметры $m,\sigma^2,\mu,\alpha,\beta,\vartheta$, вероятностный смысл и возможные значения которых будут рассмотрены позже.

**Сингулярные вероятностные функции**

Определение 3. Вероятностная функция $P$, называется ***вероятностной функцией сингулярного типа,*** если областью её определения является множество $\mathcal{D}\subseteq R$, у которого мощность $|\mathcal{D}|=c$, а мера Лебега $mes\,\mathcal{D}=0$.

Примером такого множества будет канторово множество, которое является подмножеством сегмента $U=[0;1]$. Любое число $x\in[0;1]$ можно представить в виде троичной дроби: $x=\sum_{i=1}^{\infty}\frac{a_i}{3^i}$, где $a_i=0$; 1 или 2. Множество $U$ эквивалентно множеству троичных последовательностей $T=\{\omega\}$, где $\omega=(a_1,a_2,\ldots,a_i,\ldots)$. Из этого множества выделяется подмножество $\mathcal{T}$ двоичных последовательностей $\mathcal{T}=\{\omega\}$, где $\omega=(a_1,a_2,\ldots,a_i,\ldots)$, здесь $a_i=0$ или 2. Множество $\mathcal{T}$ имеет мощность $|\mathcal{T}|=c$ и меру Лебега $mes\,\mathcal{T}=0$. Множество чисел $\mathcal{D}$, $\mathcal{D}\subset U$, эквивалентное множеству двоичных последовательностей $\mathcal{T}$, будет носителем вероятностной меры $P$ сингулярного типа.

Записать аналитический вид функции $P$ нельзя. Но можно описать процесс её построения.

Напомним, что каждому элементу множества $\mathcal{T}$ соответствует точка сегмента $[0;1]$, в троичном представлении которой $x=\sum_{i=1}^{\infty}\frac{a_i}{3^i}$ коэффициенты $a_i$ равны только 0 или 2.

Первый шаг. Разделим полуинтервал $[0;1)$ на три равные части: $\left[\frac{0}{3};\frac{1}{3}\right)$, $\left[\frac{1}{3};\frac{2}{3}\right)$ и $\left[\frac{2}{3};\frac{3}{3}\right)$. В точках $x_1=\frac{0}{3^1}$ и $x_2=\frac{2}{3^1}$ зададим значения вероятностной функции $P$: $P_1\left(\frac{0}{3^1}\right)=P_1\left(\frac{2}{3^1}\right)=\frac{1}{2}$. Точкам $x$, принадлежащим полуинтервалу $\left[\frac{1}{3};\frac{2}{3}\right)$, соответствуют троичные дроби $x=\sum_{i=1}^{\infty}\frac{a_i}{3^i}$, у которых $a_1=1$. Этот полуинтервал мы больше рассматривать не будем.

Второй шаг. Разделим каждый полуинтервал $\left[\frac{0}{3};\frac{1}{3}\right)$ и $\left[\frac{2}{3};\frac{3}{3}\right)$ на три равные части: $\left[\frac{0}{3^2};\frac{1}{3^2}\right)$, $\left[\frac{1}{3^2};\frac{2}{3^2}\right)$, $\left[\frac{2}{3^2};\frac{3}{3^2}\right)$и $\left[\frac{6}{3^2};\frac{7}{3^2}\right)$,$\left[\frac{7}{3^2};\frac{8}{3^2}\right)$,$\left[\frac{8}{3^2};\frac{9}{3^2}\right)$. Зададим значения $\boldsymbol{P}$: $\boldsymbol{P_2}\left(\frac{0}{3^2}\right)=\boldsymbol{P}_2\left(\frac{2}{3^2}\right)=\boldsymbol{P_2}\left(\frac{6}{3^2}\right)=\boldsymbol{P}_2\left(\frac{8}{3^2}\right)=\frac{1}{2^2}$. Ясно, что $\boldsymbol{x_1}=\frac{0}{3}=\frac{0}{3}+\frac{0}{3^2}$, $\boldsymbol{x_2}=\frac{2}{3^2}=\frac{0}{3}+\frac{2}{3^2}$, $\boldsymbol{x_3}=\frac{6}{3^2}=\frac{2}{3}+\frac{0}{3^2}$, $\boldsymbol{x_4}=\frac{8}{3^2}=\frac{2}{3}+\frac{2}{3^2}$, то есть это точки, которым соответствуют последовательности $\boldsymbol{\omega}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_i,\ldots)$ множества $\mathcal{T}$, у которых $\boldsymbol{a_1}$ и $\boldsymbol{a_2}$ равны или 0, или 2, а все остальные $\boldsymbol{a_i}$ равны 0. Точкам полуинтервалов $\left[\frac{1}{3^2};\frac{2}{3^2}\right)$ и $\left[\frac{7}{3^2};\frac{8}{3^2}\right)$ соответствуют последовательности $\boldsymbol{\omega}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_i,\ldots)$, в которых $\boldsymbol{a_2}$ равно 1. Эти полуинтервалы мы больше рассматривать не будем.

На третьем шаге первый, третий, четвёртый и шестой полуинтервалы снова разделим на три равные части. Получим четыре тройки полуинтервалов, длины которых будут равны $\frac{1}{3^3}$. В левой точке первого и третьего полуинтервала каждой тройки (всего таких точек будет $2^3=8$ штук) зададим значения $\boldsymbol{P}$: $\boldsymbol{P_3}=\frac{1}{2^3}$. Средние интервалы каждой тройки больше рассматривать не будем.

На $\boldsymbol{i}$-том шаге получим $2^{\boldsymbol{i}-1}$ троек полуинтервалов. В $2^{\boldsymbol{i}}$ точках, в троичном представлении которых $\boldsymbol{a_3}$ равно 0 или 2, зададим $\boldsymbol{P_i}=\frac{1}{2^{\boldsymbol{i}}}$ и исключим из рассмотрения $2^{\boldsymbol{i}-1}$ полуинтервалов, точки которых в троичном представлении имеют $\boldsymbol{a_3}$=1. Длина каждого такого интервала равна $\frac{1}{3^{\boldsymbol{i}}}$. А длина всех исключённых из рассмотрения интервалов равна $\frac{2^{\boldsymbol{i}-1}}{3^{\boldsymbol{i}}}$.

Такое деление на каждом шаге получающихся интервалов и задание функции $\boldsymbol{P}$ продолжим счётное число раз. В результате получим множество чисел $\mathcal{D}\subseteq\boldsymbol{R}$, которое будет эквивалентно множеству двоичных последовательностей $\mathcal{T}$, мощность которого равна $2^{a}=\boldsymbol{c}$. Значит мощность $|\mathcal{D}|=\boldsymbol{c}$.

Ясно, что на каждом $\boldsymbol{i}$-том шаге будет выполняться $\mathcal{D}\subset\boldsymbol{U}\setminus\boldsymbol{U_i}$, где $\{\boldsymbol{U_i}\}$, $\boldsymbol{i}=1,2,\ldots$, расширяющаяся последовательность множеств. Каждое $\boldsymbol{U_i}$ состоит из полуинтервалов, которые исключаются из рассмотрения на первом, втором, …, $\boldsymbol{i}$-том шаге. Следовательно, $\boldsymbol{mes}\,\mathcal{D}=\lim\limits_{\boldsymbol{i}\to\infty}[\boldsymbol{mes}(\boldsymbol{U}\setminus\boldsymbol{U_i})]=1-\sum\limits_{\boldsymbol{i}=1}^{\infty}\frac{2^{\boldsymbol{i}-1}}{3^{\boldsymbol{i}}}=0$.

На элементах множества $\mathcal{D}$ вероятностная функция $\boldsymbol{P}$ принимает значение

$\boldsymbol{P} = \lim\limits_{i\to\infty}\sum\limits_{j=1}^{i}\boldsymbol{P}_j = \lim\limits_{i\to\infty}\sum\limits_{j=1}^{i}\dfrac{\frac{1}{2}\cdot\left(1-\frac{1}{2^i}\right)}{1-\frac{1}{2}} = 1$, хотя в каждой точке множества $\mathcal{D}$ значение вероятностной функции $\boldsymbol{P}$ будет равно нулю.

***Комментарий***

Тремя рассмотренными типами вероятностных функций исчерпывается возможное многообразие функций, которые могут быть определены на $\boldsymbol{\sigma}$-алгебре борелевских множеств $\mathcal{B}(\boldsymbol{R})$. Точнее, произвольная вероятностная функция $\boldsymbol{P}$ на измеримом пространстве $<\boldsymbol{R},\mathcal{B}(\boldsymbol{R})>$ может быть представлена в виде линейной комбинации $\boldsymbol{P} = \boldsymbol{\alpha}\cdot\boldsymbol{P_1} + \boldsymbol{\beta}\cdot\boldsymbol{P}_2 + \boldsymbol{\gamma}\cdot\boldsymbol{P}_3$, где $\boldsymbol{P}_1$- дискретная, $\boldsymbol{P}_2$ - абсолютно непрерывная, $\boldsymbol{P}_3$ - сингулярная вероятностные функции; коэффициенты $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ - неотрицательные числа и $\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\gamma}=1$.

## Измеримое пространство $<\boldsymbol{R^n},\mathcal{B}(\boldsymbol{R^n})>$

***Комментарий***

Как и в рассмотренном случае измеримого пространства $<\boldsymbol{R},\mathcal{B}(\boldsymbol{R})>$ вероятностная функция $\boldsymbol{P}$, определяемая на элементах борелевской алгебры $\mathcal{B}(\boldsymbol{R^n})$, может быть трёх типов: дискретного, непрерывного или сингулярного. Тип функции $\boldsymbol{P}$ определяется мощностью и мерой Лебега области её определения. Рассмотрим определения этих типов и примеры вероятностных функций $\boldsymbol{P}$, для случая $\boldsymbol{n}=2$. Случай $<\boldsymbol{R^2},\mathcal{B}(\boldsymbol{R^2})>$ - наиболее удобный при первом знакомстве с многомерными вероятностными пространствами.

## Теория

## Дискретные вероятностные функции

Определение 4. Вероятностная функция $\boldsymbol{P}$, называется ***вероятностной функцией дискретного типа***, если областью её определения является не более чем счётное множество точек $\{(\boldsymbol{x_i};\boldsymbol{y_k})\}$, $\boldsymbol{i}=1,2,\ldots,\boldsymbol{n},\ldots;\boldsymbol{k}=1,2,\ldots,\boldsymbol{m},\ldots$, плоскости $\boldsymbol{R^2}$. Значения функции $\boldsymbol{P}$ задаются набором положительных чисел $\{\boldsymbol{p_{i,k}}\}$, $\boldsymbol{i}=1,2,\ldots,\boldsymbol{n},\ldots;\boldsymbol{k}=1,2,\ldots,\boldsymbol{m},\ldots$, то есть $\boldsymbol{P}((\boldsymbol{x_i};\boldsymbol{y_k}))=\boldsymbol{p_{i,k}}$. При этом всегда сумма всех возможных значений функции $\boldsymbol{P}$ равна единице: $\sum\limits_{i=1}^{n,\infty}\sum\limits_{k=1}^{m,\infty}\boldsymbol{p_{i,k}}=1$.

Множество точек плоскости $\{(\boldsymbol{x_i};\boldsymbol{y_k})\}$, $\boldsymbol{i}=1,2,\ldots,\boldsymbol{n},\ldots;\boldsymbol{k}=1,2,\ldots,\boldsymbol{m},\ldots$ называется *носителем* вероятностной меры $\boldsymbol{P}$, а набор положительных чисел $\{\boldsymbol{p_{i,k}}\}$, называются *дискретным распределением вероятностей*.

Вероятность $\boldsymbol{P}(\boldsymbol{A})$ наступления случайного события $\boldsymbol{A}$, являющегося элементом борелевской алгебры событий $\mathcal{B}(\boldsymbol{R^2})$, будет равна $\boldsymbol{P}(\boldsymbol{A})=\sum\limits_{(x_i;y_k)\in A}\boldsymbol{p_{i,k}}$.

***Комментарий***

Примеры вероятностных функций большинства многомерных дискретных распределений, определяемых на измеримом пространстве $<\boldsymbol{R^n},\mathcal{B}(\boldsymbol{R^n})>$, аналогичны примерам, которые приведены выше при рассмотрении измеримого пространства <***R***,$\mathcal{B}$(***R***)>. Рассмотрим некоторые из них, наиболее часто встречающиеся в практических задачах.

**Теория**

1. Дискретное равномерное распределение вероятностей.

Это распределение встречается при проведении испытаний, в которых множество элементарных исходов имеет конечное число элементов, которые могут быть интерпретированы как точки на плоскости. Все элементарные исходы – равновозможные. То есть, область определения функции $\boldsymbol{P}$ – конечное множество точек плоскости $\{(x_i\,;y_k)\}$, $i = 1,2,...,n;k = 1,2,...,m$. Все положительные числа $p_{i,k}$, значения функции $\boldsymbol{P}$, - одинаковые, то есть для любой пары индексов $(i,k)$ вероятность $P((x_i\,;y_k)) = p_{i,k} = \dfrac{1}{n\cdot m}$.

2. Гипергеометрическое распределение вероятностей.

Испытания, в которых возникает гипергеометрическое распределение, кратко можно описать так. Имеется $M_1$ элементов первого вида, $M_2$ элементов второго вида, и $M_3$ элементов третьего вида, всего $M = M_1 + M_2 + M_3$ элементов. Производится выбор $m$ элементов без возвращения. Получающийся неупорядоченный набор содержит $k$ элементов первого вида, $l$ элементов второго вида и $m - (k + l)$ элементов третьего вида. При этом должны выполняться неравенства: $0 \le k \le M_1$, $0 \le l \le M_2$ и $0 \le m - (k + l) \le M_3$. Элементарными исходами такого испытания будут получающиеся наборы $m$ элементов, в которых порядок чередования типов вида – не важен.. Определим область определения вероятностной функции $\boldsymbol{P}$ как множество точек $\{x_k\,;y_l\}$, где $x_k = k$, $y_l = l$ - возможные значения количества элементов первого и второго вида. Положительные числа $\{p_{k,l}\}$ определим так: $p_{k,l} = P(k;l) = \dfrac{C_{M_1}^k \cdot C_{M_2}^l \cdot C_{M_3}^{m-(k+l)}}{C_M^m}$.

3. Полиномиальное распределение вероятностей.

Это распределение является обобщением биномиального распределения в случае, когда каждое из $n$ повторных независимых испытаний имеет не два, а три: $A$, $B$ и $C$, или более, несовместных исхода. Если исходов будет три, то $A \cup B \cup C = U$, то есть исходы составляют ПГС. Следовательно, всегда будет выполняться $P(A) + P(B) + P(C) = p + q + r = 1$. Областью определения вероятностной функции $\boldsymbol{P}$ будет множество точек $\{x_k\,;y_l\}$, где $x_k = k$, $y_l = l$. Здесь $0 \le k,l$ и $k + l \le n$. Множество возможных значений $\{p_{k,l}\}$ вероятностной функции $\boldsymbol{P}$ определяются проведением $n$ раз повторных независимых испытаний, или - испытаний по схеме Бернулли с тремя исходами. Здесь для каждой возможной пары индексов $(k,l)$: $p_{k,l} = P(k,l) = \dfrac{n!}{k!\cdot l!\cdot(n-(k+l))!}\,p^k \cdot q^l \cdot r^{n-(k+l)}$. Или, что встречается гораздо реже,: $p_{k,l} = P(k,l) = C_n^k \cdot C_{n-k}^l p^k \cdot q^l \cdot r^{(n-k)-l}$.

## Теория

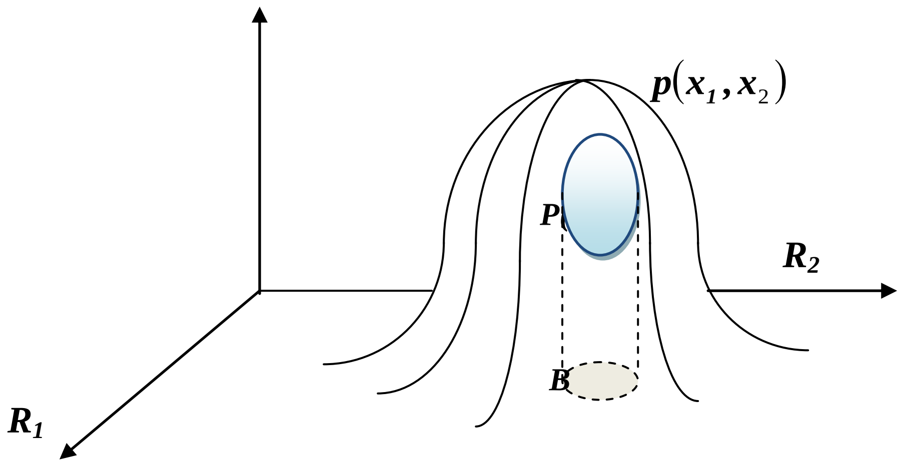
## Абсолютно-непрерывные вероятностные функции

Определение 5. Вероятностная функция $P$ называется ***вероятностной функцией абсолютно-непрерывного типа***, если областью её определения является множество $G \subseteq R^n$, у которого мощность $|G|=c$, а мера Лебега $mesG > 0$. Значения функции $P$ для любого случайного события $B$, являющегося элементом борелевской алгебры событий $\mathcal{B}(R^n)$, определяются как значения $n$-кратного интеграла $\iint\limits_B \ldots \int p(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n = P(B)$.

Здесь $p(x_1, x_2, \ldots, x_n)$ - неотрицательная, кусочно-непрерывная, интегрируемая по Риману функция, такая, что $\iint\limits_{R^n} \ldots \int p(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n = 1$. Функция $p(x_1, x_2, \ldots, x_n)$ называется ***плотностью*** вероятности.

***Комментарий***

Плотность вероятности $p(x_1, x_2, \ldots, x_n)$ при $n$=2 интерпретируется как «поверхность» над плоскостью $R_1 \times R_2$, ограничивающая единичный объём.




Вероятность $P(B)$ наступления случайного события $B$, являющегося элементом борелевской алгебры событий $\mathcal{B}(R^2)$, будет равна $P(B) = \iint\limits_B p(x_1, x_2) dx_1 dx_2$ и интерпретируется как величина объёма «цилиндра», основанием которого является множество $B \subset R^2$.

Приведём два примера абсолютно-непрерывных вероятностных функций.

1. Непрерывное равномерное распределение.

Областью определения плотности вероятности $p(x_1, x_2)$ является множество $G \subset R^2$, имеющего мощность $|G|=c$ и конечную меру Лебега $mesG > 0$. Плотность вероятности задаётся двойным равенством:

$$p(x_1, x_2) = \begin{cases} 0, & \textit{если} \quad (x_1, x_2) \notin G \\ \dfrac{1}{mesG}, & \textit{если} \quad (x_1, x_2) \in G \end{cases}.$$

2. Нормальное (Гаусса) распределение.

Пусть $\boldsymbol{\Sigma}$ - положительно определённая симметрическая матрица порядка $n \times n$:

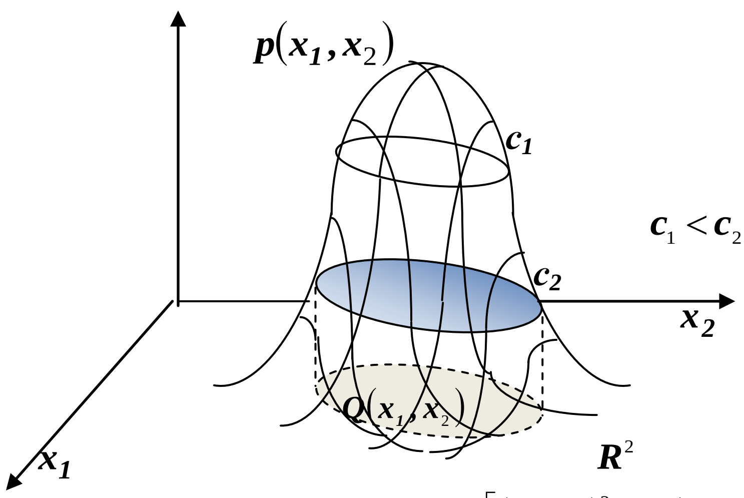
$$\Sigma = \begin{pmatrix} \sigma_1^2 & \mu_{1,2} & \dots & \mu_{1,n} \\ \mu_{2,1} & \sigma_2^2 & \dots & \mu_{2,n} \\ \dots & \dots & \dots & \dots \\ \mu_{n,1} & \mu_{n,2} & \dots & \sigma_n^2 \end{pmatrix}$$

, здесь $\mu_{i,k} = \mu_{k,i}$. Обозначим: $\left|\Sigma^{-1}\right| \equiv det\left(\Sigma^{-1}\right)$, тогда плотность $n$-мерного *гауссовского*, или *нормального распределения* имеет вид:

$$p(x_1, x_2, \dots, x_n) = \frac{\left|\Sigma^{-1}\right|^{1/2}}{(2\pi)^{n/2}} \cdot e^{-\frac{1}{2}\left[(x-m)' \cdot \Sigma^{-1} \cdot (x-m)\right]}.$$

Здесь $(x - m)$ - матрица порядка $n \times 1$: $(x - m) = \begin{pmatrix} x_1 - m_1 \\ \dots\dots \\ x_n - m_n \end{pmatrix}$, $(x - m)'$ – транспонированная матрица $(x - m)$. Матрица $\boldsymbol{\Sigma}$ называется ковариационной матрицей, вектор $m = (m_1, m_2, \dots, m_n)$ - вектор *средних значений*. Смысл параметров $m_i$, $\sigma_i^2$ и $\mu_{i,k}$ будет объяснён позже.

В случае $n=2$ обозначим $\mu_{1,2} = \mu_{2,1} = \rho \cdot \sigma_1 \cdot \sigma_2$ и вычислим: $\left|\Sigma^{-1}\right| = \dfrac{1}{\sigma_1^2 \cdot \sigma_2^2 \cdot (1 - \rho^2)}$. Тогда плотность $p(x_1, x_2)$ приводится к виду:




$$p(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-m_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-m_1)(x_2-m_2)}{\sigma_1\sigma_2} + \frac{(x_2-m_2)^2}{\sigma_2^2}\right]}$$

Плотность нормального распределения $p(x_1, x_2)$ сохраняет постоянное значение на эллипсах:

$$\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-m_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-m_1)(x_2-m_2)}{\sigma_1\sigma_2} + \frac{(x_2-m_2)^2}{\sigma_2^2}\right] = c^2,$$

которые называются *эллипсами равных вероятностей.*

**Практика**

Обозначим $Q(x_1, x_2, c^2)$, $Q \subset R^2$, область, ограниченную эллипсом равных вероятностей. Ясно, что случайное событие $Q(x_1, x_2, c^2)$ - борелевское множество, то есть $Q(x_1, x_2, c^2) \in \mathcal{B}(R^2)$. Вероятность $P(Q(x_1, x_2, c^2))$ будет численно равна объёму «цилиндра», основанием которого является множество $Q(x_1, x_2, c^2)$, а длина высоты равна $c^2$, то есть вероятность зависит от величины $c^2$: $P(Q(x_1, x_2, c^2)) = \iint\limits_{Q(x_1, x_2, c^2)} p(x_1, x_2 dx_1 dx_2)$.

Переходя к полярным координатам, вычислим этот интеграл и получим:

$$P(Q(x_1, x_2, c^2)) = P(c^2) = 1 - e^{-c^2}.$$

В частности, если $c_1^2 = 1$, то $P(1) \approx 0{,}632$; если $c_1^2 = 2$, то $P(2) \approx 0{,}865$; если $c_1^2 = 4$, то $P(4) \approx 0{,}982$.

# ГЛАВА ТРЕТЬЯ
# Случайная величина
## §1.III Случайная величина
### Практика

Рассмотрим три примера.

Пример 1. Из колоды карт (36 шт.) наудачу, без возвращения извлекаются шесть карт. Определить вероятность того, что в полученном наборе карт будут две «картинки».

Пример 2. Экзаменационный билет содержит три из 30 вопросов экзаменационной программы. Студент успел хорошо подготовить 25 вопросов из этой программы. Какова вероятность того, что в наудачу выбранном билете будут два вопроса из числа хорошо подготовленных этим студентов вопросов?

Пример 3. Пусть $A_k$ - случайное событие: «В игре «Спортлото ***6*** из ***49***» угадано $k$ чисел». Определить вероятности наступления случайных событий $A_k$, если $k = 1,2,3,4,5,6$.

В этих примерах вероятности наступлений событий вычисляются с помощью классического определения вероятности. Для этого сначала определяются элементарные исходы, и подсчитывается общее количество исходов ($n$), которые могут реализоваться в результате проведения испытания. Проверяется равновозможность реализации при проведении испытания каждого из названных элементарных исходов. Затем записывается название случайного события ($A$), вероятность наступления которого надо определить, и подсчитывается количество исходов благоприятствующих наступлению этого случайного события ($\boldsymbol{m}$). Если элементарные исходы равновозможные, то применяем классическое определение вероятности: $\boldsymbol{P}(\boldsymbol{A})=\dfrac{\boldsymbol{m}}{\boldsymbol{n}}$.

Ответы: 1. $\boldsymbol{P}(\boldsymbol{A})=\dfrac{\boldsymbol{C}_{12}^{2}\cdot\boldsymbol{C}_{24}^{4}}{\boldsymbol{C}_{36}^{6}}$, 2. $\boldsymbol{P}(\boldsymbol{A})=\dfrac{\boldsymbol{C}_{25}^{2}\cdot\boldsymbol{C}_{5}^{1}}{\boldsymbol{C}_{30}^{3}}$, 3. $\boldsymbol{P}(\boldsymbol{A_k})=\dfrac{\boldsymbol{C}_{6}^{\boldsymbol{k}}\cdot\boldsymbol{C}_{43}^{6-\boldsymbol{k}}}{\boldsymbol{C}_{49}^{6}}$.

### *Комментарий*

Обратим внимание на общность этих примеров. Во всех примерах нас интересует количественный результат этих испытаний. При этом нам не важна природа наступившего события: содержание получившегося набора карт, названия вопросов взятого экзаменационного билета, конкретный перечень угаданных чисел из шести чисел, которые попали в тираж. Математический аппарат подсчёта, как общего числа возможных исходов испытания, так и числа исходов, которые благоприятствуют наступлению случайного события, - одинаков. Это – комбинаторика.

Эти выводы приводят нас к мысли о возможности введения некоторого общего понятия, которое не будет зависеть от «природы» элементарных исходов проводимого испытания, но будет фиксировать числовые результаты этого испытания. Таким понятием, объединяющее всё разнообразие проводимых испытаний по их количественному результату, определяемому общей математической моделью подсчёта, является понятие ***случайной величины***.

Определение термина «случайная величина» в теории вероятностей аналогично определению термина «функция» в математическом анализе. Напомним определение функции.

Пусть $\boldsymbol{X}$ и $\boldsymbol{Y}$ - два числовых множества. *Функция* – это *закон, правило*, по которому каждому элементу $\boldsymbol{x}$ множества $\boldsymbol{X}$ ставится в соответствие один, и только один, элемент $\boldsymbol{y}$ множества $\boldsymbol{Y}$. То есть функция $\boldsymbol{f}(\boldsymbol{x})$ - это отображение по некоторому правилу множества $\boldsymbol{X}$ во множество $\boldsymbol{Y}$. Кратко записывается так: $\boldsymbol{f}:\boldsymbol{X}\to\boldsymbol{Y}$. При этом не исключается, что двум разным элементам $\boldsymbol{x}_1$ и $\boldsymbol{x}_2$ множества $\boldsymbol{X}$ ставится в соответствие один и тот же элемент $\boldsymbol{y}$ множества $\boldsymbol{Y}$.

### Теория

Пусть $<\Omega,\mathcal{A},\boldsymbol{P}>$ - вероятностное пространство и $<\boldsymbol{R},\mathcal{B}(\boldsymbol{R})>$ измеримое пространство.

Определение. ***Случайной величиной*** *называется измеримое отображение множества элементарных исходов* $\Omega$ *во множество действительных чисел* $\boldsymbol{R}$.

Это отображение множества $\Omega$ во множество $\boldsymbol{R}$, называемое *случайной величиной*, будем обозначать $\xi(\omega)$ и кратко записывать $\xi(\omega):\Omega\to\boldsymbol{R}$.

Из определения следует, что, так как значениями случайной величины являются действительные числа, то *случайная величина* $\xi(\omega)$ *– это функция, реализующая наш количественный интерес к результату опыта*.

Любой элемент $\boldsymbol{B}$ алгебры борелевских множеств $\mathcal{B}(\boldsymbol{R})$, являясь подмножеством множества действительных чисел $\boldsymbol{R}$, рассматривается как случайное событие $\{\xi(\omega)\in \boldsymbol{B}\}$. У подмножества $\boldsymbol{B}$ отображением $\xi(\omega)\colon \Omega \to \boldsymbol{R}$ есть прообраз $\boldsymbol{A}=\xi^{-1}(\boldsymbol{B})$. Прообраз $\boldsymbol{A}$ является подмножеством множества элементарных исходов $\Omega$, то есть подмножество $\boldsymbol{A}$ – это элемент алгебры случайных событий $\mathcal{A}$ и мы можем определить $\boldsymbol{P}(\boldsymbol{A})$ - вероятность наступления случайного события $\boldsymbol{A}$.

Измеримость отображения $\xi(\omega)$ множества $\Omega$ во множество $\boldsymbol{R}$ мы понимаем так: *вероятность наступления случайного события* $\{\xi(\omega)\in \boldsymbol{B}\}$ *равна вероятности наступления случайного события* $\boldsymbol{A}$, *то есть* $\boldsymbol{P}(\xi(\omega)\in \boldsymbol{B})=\boldsymbol{P}(\boldsymbol{A})$.

То есть, задавая правило $\xi(\omega)$ отображения множества $\Omega$ во множество $\boldsymbol{R}$, мы на измеримом пространстве <$\boldsymbol{R}$,$\mathcal{B}(\boldsymbol{R})$> задаём вероятностную функцию $\boldsymbol{P}$, соответствующую аксиомам А.Н.Колмогорова, и получаем вероятностное пространство <$\boldsymbol{R}$,$\mathcal{B}(\boldsymbol{R})$,$\boldsymbol{P}$>. Вероятностную функцию $\boldsymbol{P}$ будем называть *распределением вероятностей случайной величины* $\xi(\omega)$.

***Комментарий***

Определяя отображение $\xi(\omega)\colon \Omega \to \boldsymbol{R}$, мы не можем записать его в аналитической форме $\xi = f(\omega)$, то есть подобно функции $y = f(x)$, как это делается в математическом анализе. Мы должны записывать чёткое название случайной величины, то есть *записывать словами* «правило», по которому каждому элементарному исходу $\omega$ из множества $\Omega$ ставится в соответствие число $\boldsymbol{x}$ из множества $\boldsymbol{R}$.

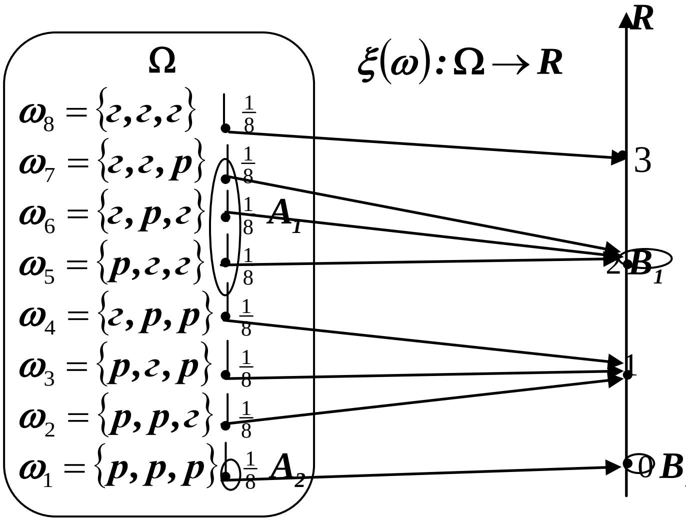
Сколько правил отображения $\xi(\omega)$, функций отображающих множество $\Omega$ во множество $\boldsymbol{R}$, мы придумаем, столько вероятностных пространств <$\boldsymbol{R}$,$\mathcal{B}(\boldsymbol{R})$,$\boldsymbol{P}$> мы построим.

***Комментарий***

Пример 4. Напомним испытание, которое мы рассматривали в самом начале знакомства с теорией вероятностей. Наудачу три раза подбрасывается монета. Был определён элементарный исход $\omega_i$: последовательность, состоящая из трёх букв «*г*» и «*р*», соответствующая результату трёх подбрасываний. Множество элементарных исходов состоит из восьми элементов: $\Omega=\{\omega_i\}$, $\boldsymbol{i}$=1,2,…,8. Все элементарные исходы – равновозможные, то есть $p(\omega_i)=\frac{1}{8}$. Алгебра $\mathcal{A}$ всех случайных событий, которые могут наступить при проведении этого испытания, состоит из $|\mathcal{A}|=2^{|\Omega|}=2^8$ элементов. Вероятностное пространство <$\Omega$,$\mathcal{A}$,$\boldsymbol{P}$> - построено.

Случайная величина $\xi(\omega)$ определяется целью проведения испытания. Если нас интересует только количество выпавших «гербов» и нам не важно, в каком порядке они выпадали, то определим случайную величину так:

$\xi(\omega)$ - количество «гербов» выпавших в результате трёх подбрасываний.




Ясно, что случайная величина $\xi(\omega)$ может принять три значения: 0,1,2 и 3. То есть восемь элементарных исходов множества $\boldsymbol{\Omega}$ отображаются в четыре точки множества $\boldsymbol{R}$.

Элементарные исходы – равновозможные, то есть $\boldsymbol{p}(\boldsymbol{\omega_i}) = \frac{1}{8}$. Измеримость отображения $\xi(\omega)$ позволяет определить, например, вероятности наступления событий: $\boldsymbol{B_1}$ - «*герб*» выпал два раза и $\boldsymbol{B_2}$ - «*герб*» ни разу не выпал. Множества $\boldsymbol{B_1}$ и $\boldsymbol{B_2}$ - одноточечные, борелевские , то есть - это элементы алгебры $\mathcal{B}(\boldsymbol{R})$. Прообразами этих множеств будут случайные события $\boldsymbol{A_1}$ и $\boldsymbol{A_2}$ - элементы алгебры $\mathcal{A}$. Так как $\boldsymbol{A_1} = \{\omega_5, \omega_6, \omega_7\}$ и $\boldsymbol{A}_2 = \{\omega_1\}$, то $\boldsymbol{P}(\xi(\omega) \in \boldsymbol{B}_1) = \boldsymbol{P}(\boldsymbol{A}_1) = \dfrac{3}{8}$ и $\boldsymbol{P}(\xi(\omega) \in \boldsymbol{B}_2) = \boldsymbol{P}(\boldsymbol{A}_2) = \dfrac{1}{8}$.

### Практика

Вернёмся к первым трём, ранее рассмотренным, примерам.

В примере 1, проводя испытание, то есть, выбирая указанным способом шесть карт из 36, мы определяем элементарный исход: «$\boldsymbol{\omega}$ - набор из шести карт различных номиналов». Множество элементарных исходов $\boldsymbol{\Omega}$ состоит из $\boldsymbol{n} = |\boldsymbol{\Omega}| = \boldsymbol{C}_{36}^{6}$ элементов. Случайное событие $\boldsymbol{A}$ формулируется так: «В получившемся наборе карт имеются две «картинки». Проверяя: наступило или нет случайное событие $\boldsymbol{A}$, мы будем подсчитывать количество «картинок» среди имеющегося набора из шести карт. То есть, нас интересует количественный результат опыта. Следовательно, случайную величину $\xi(\omega)$ - правило отображения множества элементарных исходов $\boldsymbol{\Omega}$ во множество действительных чисел $\boldsymbol{R}$, мы будем формулировать так: «$\xi(\omega)$ - *количество «картинок» среди извлечённых шести карт*».

Если случайная величина $\xi(\omega)$ примет значение 2, то мы будем говорить, что случайное событие $\boldsymbol{A}$ наступило.

Нас интересует вероятность $\boldsymbol{P}(\xi(\omega) \in \boldsymbol{B})$, где $\boldsymbol{B} = \{2\}$ - одноточечное подмножество множества $\boldsymbol{R}$. Прообразом одноточечного подмножества $\boldsymbol{B}=$ $\{\xi(\omega) = 2\}$ будет подмножество $\boldsymbol{A}$ – элемент алгебры $\mathcal{A}$, которое $\boldsymbol{A}$ состоит из $\boldsymbol{m} = |\boldsymbol{A}| = \boldsymbol{C}_{12}^{2} \cdot \boldsymbol{C}_{24}^{4}$ элементарных исходов, каждый из которых соответствует некоторому конкретному набору карт, содержащему две «картинки».

Все элементарные исходы $\boldsymbol{\omega_i}$ – равновозможные, поэтому $\boldsymbol{p}(\boldsymbol{\omega_i}) = \dfrac{1}{\boldsymbol{C}_{36}^{6}}$.

Следовательно, вероятность наступления случайного события $\boldsymbol{A}$ равна

$P(A)=\frac{C_{12}^2\cdot C_{24}^4}{C_{36}^6}$. Измеримость отображения $\xi(\omega):\Omega\to R$ означает, что случайная величина $\xi(\omega)$ примет значение 2 с вероятностью, равной вероятности наступления случайного события $A$: $P(\xi(\omega)=2)=P(A)=\frac{C_{12}^2\cdot C_{24}^4}{C_{36}^6}$.

В примере 2, проводя испытание, то есть, выбирая экзаменационный билет, содержащий три из 30 возможных вопросов, мы определяем элементарный исход: «$\omega$ - набор из трёх вопросов программы». Множество элементарных исходов $\Omega$ состоит из $n=|\Omega|=C_{30}^3$ элементов. Случайное событие $A$ формулируется так: «В выбранном билете имеются два хорошо подготовленных вопроса».

Нас интересует количественный результат опыта. Следовательно, случайную величину $\xi(\omega)$ - правило отображения множества элементарных исходов $\Omega$ во множество действительных чисел $R$, мы будем формулировать так: «$\xi(\omega)$ - *среди трёх вопросов билета есть два хорошо подготовленных*». Нам нужно определить вероятность наступления случайного события $B=\{\xi(\omega)=2\}$. Если случайная величина $\xi(\omega)$ примет значение 2, то мы будем говорить, что случайное событие $A$ наступило. Значит: $P(\xi(\omega)=2)=P(A)=\frac{C_{25}^2\cdot C_5^1}{C_{30}^3}$.

В примере 3, проводя испытание, то есть, заполняя карточку «Спортлото ***6*** из ***49***» мы отмечаем шесть чисел из 49. Элементарным исходом будет: «$\omega$ - набор из шести чисел». Множество элементарных исходов $\Omega$ состоит из $n=|\Omega|=C_{49}^6$ элементов. Случайное событие $A_k$, где $k$=0,1,2,3,4,5,6, формулируется так: «В заполненной карточке мы отгадали $k$ из шести чисел». Нас интересует количественный результат опыта. Следовательно, случайную величину $\xi(\omega)$ - правило отображения множества элементарных исходов $\Omega$ во множество действительных чисел $R$, мы будем формулировать так: «$\xi(\omega)$ - *среди шести чисел есть **k** отгаданных чисел*». Нам нужно определить вероятность наступления случайного события $B_k=\{\xi(\omega)=k\}$, где $k$=0,1,2,3,4,5,6. Значит: $P(\xi(\omega)=k)=P(A_k)=\frac{C_6^k\cdot C_{49-k}^{6-k}}{C_{49}^6}$.

### *Комментарий*

Все эти примеры показывают, что, когда мы имеем дело с какой-либо конкретной случайной величиной $\xi(\omega)$, мы можем забыть о природе элементарных исходов и самом множестве элементарных исходов $\Omega$. Это могут быть: число студентов, успешно сдавших последнюю сессию; количество гектаров, занятых посевами пшеницы; количество километров, которые проехал автомобилист в течение дня, количество израсходованных в течение месяца киловатт-часов электроэнергии; количество телефонных вызовов, поступивших на станцию скорой помощи, в течение одних суток.

Решая задачу определения вероятности наступления события, мы рассматриваем: числовую прямую $\boldsymbol{R}$, $\boldsymbol{\sigma}-$алгебру борелевских множеств $\mathcal{B}(\boldsymbol{R})$ и $\boldsymbol{\sigma}-$аддитивную вероятностную функцию.

Случайная величина $\xi(\omega)$ – это абстрактное понятие. При проведении испытания никто её не видит, не ощущает своими органами чувств. Случайная величина $\xi(\omega)$ проявляется в своих возможных значениях, которые мы видим, ощущаем органами своих чувств.

## Практика

Рассматривая в примере 4 испытание, в котором монета подбрасывается три раза, мы можем определить случайные величины: $\eta(\omega)$ - количество «гербов» выпавших в результате трёх подбрасываний кратно числу два; $\zeta(\omega)$ - остаток от деления количества выпавших «гербов» на число 3.

Естественно принять в качестве возможных значений случайной величины $\eta(\omega)$ одноточечные множества $\boldsymbol{B_0}=\{0\}$, если количество выпавших «гербов» не будет кратно числу два, и $\boldsymbol{B_1}=\{1\}$, если количество выпавших гербов будет кратно числу два. Прообразами этих множеств будут $\boldsymbol{A_0}=\boldsymbol{\eta}^{-1}(\boldsymbol{B}_0)=\{\boldsymbol{\omega}_2,\boldsymbol{\omega}_3,\boldsymbol{\omega}_4,\boldsymbol{\omega}_8\}$ и $\boldsymbol{A_1}=\boldsymbol{\eta}^{-1}(\boldsymbol{B}_1)=\{\boldsymbol{\omega}_1,\boldsymbol{\omega}_5,\boldsymbol{\omega}_6,\boldsymbol{\omega}_7\}$. Вероятностная функция $\boldsymbol{P}$ на алгебре $\mathcal{B}(\boldsymbol{R})$ принимает ненулевое значение на множествах $\boldsymbol{B_0}$ и $\boldsymbol{B_1}$: $\boldsymbol{P}(\eta(\omega)=0)=\boldsymbol{P}(\boldsymbol{B}_0)=\boldsymbol{P}(\boldsymbol{A}_0)=\dfrac{4}{8}$ и $\boldsymbol{P}(\eta(\omega)=1)=\boldsymbol{P}(\boldsymbol{B}_1)=\boldsymbol{P}(\boldsymbol{A}_1)=\dfrac{4}{8}$.

Случайная величина $\zeta(\omega)$ может принять значения $\boldsymbol{B_0}=\{0\}$, $\boldsymbol{B_1}=\{1\}$ и $\boldsymbol{B_2}=\{2\}$. Прообразами этих множеств будут $\boldsymbol{A_0}=\boldsymbol{\eta}^{-1}(\boldsymbol{B}_0)=\{\boldsymbol{\omega}_1,\boldsymbol{\omega}_8\}$, $\boldsymbol{A_1}=\boldsymbol{\eta}^{-1}(\boldsymbol{B}_1)=\{\boldsymbol{\omega}_2,\boldsymbol{\omega}_3,\boldsymbol{\omega}_4\}$ и $\boldsymbol{A_2}=\boldsymbol{\eta}^{-1}(\boldsymbol{B}_2)=\{\boldsymbol{\omega}_5,\boldsymbol{\omega}_6,\boldsymbol{\omega}_7\}$. Следовательно, вероятностная функция $\boldsymbol{P}$ принимает три значения: $\boldsymbol{P}(\zeta(\omega)=0)=\dfrac{2}{8}$, $\boldsymbol{P}(\zeta(\omega)=1)=\dfrac{3}{8}$ и $\boldsymbol{P}(\zeta(\omega)=2)=\dfrac{3}{8}$.

## §2.III Типы и примеры случайных величин

## Теория

Заявленное в формулировке случайной величины правило отображения множества элементарных исходов во множество действительных чисел определяет область определения и вид вероятностной функции $\boldsymbol{P}$, которая является третьим элементом вероятностного пространства <$\boldsymbol{R}$,$\mathcal{B}(\boldsymbol{R})$,$\boldsymbol{P}$>. Поэтому тип случайной величины определяется типом вероятностной функции $\boldsymbol{P}$.

<u>Определение 1.</u> Случайная величина $\xi(\omega)$ называется случайной величиной ***дискретного типа***, если множеством её возможных значений будет не более чем счётное множество действительных чисел – точек числовой прямой $\{\boldsymbol{x_k}\}$, $\boldsymbol{k}$=1,2,3,…,$\boldsymbol{n}$,… $\infty$. Значениями $\boldsymbol{P}(\xi(\omega)=\boldsymbol{x_k})=\boldsymbol{p_k}$ случайной величины является

набор положительных чисел $\{p_k\}$, $k$=1,2,3,…,$n$,… $\infty$. При этом $\sum\limits_{k=1}^{n,\infty} p_k = 1$.

Набор чисел $\{p_k\}$, $k$=1,2,3,…,$n$,… $\infty$, называется *дискретным распределением вероятностей.* Наступление события $B$, являющегося элементом алгебры борелевских множеств $\mathcal{B}(R)$, означает, что случайная величина $\xi(\omega)$ приняла значение $x_k$, которое принадлежит $B$. Таких значений может быть несколько, а поэтому $P(\xi(\omega) \in B) = \sum\limits_{x_k \in B} p_k$.

***Комментарий***

Ясно, что случайные величины, рассмотренные в четырёх примерах предыдущего параграфа **§1.III,** являются случайными величинами дискретного типа.

Для задания случайной величины дискретного типа необходимо записать множество её возможных значений $\{x_k\}$ и вероятности $\{p_k\}$, с которыми случайная величина $\xi(\omega)$ принимает эти значения. То есть вся информация о случайной величине может быть записана в виде таблицы, в первой строке которой записываются возможные значения $\xi(\omega)$, а во второй – соответствующие вероятности. Такая таблица называется *рядом распределения* случайной величины $\xi(\omega)$ дискретного типа.

| $\xi(\omega)$ | $x_1$ | $x_2$ | $x_3$ | ….. | $x_n$ | ….. |
|---|---|---|---|---|---|---|
| $p_k$ | $p_1$ | $p_2$ | $p_3$ | ….. | $p_n$ | ….. |

Примеры законов распределений вероятностей случайных величин дискретного типа  наиболее часто встречающихся в практических задачах приведены в **§2.II.1**.

**Теория**

<u>Определение 2.</u> Случайная величина $\xi(\omega)$ называется случайной величиной ***непрерывного типа***, если множеством её возможных значений будет множество $G$ действительных чисел ($G \subseteq R$), мощность которого равна мощности континуум $c$, а мера Лебега этого множества – положительное число, то есть $mesG>0$. Распределение вероятностей возможных значений случайной величины $\xi(\omega)$ определяется неотрицательной, кусочно-непрерывной функцией $p(x)$, называемой *плотностью вероятности*. При этом всегда $\int\limits_G p(x)dx = 1$.

Аналогично определению вероятности наступления события $B$ для дискретной случайной величины, определяется вероятность наступления события $B$ для случайной величины непрерывного типа: $P(\xi(\omega) \in B) = \int\limits_B p(x)dx$.

Примеры распределений вероятностей случайных величин непрерывного типа  наиболее часто встречающихся в практических задачах приведены в **§2.II.2**.

***Комментарий***

На измеримом пространстве $\langle R, \mathcal{B}(R) \rangle$ определяются вероятностные функции трёх типов: дискретная, абсолютно непрерывная и сингулярная. Ясно, что случайной величиной сингулярного типа будет случайная величина, множество значений которой будет борелевским множеством, мощность которого равна мощности континуум, а мера Лебега – равна нулю. Записать аналитически распределение вероятностей сингулярной случайной величины мы не можем.

## §3.III Функция распределения случайной величины

### Теория

Пусть $\xi$ - случайная величина и $<\boldsymbol{R},\mathcal{B}(\boldsymbol{R}),\boldsymbol{P}>$ - соответствующее ей вероятностное пространство. Обозначим $\boldsymbol{S_x}$ множество $(-\infty;\boldsymbol{x})$, где $\boldsymbol{x}$ – любое действительное число. Ясно, что $\boldsymbol{S_x}=(-\infty;\boldsymbol{x})$ - борелевское множество, то есть $\boldsymbol{S_x}\in\mathcal{B}(\boldsymbol{R})$. Значит, для любого значения $\boldsymbol{x}$ мы можем определить вероятность наступления события $\{\xi\in\boldsymbol{S_x}\}$, то есть вычислить: $\boldsymbol{P}(\xi\in\boldsymbol{S_x})$. Эта вероятность будет функцией аргумента $\boldsymbol{x}$, обозначим её $\boldsymbol{F}(\boldsymbol{x})$.

Определение. Функцию $\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{P}(\xi(\omega)\in\boldsymbol{S_x})$ будем называть - ***функцией распределения*** *случайной величины* $\xi$.

Так определённая функция $\boldsymbol{F}(\boldsymbol{x})$ обладает следующими свойствами:

1) функция $\boldsymbol{F}(\boldsymbol{x})$ - неубывающая функция;

2) функция $\boldsymbol{F}(\boldsymbol{x})$ - непрерывная слева, то есть для любого $\boldsymbol{x}\in\boldsymbol{R}$ справедливо $\boldsymbol{F}(-\boldsymbol{x})=\boldsymbol{F}(\boldsymbol{x})$;

3) $\lim\limits_{x\downarrow-\infty}\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{F}(-\infty)=0$ и $\lim\limits_{x\uparrow+\infty}\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{F}(+\infty)=1$.

Доказательство первого свойства очевидно. Если $\boldsymbol{x_1}<\boldsymbol{x_2}$, то $\boldsymbol{S_{x_2}}=\boldsymbol{S_{x_1}}\bigcup[\boldsymbol{x_1};\boldsymbol{x_2})$. Значит $\boldsymbol{F}(\boldsymbol{x_2})=\boldsymbol{P}(\xi\in\boldsymbol{S_{x_2}})=\boldsymbol{P}(\xi\in\boldsymbol{S_{x_1}}\bigcup[\boldsymbol{x_1};\boldsymbol{x_2}))$, а так как вероятностная функция $\boldsymbol{P}$ – аддитивная функция, то

$$\boldsymbol{F}(\boldsymbol{x_2})=\boldsymbol{P}(\xi\in\boldsymbol{S_{x_1}})+\boldsymbol{P}(\xi\in[\boldsymbol{x_1};\boldsymbol{x_2}))=\boldsymbol{F}(\boldsymbol{x_1})+\boldsymbol{P}(\xi\in[\boldsymbol{x_1};\boldsymbol{x_2})).$$

А так как значение любой вероятности всегда – неотрицательное число, то есть всегда: $\boldsymbol{P}(\xi\in[\boldsymbol{x_1};\boldsymbol{x_2}))\geq 0$, то $\boldsymbol{F}(\boldsymbol{x_2})\geq\boldsymbol{F}(\boldsymbol{x_1})$.

Доказательства второго и третьего свойства основаны на свойствах непрерывности вероятностной функции $\boldsymbol{P}$.

Пусть $\boldsymbol{x}\in\boldsymbol{R}$, $\boldsymbol{S_x}\in\mathcal{B}(\boldsymbol{R})$, и $\{\boldsymbol{x_k}\}$ – произвольная числовая последовательность такая что $\boldsymbol{x_k}\uparrow\boldsymbol{x}$. Эта последовательность индуцирует нам расширяющуюся последовательность множеств $\{\boldsymbol{S_{x_k}}\}$: $\boldsymbol{S_{x_k}}\subset\boldsymbol{S_{x_{k+1}}}$, при этом $\bigcup\limits_{k=1}^{\infty}\boldsymbol{S_{x_k}}=\boldsymbol{S_x}$. Вероятностная функция $\boldsymbol{P}$ – *непрерывная «сверху»* функция.

Следовательно, $\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{P}(\boldsymbol{S_x})=\lim\limits_{x_k\uparrow x}\boldsymbol{P}(\boldsymbol{S_{x_k}})=\lim\limits_{x_k\uparrow x}\boldsymbol{F}(\boldsymbol{x_k})$. Так как числовая последовательность $\{\boldsymbol{x_k}\}$ – произвольная, то, по определению Гейне непрерывности функции, заключаем: $\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{F}(\boldsymbol{x}-0)$.

Пусть $\{\boldsymbol{x_k}\}$ – произвольная числовая последовательность такая что $\boldsymbol{x_k}\downarrow-\infty$.

Эта последовательность индуцирует нам сужающуюся последовательность множеств $\{S_{x_k}\}$: $S_{x_k} \supset S_{x_{k+1}}$, при этом $\bigcap_{k=1}^{\infty} S_{x_k} = \varnothing$. Вероятностная функция $P$ – *непрерывная «в нуле»* функция. Следовательно, учитывая, что последовательность $\{x_k\}$ – произвольная, получаем:

$$F(-\infty) = \lim_{x_k \downarrow -\infty} P(S_{x_k}) = P(\varnothing) = 0.$$

Выбирая произвольную последовательность $\{x_k\}$ такую что $x_k \uparrow +\infty$, индуцируем расширяющуюся последовательность множеств $\{S_{x_k}\}$, где $\bigcup_{k=1}^{\infty} S_{x_k} = R$. Так как функция $P$ – *непрерывная «сверху»* и нормированная функция, то: $F(+\infty) = P(R) = 1$.

**Практика**

Заметим, что, так как $S_x = (-\infty; x)$, то случайное событие $\{\xi \in S_x\}$ - «значение случайной величины $\xi$ принадлежит множеству $S_x$» можно записать так $\{\xi < x\}$ - «значение случайной величины $\xi$ меньше числа $x$». Тогда определение функции распределения случайной величины $\xi$ можно записать в упрощённой форме: $F(x) = P(\xi < x)$.

Построим примеры функций распределения $F(x)$ случайных величин дискретного и непрерывного типа.

Пример 1. Пусть случайная величина $\xi$ - количество «гербов» выпавших при двух подбрасываниях монеты. Закон распределения случайной величины $\xi$ дискретного типа задаётся в виде ряда распределения. Запишем для рассматриваемого испытания ряд распределения этой случайной величины:

| $\xi(\omega)$ | 0 | 1 | 2 |
|---|---|---|---|
| $p_k$ | $\frac{1}{4}$ | $\frac{2}{4}$ | $\frac{1}{4}$ |

Областью определения функции распределения $F(x)$ будет вся числовая ось. Ясно, что:

1) если $x \le 0$, то вероятность события $\{\xi < x\}$ равна нулю, то есть $F(x) = 0$;

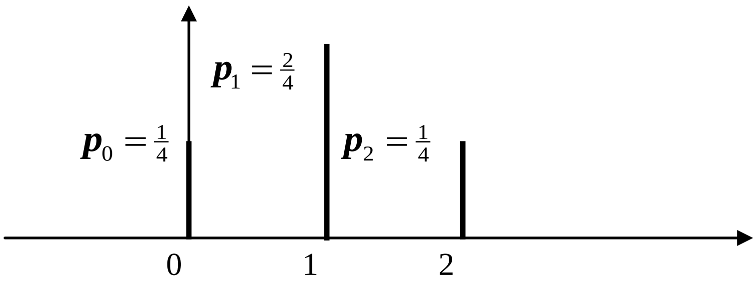
2) если $0 < x \le 1$, то вероятность события $\{\xi < x\}$ будет равна вероятности события $\{\xi = 0\}$, то есть для $0 < x \le 1$ функция распределения принимает значение $F(x) = p_0 = \frac{1}{4}$;

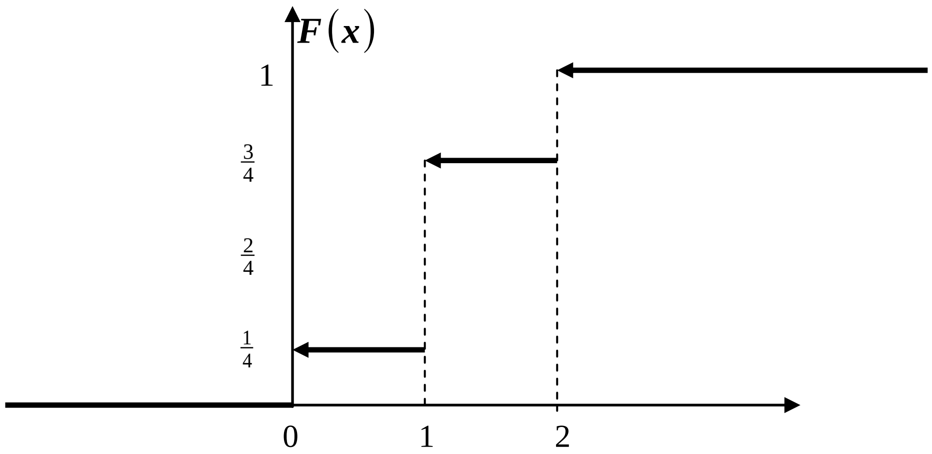
3) если $1 < x \le 2$, то вероятность события $\{\xi < x\}$ будет равна сумме

вероятностей событий $\{\xi = 0\}$ и $\{\xi = 1\}$, то есть для $1 < x \le 2$ функция распределения принимает значение $F(x) = p_0 + p_1 = \frac{1}{4} + \frac{2}{4} = \frac{3}{4}$;

4) если $2 < x$, то вероятность события $\{\xi < x\}$ будет равна сумме вероятностей событий $\{\xi = 0\}$, $\{\xi = 1\}$ и $\{\xi = 2\}$, то есть для $x > 2$ функция распределения принимает значение $F(x) = p_0 + p_1 + p_2 = \frac{1}{4} + \frac{2}{4} + \frac{1}{4} = 1.$

Сделаем геометрическую иллюстрацию закона распределения вероятностей и функции распределения случайной величины $\xi$.




Все три свойства функции распределения в этом примере – выполнены.

Анализируя процесс построения и вид графика построенной функции распределения $F(x)$, можем перечислить особенности функции распределения дискретной случайной величины. У случайной величины дискретного типа функция распределения кусочно-непрерывная. График функции $F(x)$ имеет разрывы первого рода. Абсциссами точек разрыва являются возможные значения рассматриваемой случайной величины: $x_0 = 0,\ \ x_1 = 1,\ \ x_2 = 2$. Величины «скачков» графика функции в точках разрыва равны соответствующим вероятностям этих значений случайной величины: $p_0 = \frac{1}{4},\ \ p_1 = \frac{2}{4},\ \ p_2 = \frac{1}{4}$.

Пример 2. Пусть случайная величина $\xi$ - непрерывного типа и имеет равномерное распределение вероятностей на области $\boldsymbol{G}$. Это означает, что, если областью её возможных значений $\boldsymbol{G}$ является сегмент $[\boldsymbol{a};\boldsymbol{b}]$, то плотность вероятности задаётся двойным равенством: $\boldsymbol{p}(\boldsymbol{x})=\begin{cases} 0, & \boldsymbol{x}\notin[\boldsymbol{a},\boldsymbol{b}] \\ \dfrac{1}{\boldsymbol{b}-\boldsymbol{a}} & \boldsymbol{x}\in[\boldsymbol{a},\boldsymbol{b}] \end{cases}$.

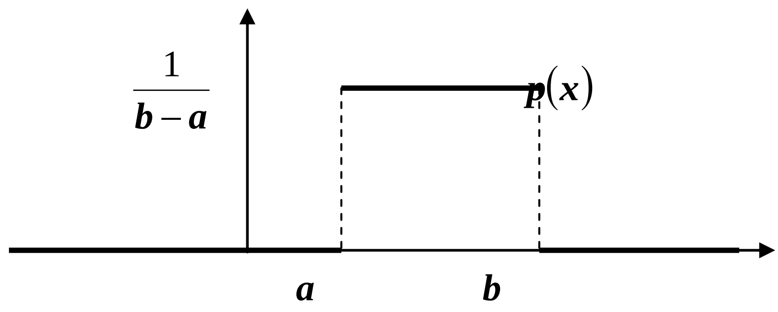
Руководствуясь определением функции распределения, определим её вид для рассматриваемой плотности $\boldsymbol{p}(\boldsymbol{x})$.

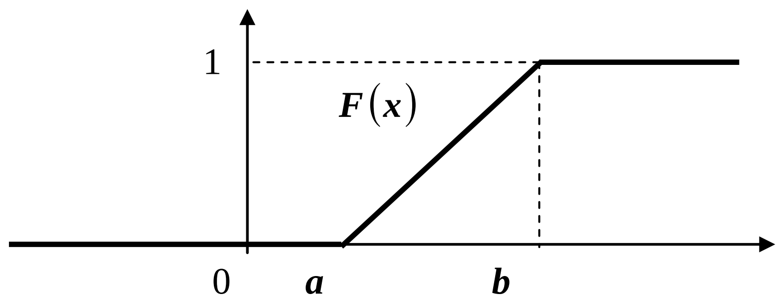
1) Если $\boldsymbol{x}\le\boldsymbol{a}$, то $\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{P}(\xi<\boldsymbol{x})=\int\limits_{-\infty}^{x}0\boldsymbol{dt}=0$.

2) Если $\boldsymbol{a}<\boldsymbol{x}\le\boldsymbol{b}$, то $\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{P}(\xi<\boldsymbol{x})=\int\limits_{-\infty}^{a}0\boldsymbol{dt}+\int\limits_{a}^{x}\frac{1}{\boldsymbol{b}-\boldsymbol{a}}\boldsymbol{dt}=\frac{\boldsymbol{x}-\boldsymbol{a}}{\boldsymbol{b}-\boldsymbol{a}}$.

3) Если $\boldsymbol{x}>\boldsymbol{b}$, то $\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{P}(\xi<\boldsymbol{x})=\int\limits_{-\infty}^{a}0\boldsymbol{dt}+\int\limits_{a}^{b}\frac{1}{\boldsymbol{b}-\boldsymbol{a}}\boldsymbol{dt}+\int\limits_{b}^{x}0\boldsymbol{dt}=\frac{\boldsymbol{b}-\boldsymbol{a}}{\boldsymbol{b}-\boldsymbol{a}}=1$.

Сделаем геометрическую иллюстрацию плотности вероятности $\boldsymbol{p}(\boldsymbol{x})$ и функции распределения $\boldsymbol{F}(\boldsymbol{x})$ случайной величины $\xi$.




Все три свойства функции распределения – выполнены.

**Практика**

Пример1.3.III Интервал движения автобуса некоторого маршрута составляет $\boldsymbol{a}$ минут. Обозначим случайную величину $\xi$ - время прибытия пассажира на место остановки автобуса. Построить закон распределения вероятностей случайной величины $\xi$. Определить вероятности событий: *а)* время ожидания автобуса будет в интервале от двух до четырёх минут; *б)* время ожидания автобуса будет в интервале от пяти до семи минут, если $\boldsymbol{a}$=10.

Так как мы не знаем: сколько минут прошло после отхода от остановки предыдущего автобуса, то областью возможных значений случайной величины $\xi$ будет сегмент $G(x)=[0;a]\subset R_1$. Любой диапазон времени ожидания $\xi\in[x;x+\Delta x)$, принадлежащий сегменту $G(x)=[0;a]$, равновозможен и не зависит от его расположении в интервале $G(x)=[0;a]$, а поэтому плотность вероятности $p(x)$ случайной величины $\xi$ будет иметь вид:

$$p(x)=\begin{cases}0, & x\notin[0;a]\\ \dfrac{1}{a} & x\in[0,a]\end{cases} \quad \text{или, если } a=10, \quad p(x)=\begin{cases}0, & x\notin[0;10]\\ \dfrac{1}{10} & x\in[0,10]\end{cases}.$$

Функция распределения $F(x)$ случайной величины $\xi$ будет иметь вид:

$$F(x)=\begin{cases}0, & \text{если} \quad x\le 0;\\ \dfrac{x}{a}, & \text{если } 0<x\le a;\\ 1, & \text{если} \quad x>a.\end{cases}$$

Так как вероятность любого события - интервала $[\alpha;\beta)$ - равна приращению значения функции распределения на этом интервале: $P(\alpha\le\xi<\beta)=F(\beta)-F(\alpha)$, то очевидно, что $P(2\le\xi<4)=P(5\le\xi<7)=\dfrac{2}{10}$. Вероятность того, что выбранная длительность интервала времени ожидания автобуса составит 2 минуты, не зависит от положения этого интервала на отрезке $[0;10]$.

***Комментарий***

У случайной величины непрерывного типа функция распределения *абсолютно непрерывная функция*.

Напомним определение абсолютно непрерывной функции.

Функция $f(x)$, определённая на сегменте $[a;b]$, называется абсолютно непрерывной функцией, если для любого, сколь угодно малого, $\varepsilon>0$ существует $\delta>0$ такое, что для любой конечной или счётной системы попарно непересекающихся интервалов $\{(a_i;b_i)\}$, принадлежащих $[a;b]$, для которой $\sum\limits_i(b_i-a_i)<\delta$, будет выполняться $\sum\limits_i|f(b_i)-f(a_i)|<\varepsilon$.

Ясно, что если в этом определении вместо системы попарно непересекающихся интервалов рассматривать только один произвольный интервал $(\alpha;\beta)$, содержащийся в $[a;b]$, и требовать для $(\beta-\alpha)<\delta$ выполнения $|f(\beta)-f(\alpha)|<\varepsilon$, то мы получим обычное определение Коши непрерывной функции $f(x)$, определённой на отрезке $[a;b]$.

Если случайная величина $\xi$ сингулярного типа, то мы не можем записать аналитически её вероятностную функцию $P$, а, следовательно, и функцию распределения $F(x)$. Однако, приведённый в **§2.II.1** процесс построения вероятностной функции позволяет, по аналогии с процессом построения функции распределения дискретной случайной величины, провести процесс построения графика функции распределения сингулярной случайной величины. Получающийся график функции $F(x)$ называется «канторовой лестницей». Эта функция распределения будет *непрерывной функцией*.

## Теория

Определение функции распределения $\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{P}(\xi(\omega)\in \boldsymbol{S}_x)$ позволяет утверждать, что каждой вероятностной функции $\boldsymbol{P}$ соответствует некоторая функция распределения $\boldsymbol{F}(\boldsymbol{x})$. Справедливо и обратное утверждение.

Пусть $\boldsymbol{F}(\boldsymbol{x})$ - функция распределения, определённая на числовой прямой $\boldsymbol{R}$. Тогда на измеримом пространстве $<\boldsymbol{R},\mathcal{B}(\boldsymbol{R})>$ по ней можно построить и притом единственным образом вероятностную функцию $\boldsymbol{P}$.

Вспомним, что при построении $\boldsymbol{\sigma}-$алгебры борелевских множеств $\mathcal{B}(\boldsymbol{R})$ мы в качестве системы подмножеств $\boldsymbol{\mathsf{A}}$ множества $\boldsymbol{R}$ взяли систему всех полуинтервалов $\{[\boldsymbol{a},\boldsymbol{b})\}$, где $\boldsymbol{a}$ и $\boldsymbol{b}$ – действительные числа. Применяя к элементам этой системы не более чем счётное число операций объединения, пересечения и дополнения, мы получали различные множества, являющиеся подмножествами множества $\boldsymbol{R}$. Эти множества называются борелевскими, совокупность всех таких множеств мы обозначили $\mathcal{B}(\boldsymbol{R})$ и назвали *алгеброй борелевских множеств.* Алгебра $\mathcal{B}(\boldsymbol{R})$ - минимальная $\boldsymbol{\sigma}$-алгебра, содержащая систему $\boldsymbol{\mathsf{A}}$.

Чтобы показать процесс построения вероятностной функции $\boldsymbol{P}$ по функции распределения $\boldsymbol{F}(\boldsymbol{x})$, сначала определим для любого элемента $[\boldsymbol{a};\boldsymbol{b})$ системы подмножеств $\boldsymbol{\mathsf{A}}$ его вероятностную *меру*: $\boldsymbol{P}([\boldsymbol{a};\boldsymbol{b}))=\boldsymbol{F}(\boldsymbol{b})-\boldsymbol{F}(\boldsymbol{a})$. То есть вероятность случайного события $\{[\boldsymbol{a},\boldsymbol{b})\}$ - это приращение значения функции $\boldsymbol{F}(\boldsymbol{x})$ на полуинтервале $[\boldsymbol{a};\boldsymbol{b})$.

Пополним систему подмножеств $\boldsymbol{\mathsf{A}}$ до алгебры $\mathcal{A}$, добавив множества, являющиеся конечными суммами непересекающихся полуинтервалов системы $\boldsymbol{\mathsf{A}}$, то есть любой элемент алгебры $\mathcal{A}$ представим в виде $\boldsymbol{A}=\sum\limits_{\boldsymbol{k}=1}^{\boldsymbol{n}}[\boldsymbol{a}_{\boldsymbol{k}};\boldsymbol{b}_{\boldsymbol{k}})$.

Определим на этой алгебре функцию множеств $\boldsymbol{P}_0(\boldsymbol{A})=\sum\limits_{\boldsymbol{k}=1}^{\boldsymbol{n}}(\boldsymbol{F}(\boldsymbol{b}_{\boldsymbol{k}})-\boldsymbol{F}(\boldsymbol{a}_{\boldsymbol{k}}))$.

Функция $\boldsymbol{P}_0$ - конечно-аддитивная и нормированная функция множеств из $\mathcal{A}$. Можно показать, что функция $\boldsymbol{P}_0$ – *непрерывная «в нуле»* функция. Следовательно, функция $\boldsymbol{P}_0$ - $\boldsymbol{\sigma}-$*аддитивная.*

Существование и единственность вероятностной функции $\boldsymbol{P}$, определяемой функцией распределения $\boldsymbol{F}(\boldsymbol{x})$, вытекает из теоремы Каратеодори о продолжении меры. Суть этой теоремы состоит в следующем: Если $<\boldsymbol{R},\mathcal{A},\boldsymbol{P}_0>$ вероятностное пространство и $\mathcal{B}(\boldsymbol{R})$ наименьшая $\boldsymbol{\sigma}$-алгебра, содержащая алгебру $\mathcal{A}$, то существует и при том единственная вероятностная функция $\boldsymbol{P}$ на $<\boldsymbol{R},\mathcal{B}(\boldsymbol{R})>$, являющаяся продолжением функции $\boldsymbol{P}_0$ с алгебры $\mathcal{A}$ на $\boldsymbol{\sigma}$-алгебру $\mathcal{B}(\boldsymbol{R})$, то есть такая, что $\boldsymbol{P}(\boldsymbol{A})=\boldsymbol{P}_0(\boldsymbol{A})$, если $\boldsymbol{A}\in\mathcal{A}$.

Возможность практического определения $\boldsymbol{P}(\boldsymbol{A})$ путём применением определения $\boldsymbol{P}([\boldsymbol{a;b}))=\boldsymbol{F}(\boldsymbol{b})-\boldsymbol{F}(\boldsymbol{a})$ для любого случайного события $\boldsymbol{A}$, являющегося элементом $\boldsymbol{\sigma}$-алгебры $\mathcal{B}(\boldsymbol{R})$, хорошо просматривается, если вспомнить метод построения элементов алгебры $\mathcal{B}(\boldsymbol{R})$. Этот метод заключается в том, что мы применяем не более чем счётное число операций объединения, пересечения и дополнения к элементам системы подмножеств $\boldsymbol{A}$ или, что очевидно, к элементам алгебры $\mathcal{A}$.

Определив по функции распределения $\boldsymbol{F}(\boldsymbol{x})$ вероятностную функцию $\boldsymbol{P}$: $\boldsymbol{P}([\boldsymbol{a;b}))=\boldsymbol{F}(\boldsymbol{b})-\boldsymbol{F}(\boldsymbol{a})$, мы строим функцию, которая называется вероятностной функцией Лебега – Стилтьеса на <$\boldsymbol{R}$,$\mathcal{B}(\boldsymbol{R})$>. В частности для случайных событий $(\boldsymbol{a,b})$, $(\boldsymbol{a,b}]$, $[\boldsymbol{a,b}]$ и $\{\boldsymbol{a}\}$, построенных по элементам системы подмножеств $\boldsymbol{A}$, получим:

$$\boldsymbol{P}((\boldsymbol{a,b}))=\boldsymbol{F}(\boldsymbol{b})-\boldsymbol{F}(\boldsymbol{a}+0);\quad \boldsymbol{P}((\boldsymbol{a,b}])=\boldsymbol{F}(\boldsymbol{b}+0)-\boldsymbol{F}(\boldsymbol{a}+0);$$
$$\boldsymbol{P}([\boldsymbol{a,b}])=\boldsymbol{F}(\boldsymbol{b}+0)-\boldsymbol{F}(\boldsymbol{a});\quad \boldsymbol{P}(\{\boldsymbol{a}\})=\boldsymbol{F}(\boldsymbol{a}+0)-\boldsymbol{F}(\boldsymbol{a}).$$

Эти различия в определении вероятностей перечисленных событий надо иметь в виду, если случайная величина $\xi$ - дискретного типа.

Если функция распределения – абсолютно непрерывная, то есть, если $\boldsymbol{F}(\boldsymbol{x})$ - функция распределения случайной величины $\xi$ непрерывного типа, то во всех случаях:

$$\boldsymbol{P}((\boldsymbol{a,b}))=\boldsymbol{P}((\boldsymbol{a,b}])=\boldsymbol{P}([\boldsymbol{a,b}])=\boldsymbol{P}((\boldsymbol{a,b}])=\boldsymbol{F}(\boldsymbol{b})-\boldsymbol{F}(\boldsymbol{a}) \text{ и } \boldsymbol{P}(\{\boldsymbol{a}\})=\boldsymbol{F}(\boldsymbol{a})-\boldsymbol{F}(\boldsymbol{a})=0.$$

## §4.III Многомерная случайная величина

### *Комментарий*

Для определения важнейшего понятия теории вероятностей – понятия **случайной величины**, мы использовали аналогию понятия функции как *отображение множества* $\boldsymbol{X}$ *во множество* $\boldsymbol{Y}$, с понятием случайной величины как *отображение множества элементарных исходов* $\boldsymbol{\Omega}$ *во множество* $\boldsymbol{R}$. По умолчанию мы полагали, что множество $\boldsymbol{Y}$ имеет размерность равную единице, и использовали его аналогию с множеством действительных чисел $\boldsymbol{R}$, у которого размерность тоже равна единице.

Естественно, что в некоторых экспериментах необходимо определить отображение множества $\boldsymbol{\Omega}$ во множество $\boldsymbol{R}^n$ – $\boldsymbol{n}$-мерное евклидово пространство, которое мы понимаем как декартово произведение $\boldsymbol{n}$ экземпляров числовых осей: $\boldsymbol{R}^n=\boldsymbol{R}_1\times\boldsymbol{R}_2\times\ldots\times\boldsymbol{R}_n$.

Например. В испытании, в котором три раза подбрасывается монета, мы определили элементарный исход $\boldsymbol{\omega_i}$ как последовательность, состоящую из трёх букв «*г*» и «*р*». Построенное множество элементарных исходов состоит из восьми элементов: $\boldsymbol{\Omega}=\{\boldsymbol{\omega_i}\}$, $\boldsymbol{i}$=1,2,…,8. Все элементарные исходы – равновозможные, то есть $\boldsymbol{p}(\boldsymbol{\omega_i})=\dfrac{1}{8}$. Мы можем определить отображение $\xi$ этого множества $\boldsymbol{\Omega}$ в трёхмерное пространство $\boldsymbol{R}^3$, ставя в соответствие каждому $\boldsymbol{\omega_i}$ точку трёхмерного пространства $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})$, у которой каждая координата принимает значение 1, если при соответствующем подбрасывании выпал «*герб*» и принимает значение 0, если выпала «*решка*». Ясно, что образом множества $\Omega$ будут восемь вершин единичного куба в $\boldsymbol{R}^3$: $(0,0,0)$; $(0,0,1)$; $(0,1,0)$; ….; $(1,1,1)$.

**Теория**

Пусть $<\Omega, \mathcal{A}, \boldsymbol{P}>$ - вероятностное пространство и $<\boldsymbol{R}^n, \mathcal{B}(\boldsymbol{R}^n)>$ измеримое пространство.

<u>Определение.</u> ***Многомерной случайной величиной*** называется *измеримое отображение* множества элементарных исходов $\Omega$ в $\boldsymbol{n}$-мерное евклидово пространство $\boldsymbol{R}^n$.

Как и в одномерном случае, это отображение множества $\Omega$ в пространство $\boldsymbol{R}^n$ будем обозначать $\xi$ и кратко записывать $\xi : \Omega \to \boldsymbol{R}^n$.

Из определения следует, что значениями многомерной случайной величины $\xi$ будут точки $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n)$ $\boldsymbol{n}$-мерного евклидова пространства $\boldsymbol{R}^n$. Согласно определению для любого $\boldsymbol{B} \in \mathcal{B}(\boldsymbol{R}^n)$ вероятность события $\{\xi \in \boldsymbol{B}\}$ равна вероятности события $\boldsymbol{A} \in \mathcal{A}$, где $\boldsymbol{A} = \xi^{-1}(\boldsymbol{B})$. То есть, на измеримом пространстве $<\boldsymbol{R}^n, \mathcal{B}(\boldsymbol{R}^n)>$ мы получаем вероятностную функцию $\boldsymbol{P}$ такую, что $\boldsymbol{P}(\xi \in \boldsymbol{B}) = \boldsymbol{P}(\boldsymbol{A})$. Эту функцию $\boldsymbol{P}$ будем называть *распределением вероятностей многомерной случайной величины* $\xi$.

Пусть функция $\pi_k$ отображает пространство $\boldsymbol{R}^n$ в $\boldsymbol{k}$-тую координатную ось $\boldsymbol{R}_k$, то есть $\pi_k : \boldsymbol{R}^n \to \boldsymbol{R}_k$. Функцию $\pi_k$ называют *проекцией* (проектор) пространства $\boldsymbol{R}^n$ на $\boldsymbol{k}$-тую координатную ось $\boldsymbol{R}_k$.

Суперпозиция отображений $\pi_k \circ \xi(\omega) : \Omega \to \boldsymbol{R}^n \to \boldsymbol{R}_k$ будет привычной нам одномерной случайной величиной $\xi_k$, осуществляющей измеримое отображение множества $\Omega$ в $\boldsymbol{R}_k$. Перебирая $\boldsymbol{k} = 1,2,...,\boldsymbol{n}$, получим $\boldsymbol{n}$ одномерных случайных величин $\xi_k$. Представим многомерную случайную величину в векторной форме $\xi = (\xi_1, \xi_2, ..., \xi_n)$. Многомерную случайную величину $\xi$ будем называть *векторной* случайной величиной, а случайные величины $\xi_k$ - *компонентами* векторной случайной величины $\xi$.

Измеримость отображения множества $\Omega$ в пространство $\boldsymbol{R}^n$ определяет вероятностную функцию $\boldsymbol{P}$ векторной случайной величины $\xi$, то есть, определив случайную величину $\xi$, мы получаем вероятностное пространство $<\boldsymbol{R}^n, \mathcal{B}(\boldsymbol{R}^n), \boldsymbol{P}>$.

Тип вероятностной функции $\boldsymbol{P}$ определяется формулировкой правила отображения множества $\Omega$ в пространство $\boldsymbol{R}^n$. Векторная случайная величина может быть любого из трёх типов: дискретного, непрерывного и сингулярного.

Зная вероятностную функцию $\boldsymbol{P}$ векторной случайной величины $\xi$, мы можем определить вероятностные функции её компонент $\{\xi_k\}$, $\boldsymbol{k} = 1,2,...,\boldsymbol{n}$. Ясно, что множество $\widetilde{\boldsymbol{B}}_k = \boldsymbol{R}_1 \times \boldsymbol{R}_2 \times ... \times \boldsymbol{B}_k \times ... \times \boldsymbol{R}_n$, где $\boldsymbol{B}_k \in \mathcal{B}(\boldsymbol{R}_k)$, является борелевским множеством, то есть, $\widetilde{\boldsymbol{B}}_k$ - элемент $\sigma-$алгебры $\mathcal{B}(\boldsymbol{R}^n)$. Так как никаких ограничений на возможные значения всех компонент, кроме $\boldsymbol{k}$-той, случайной величины $\xi$ нет, то вероятность наступления случайного события $\left\{\xi \in \widetilde{\boldsymbol{B}}_{\boldsymbol{k}} = \left(\boldsymbol{R}_1 \times \boldsymbol{R}_2 \times \ldots \times \boldsymbol{B}_{\boldsymbol{k}} \times \ldots \times \boldsymbol{R}_{\boldsymbol{n}}\right)\right\}$ равна вероятности наступления случайного события $\left\{\xi_{\boldsymbol{k}} \in \boldsymbol{B}_{\boldsymbol{k}}\right\}$. Учитывая, что $\boldsymbol{B}_{\boldsymbol{k}}$ - произвольный элемент $\sigma-$алгебры $\mathcal{B}(\boldsymbol{R_k})$, получаем, что равенство $\boldsymbol{P}\left(\xi \in \left(\boldsymbol{R}_1 \times \boldsymbol{R}_2 \times \ldots \times \boldsymbol{B}_{\boldsymbol{k}} \times \ldots \times \boldsymbol{R}_{\boldsymbol{n}}\right)\right) = \boldsymbol{P}\left(\xi_{\boldsymbol{k}} \in \boldsymbol{B}_{\boldsymbol{k}}\right)$ определяет вероятностную функцию $\boldsymbol{P}_{\boldsymbol{k}}$ компоненты $\xi_{\boldsymbol{k}}$, которую будем называть $\boldsymbol{k}$-той *частной вероятностной функцией*. Перебирая значения $\boldsymbol{k} = 1,2,\ldots,\boldsymbol{n}$, получим частные вероятностные функции $\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_{\boldsymbol{n}}$ всех компонент $\left\{\xi_{\boldsymbol{k}}\right\}$ векторной случайной величины $\xi$.

Пусть векторная случайная величина дискретного типа и $\boldsymbol{n}$=2. Область её возможных значений – не более чем счётное множество точек $\left\{\left(\boldsymbol{x}_{\boldsymbol{i}}; \boldsymbol{y}_{\boldsymbol{j}}\right)\right\}$ плоскости $\boldsymbol{R^2}$. Каждое свое возможное значение случайная величина принимает с вероятностью $\boldsymbol{p}_{\boldsymbol{ij}} = \boldsymbol{P}\left(\xi \in \left(\boldsymbol{x}_{\boldsymbol{i}}; \boldsymbol{y}_{\boldsymbol{j}}\right)\right)$. Вероятностная функция $\boldsymbol{P}$ определяется набором положительных чисел $\left\{\boldsymbol{p}_{\boldsymbol{ij}}\right\}$, при этом: $\sum\limits_{\boldsymbol{i}=1}^{\boldsymbol{n},\infty} \sum\limits_{\boldsymbol{j}=1}^{\boldsymbol{m},\infty} \boldsymbol{p}_{\boldsymbol{ij}} = 1$.

Компоненты $\xi_1$ и $\xi_2$ будут случайными величинами дискретного типа. Определим частные вероятностные функции $\boldsymbol{P}_1$ и $\boldsymbol{P}_2$ этих компонент.

Для каждого значения $\boldsymbol{x}_{\boldsymbol{i}}$ компоненты $\xi_1$ случайное событие $\left\{\boldsymbol{x}_{\boldsymbol{i}} \times \boldsymbol{R}_2\right\}$ является элементом $\sigma-$алгебры $\mathcal{B}(\boldsymbol{R^2})$. Вычислим вероятность этого события:

$$\boldsymbol{P}\left(\xi \in \left\{\boldsymbol{x}_{\boldsymbol{i}} \times \boldsymbol{R}_2\right\}\right) = \boldsymbol{P}\left(\xi \in \left\{\boldsymbol{x}_{\boldsymbol{i}}; \boldsymbol{y}_1\right\} \vee \left\{\boldsymbol{x}_{\boldsymbol{i}}; \boldsymbol{y}_2\right\} \vee \ldots \vee \left\{\boldsymbol{x}_{\boldsymbol{i}}; \boldsymbol{y}_{\boldsymbol{j}}\right\} \vee \ldots \vee \left\{\boldsymbol{x}_{\boldsymbol{i}}; \boldsymbol{y}_{\boldsymbol{m}}\right\} \vee \ldots\right).$$

Одноточечные события $\left\{\boldsymbol{x}_{\boldsymbol{i}}; \boldsymbol{y}_{\boldsymbol{j}}\right\}$, $\boldsymbol{j} = 1,2,\ldots,\boldsymbol{m},\ldots$ – несовместные события, а так как вероятностная функция - аддитивная функция, то:

$$\boldsymbol{P}\left(\xi \in \left\{\boldsymbol{x}_{\boldsymbol{i}} \times \boldsymbol{R}_2\right\}\right) = \sum_{\boldsymbol{j}=1}^{\boldsymbol{m},\infty} \boldsymbol{P}\left(\xi \in \left\{\boldsymbol{x}_{\boldsymbol{i}}; \boldsymbol{y}_{\boldsymbol{j}}\right\}\right) = \sum_{\boldsymbol{j}=1}^{\boldsymbol{m},\infty} \boldsymbol{P}\left(\xi_1 = \left\{\boldsymbol{x}_{\boldsymbol{i}}\right\}; \xi_2 = \left\{\boldsymbol{y}_{\boldsymbol{j}}\right\}\right) = \sum_{\boldsymbol{j}=1}^{\boldsymbol{m},\infty} \boldsymbol{p}_{\boldsymbol{ij}}.$$

Вторая компонента $\xi_2$ обязательно примет одно из своих возможных значений. Следовательно, вычисляемая вероятность события $\left\{\boldsymbol{x}_{\boldsymbol{i}} \times \boldsymbol{R}_2\right\}$ равна вероятности того, что первая компонента примет значение $\boldsymbol{x}_{\boldsymbol{i}}$, то есть:

$$\boldsymbol{P}\left(\xi \in \left\{\boldsymbol{x}_{\boldsymbol{i}} \times \boldsymbol{R}_2\right\}\right) = \boldsymbol{P}\left(\xi_1 = \boldsymbol{x}_{\boldsymbol{i}}\right).$$

Обозначим $\sum\limits_{\boldsymbol{j}=1}^{\boldsymbol{m},\infty} \boldsymbol{p}_{\boldsymbol{ij}} = \boldsymbol{p}_{\boldsymbol{i}*}$. Значит $\boldsymbol{P}\left(\xi_1 = \boldsymbol{x}_{\boldsymbol{i}}\right) = \boldsymbol{p}_{\boldsymbol{i}*}$. Вычислив для каждого $\boldsymbol{i}$=1,2,…,$\boldsymbol{n}$,… значения $\boldsymbol{p}_{\boldsymbol{i}*}$, получим набор вероятностей $\left\{\boldsymbol{p}_{\boldsymbol{i}*}\right\}$. Легко проверить, что $\sum\limits_{\boldsymbol{i}=1}^{\boldsymbol{n},\infty} \boldsymbol{p}_{\boldsymbol{i}*} = 1$. Набор вероятностей $\left\{\boldsymbol{p}_{\boldsymbol{i}*}\right\}$ определяет по вероятностной функции $\boldsymbol{P}$ первую частную вероятностную функцию $\boldsymbol{P}_1$.

<u>Определение.</u> Набор вероятностей $\left\{\boldsymbol{p}_{\boldsymbol{i}*}\right\}$, $\boldsymbol{i}$=1,2,…,$\boldsymbol{n}$,… называется *частным распределением вероятностей первой компоненты* $\xi_1$ дискретной двумерной случайной величины $\xi = \left(\xi_1; \xi_2\right)$.

Аналогично, вычисляя для фиксированного $\boldsymbol{j}$ вероятность события $\left\{\boldsymbol{R}_1 \times \boldsymbol{y}_{\boldsymbol{j}}\right\}$, придём к равенству $P\left(\xi\in\left\{R_1\times y_j\right\}\right)=P\left(\xi_2=y_j\right)=\sum\limits_{i=1}^{n,\infty}p_{ij}$. Обозначим $\sum\limits_{i=1}^{n,\infty}p_{ij}=p_{*j}$. Вычислив для каждого $j$=1,2,…,$m$,… значения $p_{*j}$, получим набор вероятностей $\{p_{*j}\}$, при этом всегда $\sum\limits_{j=1}^{m,\infty}p_{*j}=1$. Набор вероятностей $\{p_{*j}\}$, $j$=1,2,…,$m$,…, определяет частную вероятностную функцию $P_2$.

Определение. Набор вероятностей $\{p_{*j}\}$, $j$=1,2,…,$m$,… называется *частным распределением вероятностей второй компоненты* $\xi_2$ дискретной двумерной случайной величины $\xi=(\xi_1;\xi_2)$.

### *Комментарий*

Закон распределения одномерной случайной величины дискретного типа задаётся *рядом распределения*, в котором перечисляются возможные значения случайной величины и соответствующие вероятности. Так как значениями двумерной случайной величины $\xi$ являются точки пространства $R^2$, то закон распределения $\xi$, в записи которого должны быть перечислены возможные значения $\xi$ и соответствующие вероятности, задаётся *таблицей распределения*.

| $\xi_2(\omega)$ / $\xi_1(\omega)$ | $y_1$ | $y_2$ | … | $y_j$ | … | $y_m$ | $\{p_{i*}\}$ |
|---|---|---|---|---|---|---|---|
| $x_1$ | $(x_1,y_1)$ $p_{11}$ | $(x_1,y_2)$ $p_{12}$ | … | $(x_1,y_j)$ $p_{1j}$ | … | $(x_1,y_m)$ $p_{1m}$ | $p_{1*}$ |
| $x_2$ | $(x_2,y_1)$ $p_{21}$ | $(x_2,y_2)$ $p_{22}$ | | $(x_2,y_j)$ $p_{2j}$ | | $(x_2,y_m)$ $p_{2m}$ | $p_{2*}$ |
| … | | | | … | | | … |
| $x_i$ | $(x_i,y_1)$ $p_{i1}$ | $(x_i,y_2)$ $p_{i2}$ | … | $(x_i,y_j)$ $p_{ij}$ | … | $(x_i,y_m)$ $p_{im}$ | $p_{i*}$ |
| … | | | | … | | | … |
| $x_n$ | $(x_n,y_1)$ $p_{n1}$ | $(x_n,y_2)$ $p_{n2}$ | | $(x_n,y_j)$ $p_{nj}$ | | $(x_n,y_m)$ $p_{nm}$ | $p_{n*}$ |
| $\{p_{*j}\}$ | $p_{*1}$ | $p_{*2}$ | … | $p_{*j}$ | …. | $p_{*m}$ | |

Перечисление вероятностей $\{p_{ij}\}$ возможных значений двумерной случайной величины $\xi$ в виде таблицы позволяет легко определить частные распределения вероятностей $\{p_{i*}\}$ и $\{p_{*j}\}$ компонент $\xi_1$ и $\xi_2$. Каждая вероятность $p_{i*}$ равна сумме вероятностей $p_{ij}$ записанных в $i$-той строке таблицы. Каждая вероятность $p_{*j}$ равна сумме вероятностей $p_{ij}$ записанных в $j$-том столбце таблицы.

## Практика

Пример 1.4.III. Две монеты подбрасываются наудачу один раз.

Пусть $\xi_1$ - число «гербов» на первой монете, $\xi_2$ - число «гербов» на второй монете. Построить таблицу распределения двумерной случайной величины $\xi = (\xi_1; \xi_2)$.

Ясно, что случайная величина $\xi = (\xi_1; \xi_2)$ может принять четыре значения с равными вероятностями. Составим таблицу распределения.

| $\xi_2(\omega)$ / $\xi_1(\omega)$ | $\xi_2(\omega)=0$ | $\xi_2(\omega)=1$ | $\{p_{i*}\}$ |
|---|---|---|---|
| $\xi_1(\omega)=0$ | (0;0) $p_{00} = \frac{1}{4}$ | (0;1) $p_{01} = \frac{1}{4}$ | $p_{0*} = \frac{1}{2}$ |
| $\xi_1(\omega)=1$ | (1;0) $p_{10} = \frac{1}{4}$ | (1;1) $p_{11} = \frac{1}{4}$ | $p_{1*} = \frac{1}{2}$ |
| $\{p_{*j}\}$ | $p_{*0} = \frac{1}{2}$ | $p_{*1} = \frac{1}{2}$ | |

В последнем столбце и в последней строке таблицы записаны частные распределения компонент $\xi_1$ и $\xi_2$.

Пример 2.4.III. Две монеты подбрасываются наудачу один раз.

Оставим прежним название первой компоненты случайной величины $\xi = (\xi_1; \xi_2)$ и изменим формулировку названия второй компоненты случайной величины. Пусть $\xi_1$ - число «гербов» на первой монете, $\xi_2$ - общее количество «гербов» на двух монетах.

Теперь множество возможных значений второй компоненты $\xi_2$ состоит из трёх элементов. А таблица распределения $\xi = (\xi_1; \xi_2)$ будет иметь такой вид.

| $\xi_2(\omega)$ / $\xi_1(\omega)$ | $\xi_2(\omega)=0$ | $\xi_2(\omega)=1$ | $\xi_2(\omega)=2$ | $\{p_{i*}\}$ |
|---|---|---|---|---|
| $\xi_1(\omega)=0$ | (0;0) $p_{00} = \frac{1}{4}$ | (0;1) $p_{01} = \frac{1}{4}$ | | $p_{0*} = \frac{1}{2}$ |
| $\xi_1(\omega)=1$ | | (1;1) $p_{11} = \frac{1}{4}$ | (1;2) $p_{12} = \frac{1}{4}$ | $p_{1*} = \frac{1}{2}$ |
| $\{p_{*j}\}$ | $p_{*0} = \frac{1}{4}$ | $p_{*1} = \frac{2}{4}$ | $p_{*2} = \frac{1}{4}$ | |

В первом примере обе компоненты $\xi_1$ и $\xi_2$ подчиняются бернуллиевскому распределению $B_1\left(\frac{1}{2}\right)$. Во втором примере распределение первой компоненты

$\xi_1$ - не изменилась, а компонента $\xi_2$ подчиняется биномиальному распределению $\boldsymbol{B}_2\left(\frac{1}{2}\right)$. Обсуждение полученных результатов проведём ниже.

***Комментарий***

Пусть двумерная случайная величина $\xi=(\xi_1;\xi_2)$ - непрерывного типа. Её вероятностная функция $\boldsymbol{P}$ определяется плотностью вероятности $\boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})$, при этом: $\iint\limits_{\boldsymbol{R}^2}\boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})\boldsymbol{dxdy}=1$. Компоненты $\xi_1$ и $\xi_2$ будут случайными величинами непрерывного типа.

Функции $\boldsymbol{p}_1(\boldsymbol{x})=\int\limits_{\boldsymbol{R}_2}\boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})\boldsymbol{dy}$ и $\boldsymbol{p}_2(\boldsymbol{y})=\int\limits_{\boldsymbol{R}_1}\boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})\boldsymbol{dx}$ удовлетворяют всем требованиям, предъявляемым к плотности вероятности, определяющей распределение вероятностей одномерной случайной величины непрерывного типа. Эти функции определяют вид частных вероятностных функций $\boldsymbol{P_1}$ и $\boldsymbol{P_2}$ компонент $\xi_1$ и $\xi_2$ двумерной случайной величины непрерывного типа.

Определение. Функции $\boldsymbol{p}_1(\boldsymbol{x})$ и $\boldsymbol{p}_2(\boldsymbol{y})$ называются *частными плотностями вероятностей* компонент $\xi_1$ и $\xi_2$.

**Практика**

Пример 3.4. III. (Задача о встрече)

Два лица - ***A*** и ***B*** договорились о встрече в некотором месте в любой момент времени от 12 часов до 13 часов. Каждый, пришедший на место встречи, ждёт другого в течение 15 мин., после чего уходит. Определить вероятность того, что встреча состоится.

Определим две случайные величины $\xi_1$- время прибытия лица ***A*** на место встречи и $\xi_2$- время прибытия лица ***B*** на место встречи. Рассмотрим двумерную случайную величину $\xi=(\xi_1,\xi_2)$ с компонентами $\xi_1$ и $\xi_2$. Ясно, что если ***A*** и ***B*** – лица обязательные, то есть каждый из них непременно с 12 часов до 13 часов придёт на место встречи, то областью возможных значений $\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{y})$ двумерной случайной величины $\xi=(\xi_1,\xi_2)$ будет квадрат:

$$\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{y})=\left\{(\boldsymbol{x},\boldsymbol{y})\in\boldsymbol{R}^2:0\le\boldsymbol{x}<1;0\le\boldsymbol{y}<1\right\}.$$

Так как все возможные моменты появления и лица ***A***, и лица ***B*** на месте встречи в течение часа (точка $(\boldsymbol{x},\boldsymbol{y})\in\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{y})$) равновозможные, то плотность вероятности случайной величины $\xi$ имеет вид:

$$\boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})=\begin{cases}0, & (\boldsymbol{x},\boldsymbol{y})\notin\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{y})\\ \dfrac{1}{\boldsymbol{mesQ}(\boldsymbol{x},\boldsymbol{y})}=1, & (\boldsymbol{x},\boldsymbol{y})\in\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{y})\end{cases}.$$

Частные плотности вероятности компонент $\xi_1$ и $\xi_2$ будут равны:

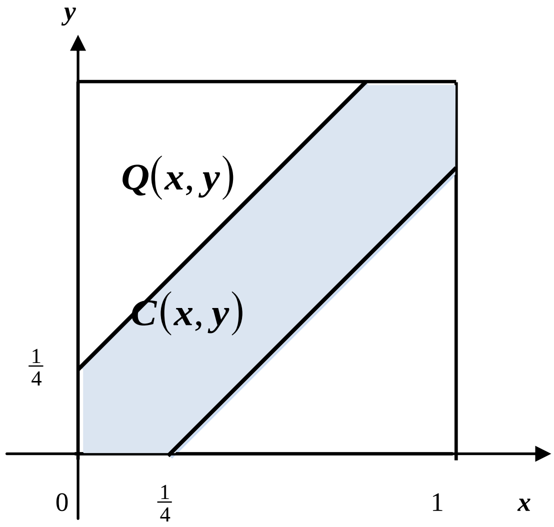
$$\boldsymbol{p}_1(\boldsymbol{x})=\int\limits_{\boldsymbol{R}_2}\boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})\boldsymbol{dy}=\int\limits_0^1\frac{1}{\boldsymbol{mesQ}(\boldsymbol{x},\boldsymbol{y})}\boldsymbol{dy}=1\text{ ,если }0\le\boldsymbol{x}<1,$$

$$p_2(y)=\int_{R_1} p(x,y)dx=\int_0^1 \frac{1}{mesQ(x,y)}dx=1 \text{ ,если } 0\le y<1.$$

Обозначим $C$ – случайное событие – «встреча лиц $A$ и $B$ состоялась».

Ясно, что встреча состоится, если модуль разности между значениями случайных величин $\xi_1$ и $\xi_2$ не будет превышать $\frac{1}{4}$ часа, то есть:

$$C=\left\{(x,y)\in Q(x,y):|x-y|\le\frac{1}{4}\right\}.$$

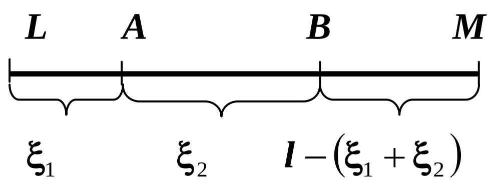



Вероятность наступления события $C$ будет равна: $P(\xi\in C)=\iint_C p(x,y)dxdy$. Так как мы интерпретируем значение вероятности наступления события $C$ как долю единичной массы, оказавшейся над подмножеством $C(x,y)$ множества $Q(x,y)$, то численно вероятность $P(\xi\in C)$ наступления события $C$ будет равна объёму шестигранной призмы, длина высоты которой равна $\frac{1}{mesQ(x,y)}=1$, а, величина площади основания равна площади шестиугольника $C(x,y)$.

Значит: $P(\xi\in C)=mesC(x,y)\cdot\frac{1}{mesQ(x,y)}$, то есть $P(\xi\in C)=\frac{7}{16}$.

Пример 4.4.III. (Задача о треугольнике)

На отрезке $LM$, длина которого равна $l$, наудачу ставятся две точки, в результате чего этот отрезок оказывается разделённым на три отрезка. Определить вероятность того, что из этих отрезков можно построить треугольник.




Ясно, что длины получающихся трёх отрезков – случайные величины.

Обозначим: $\xi_1$ - длина отрезка $\boldsymbol{LA}$, $\xi_2$ - длина отрезка $\boldsymbol{AB}$. Длина отрезка $\boldsymbol{BM}$ зависит от значений случайных величин $\xi_1$ и $\xi_2$, и она равна: $\boldsymbol{l}-(\xi_1+\xi_2)$.

Чтобы рассмотреть все возможные ситуации, которые могут возникнуть при делении отрезка на три части, достаточно рассмотреть двумерную случайную величину $\xi=(\xi_1,\xi_2)$.

Областью возможных значений $\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{y})$ случайной величины $\xi=(\xi_1,\xi_2)$ будет треугольник: $\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{y})=\{(\boldsymbol{x},\boldsymbol{y})\in \boldsymbol{R}^2 : 0\le \boldsymbol{x}; 0\le \boldsymbol{y}; \boldsymbol{x}+\boldsymbol{y}\le \boldsymbol{l}\}$.

Так как все положения точек $\boldsymbol{A}$ и $\boldsymbol{B}$ на отрезке $\boldsymbol{LM}$ – равновозможные, то двумерная случайная величина $\xi=(\xi_1,\xi_2)$ подчиняется на множестве $\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{y})$ равномерному распределению:

$$\boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})=\begin{cases} 0, & (\boldsymbol{x},\boldsymbol{y})\notin \boldsymbol{Q}(\boldsymbol{x},\boldsymbol{y}) \\ \dfrac{1}{\boldsymbol{mesQ}(\boldsymbol{x},\boldsymbol{y})}=\dfrac{2}{\boldsymbol{l}^2}, & (\boldsymbol{x},\boldsymbol{y})\in \boldsymbol{Q}(\boldsymbol{x},\boldsymbol{y}) \end{cases}$$

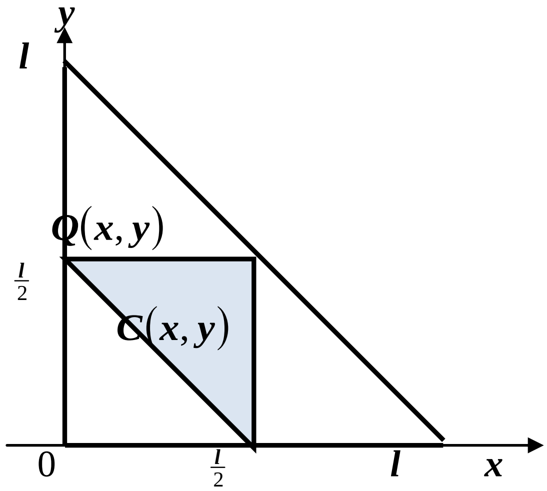
Частные плотности вероятности компонент $\xi_1$ и $\xi_2$ будут равны:

$$\boldsymbol{p}_1(\boldsymbol{x})=\int\limits_{\boldsymbol{R}_2}\boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})\boldsymbol{dy}=\int\limits_0^{\boldsymbol{l}-\boldsymbol{x}}\frac{1}{\boldsymbol{mesQ}(\boldsymbol{x},\boldsymbol{y})}\boldsymbol{dy}=\frac{2}{\boldsymbol{l}^2}\cdot(\boldsymbol{l}-\boldsymbol{x}) \quad \text{,если } 0\le \boldsymbol{x}<\boldsymbol{l},$$

$$\boldsymbol{p}_2(\boldsymbol{y})=\int\limits_{\boldsymbol{R}_1}\boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})\boldsymbol{dx}=\int\limits_0^{\boldsymbol{l}-\boldsymbol{y}}\frac{1}{\boldsymbol{mesQ}(\boldsymbol{x},\boldsymbol{y})}\boldsymbol{dx}=\frac{2}{\boldsymbol{l}^2}\cdot(\boldsymbol{l}-\boldsymbol{y}) \quad \text{,если } 0\le \boldsymbol{y}<\boldsymbol{l}.$$

Обозначим $\boldsymbol{C}$ – случайное событие – «из трёх отрезков треугольник получился»

Ясно, что из трёх наудачу взятых отрезков треугольник получится, если будет выполнено условие: сумма длин любых двух отрезков больше длины третьего отрезка.




То есть $\boldsymbol{C}(\boldsymbol{x},\boldsymbol{y})=\begin{cases} \boldsymbol{x}+\boldsymbol{y}>\boldsymbol{l}-(\boldsymbol{x}+\boldsymbol{y}), \\ \boldsymbol{x}+(\boldsymbol{l}-(\boldsymbol{x}+\boldsymbol{y}))>\boldsymbol{y}, \\ \boldsymbol{y}+(\boldsymbol{l}-(\boldsymbol{x}+\boldsymbol{y}))>\boldsymbol{x}, \end{cases}$ или $\boldsymbol{C}(\boldsymbol{x},\boldsymbol{y})=\begin{cases} \boldsymbol{x}+\boldsymbol{y}>\dfrac{\boldsymbol{l}}{2} \\ \boldsymbol{y}<\dfrac{\boldsymbol{l}}{2} \\ \boldsymbol{x}<\dfrac{\boldsymbol{l}}{2} \end{cases}$.

Подмножество $C(x,y)$ является основанием призмы, длина высоты которой равна $\dfrac{1}{mesQ(x,y)}$. Площадь этого подмножества, как видно из рисунка, составляет $\dfrac{1}{4}$ площади множества возможных значений $Q(x,y)$ случайной величины $\xi=(\xi_1,\xi_2)$. Интерпретация значения вероятности наступления события $C$ как доли единичной массы, оказавшейся над подмножеством $C(x,y)$, приводит нас к выводу, что $P(C)=\dfrac{1}{4}$.

Непосредственные вычисления значения вероятности наступления события $C$, подтверждают этот вывод: $P(\xi\in C)=\iint\limits_{C(x,y)} p(x,y)dxdy=\dfrac{2}{l^2}\int\limits_0^{\frac{l}{2}}\left(\int\limits_{\frac{l}{2}-x}^{\frac{l}{2}}dy\right)dx=\dfrac{1}{4}$.

Пример 5.4.III. Двумерная случайная величина $\xi=(\xi_1;\xi_2)$ имеет равномерное распределение вероятностей в области $Q(x,y)\subset R^2$, ограниченной эллипсом $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$. То есть плотность вероятности случайной величины $\xi$ имеет вид:

$$p(x,y)=\begin{cases} 0, & (x,y)\notin Q(x,y) \\ \dfrac{1}{\pi ab}, & (x,y)\in Q(x,y) \end{cases}.$$

Определим частные плотности вероятности $p_1(x)$ и $p_2(y)$ компонент случайной величины $\xi$.

Случайные величины $\xi_1$ и $\xi_2$ принимают значения, принадлежащие сегментам $[-a;+a]\subset R_1$ и $[-b;+b]\subset R_2$, соответственно. Здесь: $[-a;+a]=proj_{R_1}Q(x,y)$ и $[-b;+b]=proj_{R_2}Q(x,y)$. Но каждое фиксированное значение $\xi_1$ определяет соответствующее ему множество возможных значений $\xi_2$: То есть, если $\xi_1=x_0$, то множеством возможных значений случайной величины $\xi_2$ будет сегмент $\left[-b\cdot\sqrt{1-\dfrac{x_0^2}{a^2}};+b\cdot\sqrt{1-\dfrac{x_0^2}{a^2}}\right]$. Кратко запишем этот сегмент так: $[-\alpha(x_0);+\alpha(x_0)]$. Для значений $y\in[-\alpha(x_0);+\alpha(x_0)]$ плотность вероятности случайной величины $\xi$ будет равна: $p(x_0,y)=\dfrac{1}{\pi ab}$. Тогда для каждого фиксированного значения $x_0$ первой компоненты $\xi_1$ получаем значение её частной плотности вероятности:

$$p_1(x_0)=\int\limits_{-\alpha(x_0)}^{\alpha(x_0)}p(x_0,y)dy=2\cdot\frac{1}{\pi ab}\cdot\alpha(x_0)=2\cdot\frac{1}{\pi ab}\cdot b\sqrt{1-\frac{x_0^2}{a^2}}.$$

Следовательно, частная плотность вероятности первой компоненты $\xi_1$, будет иметь вид: $$p_1(x)=\begin{cases}0, & x\notin[-a;+a]\\ \dfrac{2}{\pi a}\cdot\sqrt{1-\dfrac{x^2}{a^2}}, & x\in[-a;+a]\end{cases}.$$

Аналогично, каждое фиксированное значение $\xi_2$ определяет множество возможных значений $\xi_1$. То есть, если $\xi_2=y_0$, то множеством возможных значений случайной величины $\xi_1$ будет сегмент $\left[-a\cdot\sqrt{1-\dfrac{y_0^2}{b^2}};+a\cdot\sqrt{1-\dfrac{y_0^2}{b^2}}\right]$.

Обозначим $\beta(y_0)=a\cdot\sqrt{1-\dfrac{y_0^2}{b^2}}$.

Вычисляя для каждого фиксированного $y_0$ значение частной плотности вероятности: $p_2(y_0)=\int\limits_{-\beta(y_0)}^{+\beta(y_0)}p(x,y_0)dx$, приходим к выводу, что частная плотность вероятности второй компоненты $\xi_2$ имеет вид:

$$p_2(y)=\begin{cases}0, & y\notin[-b;+b]\\ \dfrac{2}{\pi b}\cdot\sqrt{1-\dfrac{y^2}{b^2}}, & y\in[-b;+b]\end{cases}.$$

Пример 6.4.III. Двумерная случайная величина $\xi=(\xi_1,\xi_2)$ имеет нормальное распределение вероятностей, то есть плотность вероятности случайной величины $\xi$ имеет вид:

$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-m_1)^2}{\sigma_1^2}-2\rho\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2}+\frac{(y-m_2)^2}{\sigma_2^2}\right]}.$$

Найдём частную плотность вероятности первой компоненты: $p_1(x)=\int\limits_{R_2}p(x,y)dy$.

Для вычисления интеграла с помощью замены переменной интегрирования преобразуем квадратичную форму показательной функции:

$$\left[\frac{(x-m_1)^2}{\sigma_1^2}-2\rho\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2}+\frac{(y-m_2)^2}{\sigma_2^2}\right]=\frac{(x-m_1)^2}{\sigma_1^2}(1-\rho^2)+\left[\frac{y-m_2}{\sigma_2}-\frac{x-m_1}{\sigma_1}\rho\right]^2.$$

Значит: $p_1(x)=\dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{(x-m_1)^2}{2\sigma_1^2}}\cdot\int\limits_{-\infty}^{+\infty}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{y-m_2}{\sigma_2}-\frac{x-m_1}{\sigma_1}\rho\right]^2}dy$.

Заменим переменную интегрирования по формуле:

$\dfrac{1}{\sqrt{1-\rho^2}}\left[\dfrac{y-m_2}{\sigma_2}-\dfrac{x-m_1}{\sigma_1}\rho\right]=z$, тогда $dy=\sqrt{1-\rho^2}\sigma_2dz$. Проведя сокращения, получаем $p_1(x)=\frac{1}{2\pi\sigma_1}e^{-\frac{(x-m_1)^2}{2\sigma_1^2}}\cdot\int\limits_{-\infty}^{+\infty}e^{-\frac{1}{2}z^2}dz$. Из курса математического анализа мы знаем, что $\int\limits_0^{\infty}e^{-\frac{1}{2}z^2}dz=\sqrt{\frac{\pi}{2}}$. Следовательно, частная плотность вероятности будет иметь вид: $p_1(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-m_1)^2}{2\sigma_1^2}}$. То есть, первая компонента $\xi_1$ имеет нормальное распределение вероятностей, которое мы обозначаем $\mathcal{N}(m_1,\sigma_1)$. Аналогично получаем, что частная плотность вероятности второй компоненты $\xi_2$ будет иметь вид: $p_2(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-m_2)^2}{2\sigma_2^2}}$. То есть, компоненты $\xi_1$ и $\xi_2$ имеют нормальное распределение вероятностей $\mathcal{N}(m_1,\sigma_1)$ и $\mathcal{N}(m_2,\sigma_2)$ со своими параметрами, смысл которых будет установлен позже.

### *Комментарий*

Итак, зная вероятностную функцию $P$ случайного вектора $\xi=(\xi_1,\xi_2,...,\xi_n)$, мы всегда сможем найти частные вероятностные функции $P_1,P_2,...,P_n$ его компонент. Обратное оказывается не всегда верно, то есть: по частным вероятностным функциям $P_1,P_2,...,P_n$ компонент не всегда можно восстановить вероятностную функцию $P$ вектора. Что можно увидеть в рассмотренных выше примерах.

В примере 3.4.III. (Задача о встрече) случайная величина $\xi=(\xi_1,\xi_2)$ - непрерывного типа, поэтому вероятностные функции случайной величины $\xi$ и её компонент $\xi_1$ и $\xi_2$ задаются соответствующими плотностями вероятности:

$$p(x,y)=\begin{cases}0, & (x,y)\notin[0;1)\times[0;1)\\ 1, & (x,y)\in[0;1)\times[0;1)\end{cases},\quad p_1(x)=\begin{cases}0, & x\notin[0;1)\\ 1, & x\in[0;1)\end{cases},\quad p_2(y)=\begin{cases}0, & y\notin[0;1)\\ 1, & y\in[0;1)\end{cases}.$$

Здесь очевидно, что если $x\in[0;1)$ и $y\in[0;1)$, то будет справедливо: $p_1(x)\cdot p_2(y)=p(x,y)$.

В примере 4.4.III. (Задача о треугольнике) случайная величина $\xi=(\xi_1,\xi_2)$- непрерывного типа и плотности вероятности $\xi$ и её компонент $\xi_1$ и $\xi_2$ здесь имеют вид:

$$p(x,y)=\begin{cases}0, & (x,y)\notin Q(x,y)\\ \frac{2}{l^2}, & (x,y)\in Q(x,y)\end{cases},$$

$$p_1(x)=\begin{cases}0, & x\notin[0;l)\\ \frac{2}{l^2}\cdot(l-x), & x\in[0;l)\end{cases},\quad p_2(y)=\begin{cases}0, & y\notin[0;l)\\ \frac{2}{l^2}\cdot(l-y), & y\in[0;l)\end{cases}.$$

Пытаясь по частным плотностям вероятности компонент $\xi_1$ и $\xi_2$ восстановить плотность вероятности двумерной $\xi$, мы, прежде всего, должны определить область возможных значений $\xi$. Ясно, что, если $x \in [0; l)$ и $y \in [0; l)$, то $(x, y) \in [0; l) \times [0; l)$. Получающееся множество возможных значений восстанавливаемой двумерной случайной величины будет квадратом, который будет содержать в себе треугольник $Q(x, y)$. Если мы определим плотность вероятности этой двумерной случайной величины как произведение частных плотностей вероятности $p^*(x, y) = p_1(x) \cdot p_2(y)$, то приходим к выводу, что мы получаем плотность вероятности новой случайной величины $\xi^*$. Распределение вероятностей $\xi^*$ не совпадает с распределением вероятностей исходной случайной величины $\xi = (\xi_1, \xi_2)$.

Рассмотрим общий случай. Пусть $n$=2, то есть мы имеем вероятностное пространство $<R^2, \mathcal{B}(R^2), P>$ и двумерную случайную величину $\xi = (\xi_1, \xi_2)$. Пусть $\Omega \subseteq R^2$ - множество возможных значений $\xi = (\xi_1, \xi_2)$. Подмножества $C$ множества $\Omega$ - случайные события, $C \in \mathcal{B}(R^2)$. Определяя для любого случайного события $C$ его проекции на координатные оси – случайные события $A = proj_{R_1} C$ и $B = proj_{R_2} C$, мы можем, используя элементы $\sigma$-алгебры $\mathcal{B}(R^2)$, построить измеримые пространства $<R_1, \mathcal{B}(R_1)>$ и $<R_2, \mathcal{B}(R_2)>$, в которых $\mathcal{B}(R_1)$ и $\mathcal{B}(R_2)$ – алгебры случайных событий, получающихся путём применения не более чем счётного числа операций объединения пересечения, дополнения к проекциям случайных событий $C \in \mathcal{B}(R^2)$: событиям $A = proj_{R_1} C$ и $B = proj_{R_2} C$.

Полагая $P_1(\xi_1 \in A) = P(\xi \in A \times R_2)$ и $P_2(\xi_2 \in B) = P(\xi \in R_1 \times B)$, мы определяем частные вероятностные функции $P_1$ и $P_2$ компонент случайной величины $\xi$ – одномерных случайных величин $\xi_1$ и $\xi_2$.

Желая по частным вероятностным распределениям $P_1$ и $P_2$ компонент $\xi_1$ и $\xi_2$ восстановить распределение вероятностей двумерной случайной величины $\xi$, сначала построим прямое произведение построенных алгебр $\mathcal{B}(R_1) \otimes \mathcal{B}(R_2)$, элементами которого будут случайные события $C^*$, являющиеся декартовыми произведениями проекций $C$ на координатные оси:

$C^* = A \times B = proj_{R_1} C \times proj_{R_2} C$. Ясно, что $C \subseteq C^*$.

Обозначим $\Lambda$ - множество, любое подмножество которого является случайным событием $C^*$. Ясно, что $\Omega \subseteq \Lambda \subseteq R^2$ и для алгебры $\mathcal{B}(\Lambda)$ подмножеств множества $\Lambda$ будет справедливо: $\mathcal{B}(\Lambda) = \mathcal{B}(R_1) \otimes \mathcal{B}(R_2)$. Определим двумерную случайную величину $\xi^* = (\xi_1, \xi_2)$, множеством возможных значений которой будет множество $\Lambda$, а её вероятностной функцией будет функция $P^* = P_1 \cdot P_2$, определённая на $\mathcal{B}(\Lambda)$. То есть, вероятность наступления события $C^*$ определим как произведение значений частных вероятностных функций: $P^*(\xi^* \in C^*) = P_1(\xi_1 \in A) \cdot P_2(\xi_2 \in B)$.

Вероятностная функция $P^* = P_1 \cdot P_2$ удовлетворяет всем требованиям аксиоматического построения вероятностного пространства $<\boldsymbol{R^2},\mathcal{B}(\boldsymbol{R^2}),\boldsymbol{P}^*>$.

Построенную таким образом случайную величину $\xi^*=(\xi_1,\xi_2)$ будем называть *составной* случайной величиной.

В примере 5.4.III. множеством $\Lambda$ возможных значений составной случайной величины $\xi^*$ будет прямоугольник $\Lambda=[-\boldsymbol{a};+\boldsymbol{a}]\times[-\boldsymbol{b};\boldsymbol{b}]$. Распределение вероятностей $\boldsymbol{P}^*$ случайной величины $\xi^*$ определяется плотностью вероятности $\boldsymbol{p}^*(\boldsymbol{x},\boldsymbol{y})$, являющейся произведением частных плотностей вероятности:

$$\boldsymbol{p}_1(\boldsymbol{x})=\begin{cases}0, & \boldsymbol{x}\notin[-\boldsymbol{a};+\boldsymbol{a}]\\ \dfrac{2}{\pi\boldsymbol{a}}\cdot\sqrt{1-\dfrac{\boldsymbol{x}^2}{\boldsymbol{a}^2}}, & \boldsymbol{x}\in[-\boldsymbol{a};+\boldsymbol{a}]\end{cases} \text{ и } \boldsymbol{p}_2(\boldsymbol{y})=\begin{cases}0, & \boldsymbol{y}\notin[-\boldsymbol{b};+\boldsymbol{b}]\\ \dfrac{2}{\pi\boldsymbol{b}}\cdot\sqrt{1-\dfrac{\boldsymbol{y}^2}{\boldsymbol{b}^2}}, & \boldsymbol{y}\in[-\boldsymbol{b};+\boldsymbol{b}]\end{cases}.$$

То есть $\boldsymbol{p}^*(\boldsymbol{x},\boldsymbol{y})=\begin{cases}0, & \text{если } (\boldsymbol{x},\boldsymbol{y})\notin\Lambda\\ \boldsymbol{p}_1(\boldsymbol{x})\cdot\boldsymbol{p}_2(\boldsymbol{y}), & \text{если } (\boldsymbol{x},\boldsymbol{y})\in\Lambda\end{cases}.$

Ясно, что и в этом примере области возможных значений $\Omega$ и $\Lambda$, и распределения вероятностей $\boldsymbol{P}$ и $\boldsymbol{P}^*$ двумерных случайных величин $\xi=(\xi_1,\xi_2)$ и $\xi^*=(\xi_1,\xi_2)$ не будут совпадать, в то время как компоненты $\xi_1$ и $\xi_2$ этих двумерных случайных величин имеют одинаковые распределения вероятностей.

***Комментарий***

Мы приходим к следующему выводу.

В черчении по имеющейся геометрической фигуре мы можем построить три её проекции: вид в анфас, вид в профиль и вид сверху. А затем, по этим трём проекциям, мы можем начертить геометрическую фигуру, которая совпадёт с начальной геометрической фигурой.

В теории вероятностей такое не всегда возможно. По проекциям случайного вектора не всегда можно восстановить начальный случайный вектор.

## §5.III Функция распределения многомерной случайной величины

### Теория

Понятие функции распределения многомерной случайной величины вводится аналогично введению функции распределения одномерной случайной величины.

Пусть $\xi$ - многомерная случайная величина и $<\boldsymbol{R^n},\mathcal{B}(\boldsymbol{R^n}),\boldsymbol{P}>$ - соответствующее ей вероятностное пространство. Обозначим $\boldsymbol{S}_x$ множество $(-\infty;\boldsymbol{x})$, где $\boldsymbol{x}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_n)$ – произвольная точка пространства $\boldsymbol{R^n}$ . Ясно что $\boldsymbol{S}_x$ - борелевское множество. Значит, для любого $\boldsymbol{x}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_n)$ мы можем определить вероятность $\boldsymbol{P}(\xi(\omega)\in\boldsymbol{S}_x)$ случайного события $\{\xi\in\boldsymbol{S}_x\}$. Если $\boldsymbol{x}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_n)$ будет изменяться, то мы получаем функцию $\boldsymbol{n}$ аргументов.

<u>Определение.</u> Функцию $\boldsymbol{n}$ аргументов $\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{P}(\xi(\omega)\in\boldsymbol{S}_x)$ будем называть - *функцией распределения **n**-мерной случайной величины* $\xi$.

Множество $\boldsymbol{S}_x$ - является декартовым произведением $\boldsymbol{n}$ интервалов

$S_{x_i}=(-\infty;x_i)$, $\quad i=1,2,...,n$: $\quad S_x=S_{x_1}\times S_{x_2}\times...\times S_{x_n}$. Значит, событие $\{\xi\in S_x\}$ можно записать так: $\{\xi\in S_x\}=\{\xi_1\in S_{x_1},\xi_2\in S_{x_2},...,\xi_n\}$ и функцию распределения $n$-мерной случайной величины $\xi$ представить в виде:

$$F(x)=F(x_1,x_2,...,x_n)=P(\xi_1<x_1,\xi_2<x_2,...,\xi_n<x_n).$$

Ограничиваясь случаем $n$=2, запишем общий вид функции распределения $F(x,y)$ двумерной случайной величины непрерывного и дискретного типа.

*а)* Если двумерная случайная величина $\xi$ - непрерывного типа и $p(x,y)$ её плотность вероятности, то $F(x,y)=\int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y}p(u,v)du dv$.

*б)* Если двумерная случайная величина $\xi$ - дискретного типа и набор положительных чисел $\{p_{ij}\}$, где $i=1,...,n,...$, и $j=1,...,m,...$, её распределение вероятностей, то $F(x,y)=\sum\limits_{x_i<x}\sum\limits_{y_j<y}p_{ij}$.

Функция распределения $F(x)$ обладает свойствами, аналогичными свойствам функции распределения одномерной случайной величины:

1) функция $F(x)$ - неубывающая функция по каждому из своих аргументов;

2) функция $F(x)$ - непрерывная слева по каждому из своих аргументов;

3*а*) $\lim\limits_{x\downarrow-\infty}F(x)=F(-\infty)=0$;

3*б* ) для любого номера $i$: $\lim\limits_{x_i\downarrow-\infty}F(x_1,...,x_n)=0$

3*в*) $\lim\limits_{x\uparrow+\infty}F(x)=F(+\infty)=1$.

Записывая $x\downarrow-\infty$, мы подразумеваем, что для всех координат аргумента $x$ выполняется $x_i\downarrow-\infty$. Аналогично, запись $x\uparrow+\infty$ означает, что для всех координат выполняется $x_i\uparrow+\infty$.

По вероятностной функции $n$-мерной случайной величины мы всегда можем найти частные вероятностные функции её компонент. Так как любое значение функции распределения $F(x)$ - это значение вероятностной функции $P$ для события $S_x=S_{x_1}\times S_{x_2}\times...\times S_{x_n}$, то мы можем определить частную функцию распределения $F_i(x_i)$ любой $i$-той компоненты $\xi_i$:

$$F_i(x_i)=P(\xi_1<+\infty,..,\xi_{i-1}<+\infty,\xi_i<x_i,\xi_{i+1}<+\infty,..,\xi_n<+\infty)=F(+\infty,..,+\infty,x_i,+\infty,..,+\infty)$$

Определение функций распределения компонент является частным случаем свойства функции распределения $n$-мерной случайной величины, называемого свойством *согласованности*, которое записывается так:

*Если для каких-либо* ***k*** *аргументов функции распределения* ***n****-мерной случайной величины выполняется* $x_j\uparrow+\infty,(j=1,...,k)$, *то получающаяся функция будет функцией распределения* ***n-k*** *– мерной случайной величины.*

***Комментарий***

Рассматривая одномерную случайную величину, мы показали, что любой вероятностной функции $P$ соответствует единственная функция распределения

$\boldsymbol{F}(\boldsymbol{x})$ и, наоборот, по функции распределения $\boldsymbol{F}(\boldsymbol{x})$, то есть по произвольной функции, обладающей свойствами 1,2,3, всегда можно единственным образом построить вероятностную функцию $\boldsymbol{P}$.

Если размерность многомерной случайной величины $\boldsymbol{n} \geq 2$, то мы всегда единственным образом по вероятностной функции $\boldsymbol{P}$ можем определить функцию распределения $\boldsymbol{F}(\boldsymbol{x})$, которая будет функцией $\boldsymbol{n}$ аргументов и будет обладать свойствами 1,2,3.

Однако обратное утверждение: «по произвольной функции $\boldsymbol{F}(\boldsymbol{x})$ $\boldsymbol{n}$ аргументов, обладающей свойствами 1,2,3, можно единственным образом построить вероятностную функцию $\boldsymbol{P}$ будет не всегда верно». То есть, обладания функцией $\boldsymbol{n}$ аргументов $\boldsymbol{F}(\boldsymbol{x})$ тремя свойствами недостаточно, чтобы она была функцией распределения $\boldsymbol{n}$-мерной случайной величины. Следовательно, по такой функции нельзя построить вероятностную функцию $\boldsymbol{P}$. Что показывает следующий пример.

Пример. Пусть функция двух аргументов $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$ определена следующим образом: $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y}) = \begin{cases} 0, & \text{если } \boldsymbol{x}+\boldsymbol{y} \leq 1 \\ 1, & \text{если } \boldsymbol{x}+\boldsymbol{y} > 1 \end{cases}$. Легко проверить, что функция $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$ обладает свойствами 1,2,3. Будем считать, что мы определяем по функции $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$ вероятностную функцию $\boldsymbol{P}$, полагая, что вероятность случайного события $\{\boldsymbol{S}_{x,y} = \boldsymbol{S}_x \times \boldsymbol{S}_y\}$, где $\boldsymbol{S}_{x,y} = (-\infty, \boldsymbol{x}) \times (-\infty, \boldsymbol{y}) \in \mathcal{B}(\boldsymbol{R}^2)$, равна значению функции $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$ в точке $(\boldsymbol{x},\boldsymbol{y})$.

Рассмотрим случайное событие - квадрат $\boldsymbol{B} = \left[\frac{1}{2},1\right] \times \left[\frac{1}{2},1\right]$. Ясно, что $\boldsymbol{B} \in \mathcal{B}(\boldsymbol{R}^2)$. Событие $\boldsymbol{B}$ можно представить так: $\boldsymbol{B} = \boldsymbol{S}_{1,1} \setminus \boldsymbol{S}_{1,\frac{1}{2}} \setminus \boldsymbol{S}_{\frac{1}{2},1} \bigcup \boldsymbol{S}_{\frac{1}{2},\frac{1}{2}}$. Так как вероятностная функция – аддитивная функция, то можем записать:

$$\boldsymbol{P}(\boldsymbol{B}) = \boldsymbol{F}(1,1) - \boldsymbol{F}\left(1,\tfrac{1}{2}\right) - \boldsymbol{F}\left(\tfrac{1}{2},1\right) + \boldsymbol{F}\left(\tfrac{1}{2},\tfrac{1}{2}\right).$$

Записывая значения функции $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$ в четырёх точках, являющихся вершинами квадрата $\boldsymbol{B} = \left[\frac{1}{2},1\right] \times \left[\frac{1}{2},1\right]$, получаем: $\boldsymbol{P}(\boldsymbol{B}) = 1-1-1+0 = -1 < 0$. Противоречие, - вероятность наступления случайного события не может быть отрицательным числом.

Значит не всякая функция $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$, обладающая свойствами 1,2,3, будет функцией распределения многомерной случайной величины.

**Теория**

Чтобы значения вероятностной функции $\boldsymbol{P}$, построенной по функции $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$, на элементах алгебры $\mathcal{B}(\boldsymbol{R}^n)$ всегда были неотрицательными числами надо потребовать, чтобы функция $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$ обладала ещё одним, четвёртым свойством, которое предохраняет функцию $\boldsymbol{P}$, от возникновения противоречий, аналогичных противоречию, возникшему в рассмотренном примере.

Пусть множество $\boldsymbol{B}=[\boldsymbol{a},\boldsymbol{b})=[\boldsymbol{a}_1,\boldsymbol{b}_1)\times[\boldsymbol{a}_2,\boldsymbol{b}_2)$ произвольный прямоугольник из системы полузакрытых прямоугольников $\boldsymbol{A}$, с помощью которых мы строим $\sigma$ – алгебру $\mathcal{B}(\boldsymbol{R}^2)$ борелевских множеств на плоскости $\boldsymbol{R}^2=\boldsymbol{R}_1\times\boldsymbol{R}_2$.

Пусть $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$ - произвольная функция, обладающая свойствами 1,2,3. Определим разностные операторы: $\Delta_{[\boldsymbol{a}_1,\boldsymbol{b}_1)}\boldsymbol{F}=\boldsymbol{F}(\boldsymbol{b}_1,\boldsymbol{y})-\boldsymbol{F}(\boldsymbol{a}_1,\boldsymbol{y})$ и $\Delta_{[\boldsymbol{a}_2,\boldsymbol{b}_2)}\boldsymbol{F}=\boldsymbol{F}(\boldsymbol{x},\boldsymbol{b}_2)-\boldsymbol{F}(\boldsymbol{x},\boldsymbol{a}_2)$. Суперпозиция этих операторов имеет вид:

$$\Delta_{[\boldsymbol{a},\boldsymbol{b})}\boldsymbol{F}=\Delta_{[\boldsymbol{a}_1,\boldsymbol{b}_1)}\left(\Delta_{[\boldsymbol{a}_2,\boldsymbol{b}_2)}\boldsymbol{F}\right)=\boldsymbol{F}(\boldsymbol{b}_1,\boldsymbol{b}_2)-\boldsymbol{F}(\boldsymbol{b}_1,\boldsymbol{a}_2)-\boldsymbol{F}(\boldsymbol{a}_1,\boldsymbol{b}_2)+\boldsymbol{F}(\boldsymbol{a}_1,\boldsymbol{a}_2).$$

Теперь ясно: для того чтобы произвольная функция $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$ была функцией распределения некоторой двумерной случайной величины необходимо, чтобы она обладала ещё и свойством:

4) для любого прямоугольника $\boldsymbol{B}=[\boldsymbol{a},\boldsymbol{b})\subset\boldsymbol{R}^2$ выполняется неравенство $\Delta_{[\boldsymbol{a},\boldsymbol{b})}\boldsymbol{F}\geq 0$.

## §6.III Независимость случайных величин

## Теория

Рассматриваются вероятностное пространство <$\boldsymbol{R}^2$,$\mathcal{B}(\boldsymbol{R}^2)$,$\boldsymbol{P}$> и $\xi=(\xi_1,\xi_2)$ - двумерная случайная величина. Функция $\boldsymbol{P}$ – вероятностная функция, которая определяет распределение вероятностей случайной величины $\xi=(\xi_1,\xi_2)$.

По вероятностной функции $\boldsymbol{P}$ , мы определим частные вероятностные функции $\boldsymbol{P}_1$ и $\boldsymbol{P}_2$ компонент $\xi_1$ и $\xi_2$, и рассмотрим два вероятностных пространства <$\boldsymbol{R}_1$,$\mathcal{B}(\boldsymbol{R}_1)$, $\boldsymbol{P}_1$> и <$\boldsymbol{R}_2$,$\mathcal{B}(\boldsymbol{R}_2)$, $\boldsymbol{P}_2$>.

Пример 1, в котором две монеты наудачу подбрасывались один раз, каждая из компонент $\xi_1$ и $\xi_2$, принимает свои возможные значения, независимо от того, какое значение приняла другая компонента. В примере 2 возможные значения второй случайной величины $\xi_2$ зависят от результата первого подбрасывания, то есть от значения, которая приняла случайная величина $\xi_1$.

Понятие **независимости** определяет ту специфичность, которая выделяет теорию вероятностей в общей теории, занимающейся исследованием измеримых пространств с мерой.

<u>Определение.</u> Случайные величины $\xi_1$ и $\xi_2$ называются ***независимыми случайными величинами***, если для любого борелевского множества $\boldsymbol{B}\in\mathcal{B}(\boldsymbol{R}^2)$, которое *является декартовым произведением* множеств $\boldsymbol{B}_1\in\mathcal{B}(\boldsymbol{R}_1)$ и $\boldsymbol{B}_2\in\mathcal{B}(\boldsymbol{R}_2)$, то есть, если $\boldsymbol{B}=\boldsymbol{B}_1\times\boldsymbol{B}_2$, будет выполняться:

$$\boldsymbol{P}((\xi_1,\xi_2)\in\boldsymbol{B})=\boldsymbol{P}_1(\xi_1\in\boldsymbol{B}_1)\cdot\boldsymbol{P}_2(\xi_2\in\boldsymbol{B}_2). \qquad (1)$$

Случайное событие $\boldsymbol{B}$ - элемент алгебры $\mathcal{B}(\boldsymbol{R}^2)$, а случайные события $\boldsymbol{B}_1\in\mathcal{B}(\boldsymbol{R}_1)$ и $\boldsymbol{B}_2\in\mathcal{B}(\boldsymbol{R}_2)$ - это проекции события $\boldsymbol{B}$ на координатные оси: $\boldsymbol{B}_1=\boldsymbol{proj}_{\boldsymbol{R}_1}\boldsymbol{B}$ и $\boldsymbol{B}_2=\boldsymbol{proj}_{\boldsymbol{R}_2}\boldsymbol{B}$. Определение независимости требует, чтобы равенство (1) выполнялось для всех событий $\boldsymbol{B}$, которые можно представить в виде: $\boldsymbol{B}=\boldsymbol{proj}_{\boldsymbol{R}_1}\boldsymbol{B}\times\boldsymbol{proj}_{\boldsymbol{R}_2}\boldsymbol{B}$.

***Комментарий***

Пусть $\xi_1$ и $\xi_2$ - произвольные случайные величины, значения которых принадлежат $\boldsymbol{R_1}$ и $\boldsymbol{R_2}$. Функции $\boldsymbol{P}_1$ и $\boldsymbol{P}_2$ - вероятностные функции этих случайных величин $\xi_1$ и $\xi_2$. Построим двумерную случайную величину $\xi^* = (\xi_1, \xi_2)$, значения которой принадлежат $\boldsymbol{R}_1 \times \boldsymbol{R}_2 = \boldsymbol{R}^2$, а распределение вероятностей этой случайной величины $\xi^*$ задаётся функцией $\boldsymbol{P}^* = \boldsymbol{P}_1 \cdot \boldsymbol{P}_2$. Такую случайную величину мы называем *составной* случайной величиной. Будем рассматривать случайные события $\boldsymbol{C}^*$, которые представимы в виде: $\boldsymbol{C}^* = \boldsymbol{A} \times \boldsymbol{B}$. Значение вероятности таких событий определяется как произведение значений частных вероятностных функций: $\boldsymbol{P}(\xi^* \in \boldsymbol{C}^*) = \boldsymbol{P}_1(\xi_1 \in \boldsymbol{A}) \cdot \boldsymbol{P}_2(\xi_2 \in \boldsymbol{B})$.

Для компонент $\xi_1$ и $\xi_2$ составной случайной величины $\xi^*$ выполнены все требования определения независимости случайных величин. Следовательно, у составной случайной величины компоненты $\xi_1$ и $\xi_2$ - всегда независимые случайные величины.

## Критерии независимости случайных величин

### Теория

Мы будем постепенно переходить к изучению множества возможных значений и распределения вероятностей на этом множестве многомерной случайной величины $\xi$ путём анализа её функции распределения $\boldsymbol{F}(\boldsymbol{x})$, которую мы определили следующим образом: $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{P}(\xi \in \boldsymbol{S}_{\boldsymbol{x}})$, где $\boldsymbol{S}_{\boldsymbol{x}} = (-\infty, \boldsymbol{x}) \in \mathcal{B}(\boldsymbol{R}^{\boldsymbol{n}})$. Поэтому нам необходимо иметь признак независимости случайных величин, которые изучаются путём анализа их функций распределения.

Будем рассматривать случай $\boldsymbol{n}$=2. Обобщение рассмотренного критерия на случай произвольного $\boldsymbol{n}$ не вызовет больших затруднений.

Итак, рассматривается двумерная случайная величина $\xi = (\xi_1, \xi_2)$. Функция двух переменных $\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{y})$ - функция распределения этой двумерной случайной величины $\xi$, $\boldsymbol{F}_1(\boldsymbol{x})$ и $\boldsymbol{F}_2(\boldsymbol{y})$ частные функции распределения компонент $\xi_1$ и $\xi_2$.

Теорема 1. *Для того чтобы случайные величины $\xi_1$ и $\xi_2$ были независимыми необходимо и достаточно, чтобы для любой точки $(\boldsymbol{x}, \boldsymbol{y}) \in \boldsymbol{R}^2$ функция распределения двумерной случайной величины была равна произведению частных функций распределения компонент:*

$$\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{F}_1(\boldsymbol{x}) \cdot \boldsymbol{F}_2(\boldsymbol{y}). \qquad (2)$$

Доказательство.

*Необходимость.* Дано: случайные величины $\xi_1$ и $\xi_2$ - независимые, то есть выполняется равенство (1).

Надо доказать, что равенство (2) выполняется для любой точки $(\boldsymbol{x}, \boldsymbol{y}) \in \boldsymbol{R}^2$.

Равенство (1) выполняется для любого борелевского множества $\boldsymbol{B} \subset \boldsymbol{R}^2$, являющегося декартовым произведением $\boldsymbol{B} = \boldsymbol{B}_1 \times \boldsymbol{B}_2$ множеств $\boldsymbol{B}_1 \subset \boldsymbol{R}_1$ и

$\boldsymbol{B}_2 \subset \boldsymbol{R}_2$. Мы можем выбрать $\boldsymbol{B} = \boldsymbol{S}_{xy}$, где точка $(\boldsymbol{x},\boldsymbol{y}) \in \boldsymbol{R}^2$ - произвольная. Множество $\boldsymbol{S}_{xy} \subset \boldsymbol{R}^2$ является декартовым произведением множеств $\boldsymbol{S}_x \subset \boldsymbol{R}_1$ и $\boldsymbol{S}_y \subset \boldsymbol{R}_2$, то есть $\boldsymbol{S}_{xy} = \boldsymbol{S}_x \times \boldsymbol{S}_y$. Но тогда равенство (1) будет иметь вид: $\boldsymbol{P}(\{(\xi_1;\xi_2) \in \boldsymbol{S}_{xy}\}) = \boldsymbol{P}_1(\{\xi_1 \in \boldsymbol{S}_x\}) \cdot \boldsymbol{P}_2(\{\xi_2 \in \boldsymbol{S}_y\})$.

Воспользуемся определением функции распределения случайной величины и получим равенство (2):

$$\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{F}_1(\boldsymbol{x}) \cdot \boldsymbol{F}_2(\boldsymbol{y}). \qquad (2)$$

*Достаточность.* Дано: для функций распределения случайных величин $\xi$, $\xi_1$ и $\xi_2$ справедливо равенство (2): $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{F}_1(\boldsymbol{x}) \cdot \boldsymbol{F}_2(\boldsymbol{y})$.

Доказать, что случайные величины $\xi_1$ и $\xi_2$ - независимые, то есть - выполняется равенство (1).

Сначала докажем, что равенство (1) выполняется, если случайное событие $\boldsymbol{B}$ - прямоугольник $[\boldsymbol{a},\boldsymbol{b})$, а его проекции на координатные оси - это полузакрытые слева интервалы $\boldsymbol{B}_1 = [\boldsymbol{a}_1,\boldsymbol{b}_1)$ и $\boldsymbol{B}_2[\boldsymbol{a}_2,\boldsymbol{b}_2)$.

Согласно четвёртому свойству функция распределения $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$ многомерной случайной величины, для любого прямоугольника $[\boldsymbol{a},\boldsymbol{b}) \subset \boldsymbol{R}^2$ разностный оператор $\Delta_{[\boldsymbol{a},\boldsymbol{b})}\boldsymbol{F}$ принимает неотрицательные значения.

Определим вероятность случайного события $\{\xi \in [\boldsymbol{a},\boldsymbol{b})\}$ как значение разностного оператора от функции $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})$: $\boldsymbol{P}(\xi \in [\boldsymbol{a},\boldsymbol{b})) = \Delta_{[\boldsymbol{a},\boldsymbol{b})}\boldsymbol{F}$ .

Согласно определению разностного оператора запишем:

$\boldsymbol{P}(\xi \in [\boldsymbol{a},\boldsymbol{b})) = \boldsymbol{F}(\boldsymbol{b}_1,\boldsymbol{b}_2) - \boldsymbol{F}(\boldsymbol{a}_1,\boldsymbol{b}_2) - \boldsymbol{F}(\boldsymbol{b}_1,\boldsymbol{a}_2) + \boldsymbol{F}(\boldsymbol{a}_1,\boldsymbol{a}_2)$. В каждой точке – вершине прямоугольника применим равенство (2):

$$\boldsymbol{P}(\xi \in [\boldsymbol{a},\boldsymbol{b})) = \boldsymbol{F}_1(\boldsymbol{b}_1) \cdot \boldsymbol{F}_2(\boldsymbol{b}_2) - \boldsymbol{F}_1(\boldsymbol{a}_1) \cdot \boldsymbol{F}_2(\boldsymbol{b}_2) - \boldsymbol{F}_1(\boldsymbol{b}_1) \cdot \boldsymbol{F}_2(\boldsymbol{a}_2) + \boldsymbol{F}_1(\boldsymbol{a}_1) \cdot \boldsymbol{F}_2(\boldsymbol{a}_2).$$

Дальнейшие преобразования не нуждаются в пояснении:

$$\boldsymbol{P}(\xi \in [\boldsymbol{a},\boldsymbol{b})) = [\boldsymbol{F}_1(\boldsymbol{b}_1) - \boldsymbol{F}_1(\boldsymbol{a}_1)] \cdot \boldsymbol{F}_2(\boldsymbol{b}_2) - [\boldsymbol{F}_1(\boldsymbol{b}_1) - \boldsymbol{F}_1(\boldsymbol{a}_1)] \cdot \boldsymbol{F}_2(\boldsymbol{a}_2);$$

$$\boldsymbol{P}(\xi \in [\boldsymbol{a},\boldsymbol{b})) = [\boldsymbol{F}_1(\boldsymbol{b}_1) - \boldsymbol{F}_1(\boldsymbol{a}_1)] \cdot [\boldsymbol{F}_2(\boldsymbol{b}_2) - \boldsymbol{F}_2(\boldsymbol{a}_2)] = \Delta_{[\boldsymbol{a}_1,\boldsymbol{b}_1)}\boldsymbol{F} \cdot \Delta_{[\boldsymbol{a}_2,\boldsymbol{b}_2)}\boldsymbol{F};$$

$$\boldsymbol{P}(\xi \in [\boldsymbol{a},\boldsymbol{b})) = \boldsymbol{P}_1(\xi_1 \in [\boldsymbol{a}_1,\boldsymbol{b}_1)) \cdot \boldsymbol{P}_2(\xi_2 \in [\boldsymbol{a}_2,\boldsymbol{b}_2)).$$

То есть, если множество $\boldsymbol{B}$ – прямоугольник $B = [a,b) = [a_1,b_1) \times [a_2,b_2)$, то равенство (1) выполняется.

Далее поступаем так же, как и в случае, когда мы рассматривали возможность по функции распределения $\boldsymbol{F}(\boldsymbol{x})$ построения вероятностной функции $\boldsymbol{P}$. К элементам системы прямоугольников $\{[\boldsymbol{a},\boldsymbol{b}]\} \subset \boldsymbol{R}^2$ применяем не более чем счётное число операций объединения, пересечения, дополнения. Из получающихся множеств выбираем только те, которые можно представить в виде декартового произведения $\boldsymbol{B} = \boldsymbol{B}_1 \times \boldsymbol{B}_2$. Пользуясь непрерывностью «сверху», «снизу» и «в нуле» вероятностной функции $\boldsymbol{P}$, убеждаемся, что если $\boldsymbol{B} = \boldsymbol{B}_1 \times \boldsymbol{B}_2$, то будет выполняться равенство (1):

$$\boldsymbol{P}((\xi_1,\xi_2) \in \boldsymbol{B}) = \boldsymbol{P}_1(\xi_1 \in \boldsymbol{B}_1) \cdot \boldsymbol{P}_2(\xi_2 \in \boldsymbol{B}_2).$$

***Комментарий***

Понятие независимости играет в определённом смысле центральную роль в теории вероятностей.

В практических задачах обычно рассматриваются случайные величины непрерывного и дискретного типов. Поэтому естественно знать критерий независимости случайных величин в форме, учитывающей тип случайной величины.

Сначала вспомним формы записи распределения вероятностей двумерной случайной величины и распределений её компонент – одномерных случайных величин.

Если двумерная случайная величина $\xi = (\xi_1, \xi_2)$ непрерывного типа, то распределение вероятностей её возможных значений описывается плотностью вероятности $\boldsymbol{p}(\boldsymbol{x}, \boldsymbol{y})$, при этом $\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{y}) = \int\limits_{-\infty}^{\boldsymbol{x}} \int\limits_{-\infty}^{\boldsymbol{y}} \boldsymbol{p}(\boldsymbol{u}, \boldsymbol{v}) \boldsymbol{dudv}$. Частные плотности вероятности её компонент $\xi_1$ и $\xi_2$ будут равны: $\boldsymbol{p}_1(\boldsymbol{x}) = \int\limits_{\boldsymbol{R}_2} \boldsymbol{p}(\boldsymbol{x}, \boldsymbol{v}) \boldsymbol{dv}$ и $\boldsymbol{p}_2(\boldsymbol{y}) = \int\limits_{\boldsymbol{R}_1} \boldsymbol{p}(\boldsymbol{u}, \boldsymbol{y}) \boldsymbol{du}$.

Если двумерная случайная величина $\xi = (\xi_1, \xi_2)$ дискретного типа, то распределение вероятностей её возможных значений описывается конечным или счётным набором вероятностей $\{\boldsymbol{p}_{\boldsymbol{i},\boldsymbol{k}}\}$, $\boldsymbol{i} = 1,2,\ldots,\boldsymbol{n},\ldots; \boldsymbol{k} = 1,2,\ldots,\boldsymbol{m},\ldots$, при этом $\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{y}) = \sum\limits_{x_i < x} \sum\limits_{y_k < y} \boldsymbol{p}_{\boldsymbol{ik}}$. Частные распределения вероятностей $\{\boldsymbol{p}_{\boldsymbol{i}*}\}$, $\boldsymbol{i} = 1,2,\ldots,\boldsymbol{n},\ldots$ и $\{\boldsymbol{p}_{*\boldsymbol{k}}\}$, $\boldsymbol{k} = 1,2,\ldots,\boldsymbol{m},\ldots$ её компонент $\xi_1$ и $\xi_2$ определяются из условий $\boldsymbol{p}_{\boldsymbol{i}*} = \sum\limits_{\boldsymbol{k}=1}^{\boldsymbol{m},\infty} \boldsymbol{p}_{\boldsymbol{ik}}$ для каждого $\boldsymbol{i}$, и $\boldsymbol{p}_{*\boldsymbol{k}} = \sum\limits_{\boldsymbol{i}=1}^{\boldsymbol{n},\infty} \boldsymbol{p}_{\boldsymbol{ik}}$ для каждого $\boldsymbol{k}$.

Познакомимся теперь с формами критерия независимости случайных величин, учитывающими тип случайной величины.

**Теория**

Теорема 2. *Для того чтобы случайные величины непрерывного типа $\xi_1$ и $\xi_2$ были независимыми необходимо и достаточно, чтобы плотность вероятности $\boldsymbol{p}(\boldsymbol{x}, \boldsymbol{y})$ двумерной случайной величины $\xi = (\xi_1, \xi_2)$ была равна произведению частных плотностей вероятности её компонент:*

$$\boldsymbol{p}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{p}_1(\boldsymbol{x}) \cdot \boldsymbol{p}_2(\boldsymbol{y}). \qquad (3)$$

Доказательство.

*Необходимость.* Дано: случайные величины $\xi_1$ и $\xi_2$ - независимые, то есть выполняется равенство (1), а следовательно, выполняется и равенство (2).

Надо доказать справедливость равенства (3).

Так как случайные величины $\xi, \xi_1$ и $\xi_2$ непрерывного типа, то в равенстве (2) заменим функции распределения $\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{F}_1(\boldsymbol{x})$ и $\boldsymbol{F}_2(\boldsymbol{y})$ их выражениями через плотности вероятностей:

$$\int_{-\infty}^{x}\int_{-\infty}^{y} p(u,v)dudv = \int_{-\infty}^{x} p_1(u)du \cdot \int_{-\infty}^{y} p_2(v)dv .$$

От правой и левой части полученного равенства возьмём смешанные производные второго порядка:

$$\frac{\partial^2}{\partial x\partial y}\int_{-\infty}^{x}\int_{-\infty}^{y} p(u,v)dudv = \frac{\partial}{\partial x}\int_{-\infty}^{x} p_1(u)du \cdot \frac{\partial}{\partial y}\int_{-\infty}^{y} p_2(v)dv .$$

Воспользуемся правилом дифференцирования интеграла с переменным верхним пределом и получим равенство (3):

$$p(x,y) = p_1(x)\cdot p_2(y).$$

*Достаточность.* Дано: для плотностей вероятности случайных величин $\xi$, $\xi_1$ и $\xi_2$ выполняется равенство (3): $p(x,y) = p_1(x)\cdot p_2(y)$.

Доказать, что выполняется равенство (2), а значит выполняется равенство (1), то есть случайные величины $\xi_1$ и $\xi_2$ - независимые.

От левой и правой части равенства (3) возьмём кратные интегралы по области $S_{xy} = (-\infty, x)\times(-\infty, y)$. Так как в правой части равенства (3) подинтегральная функция равна произведению двух функций, каждая из которых зависит только от одного аргумента, то можно, учитывая специфичность области интегрирования, от кратного интеграла перейти к повторному интегралу:

$$\int_{-\infty}^{x}\int_{-\infty}^{y} p(u,v)dudv = \int_{-\infty}^{x} p_1(u)du \cdot \int_{-\infty}^{y} p_2(v)dv .$$

Значения этих трёх интегралов для любого значения переменной $(x,y)$ являются значениями трёх функций распределения. Получаем равенство (2):

$$F(x,y) = F_1(x)\cdot F_2(y).$$

Теорема 3. *Для того чтобы случайные величины дискретного типа $\xi_1$ и $\xi_2$ были независимыми необходимо и достаточно, чтобы для любой пары индексов $(i,k)$ выполнялось:*

$$p_{ik} = p_{i*}\cdot p_{*k} . \tag{4}$$

Доказательство теоремы читатель может сделать самостоятельно. Напомним лишь, что функция распределения любой случайной величины – непрерывная слева функция.

Отсюда следует, что, если $x_i$ - возможное значение случайной величины $\xi_1$ дискретного типа, то: $p_{i*} = \Delta_{[x_i;x_{i+0})}F_1(x) = F_1(x_i+0) - F_1(x_i)$. Аналогично, если $y_k$ - возможное значение случайной величины $\xi_2$ дискретного типа, то: $p_{*k} = \Delta_{[y_k;y_k+0)}F_2(y) = F_2(y_k+0) - F_2(y_k)$. И, наконец, для любой пары индексов $(i,k)$ вероятность $p_{ik}$ определяется как значение разностного оператора:

$$p_{ik} = \Delta_{[z;z+0)}F(x,y), \text{ где } [z;z+0) = [x_i;x_i+0)\times[y_k;y_k+0).$$

**Практика**

В примере 1.4.III, где две монеты подбрасываются наудачу один раз, равенство (4) выполняется для любой из четырех пар индексов $(i,k)$. Значит случайные величины $\xi_1$ и $\xi_2$ - независимые. В примере 2.4.III. мы наблюдаем, равенство (4) выполнятся не всегда. Здесь: $p_{01}=\frac{1}{4}=p_{0*}\cdot p_{*1}=\frac{1}{2}\cdot\frac{1}{2}$ и $p_{11}=\frac{1}{4}=p_{1*}\cdot p_{*1}=\frac{1}{2}\cdot\frac{1}{2}$, но: $p_{00}=\frac{1}{4}\neq p_{0*}\cdot p_{*0}=\frac{1}{2}\cdot\frac{1}{4}$ и $p_{12}=\frac{1}{4}\neq p_{1*}\cdot p_{*2}=\frac{1}{2}\cdot\frac{1}{4}$.

Хотя и без вычисления этих вероятностей ясно, что количество гербов, выпавших на одной из монет, вовсе не зависит от количества гербов, выпавших на другой монете. В то же время, во втором примере, мы уверенно можем сказать, что общее число гербов выпавших на двух монетах зависит от результата подбрасывания первой монеты. Но значения случайной величины $\xi_1$ не определяют однозначно возможные значения случайной величины $\xi_2$, то есть, между $\xi_1$ и $\xi_2$ нет функциональной зависимости.

В то же время мы можем сказать, что возможные значения первой случайной величины $\xi_1$ влияют на возможные значения, и их вероятности, случайной величины $\xi_2$. Зависимости такого вида между случайными величинами $\xi_1$ и $\xi_2$ называются *стохастическими*.

## §7.III Интеграл Римана-Стилтьеса

### *Комментарий*

Мы рассматриваем случайные величины двух типов: дискретного и непрерывного, определяемые на вероятностном пространстве $<\Omega;\mathcal{B}(\Omega);P>$. Решая практические задачи определения вероятности наступления случайного события $\{\xi\in B\}$ мы, интерпретируя эту вероятность как долю единичной массы, находящейся над множеством $B$, в зависимости от типа случайной величины суммируем, или интегрируем $\sigma-$аддитивную функцию множеств – вероятностную функцию $P$. «Запас» борелевских множеств – элементов $\sigma-$ алгебры $\mathcal{B}(\Omega)$ очень велик. И не для всякого $B\in\mathcal{B}(\Omega)$ мы сможем подсчитать

$$P(\xi\in B)=\begin{cases}\sum\limits_{x_k\in B}p_k\\ \int\limits_B p(x)dx\end{cases}.$$

Задача осложняется, если мы приходим к необходимости вычисления значения $\sum\limits_k g(x_k)\cdot p_k$ или $\int\limits_B g(x)\cdot p(x)dx$, где $y=g(x)$ - произвольная борелевская функция. Расширение понятия интеграла Римана до интеграла Лебега позволило применить метрическую теорию функций как основу построения теории вероятностей, то есть в общем случае определять $\sigma-$ аддитивную функцию множеств $\mu(B)$ на элементах $\sigma-$алгебры $\mathcal{B}(\Omega)$.

Рассматривая только непрерывные функции $y=g(x)$ и множества $B$ «простого» вида, мы будем при решении задач вместо интеграла Лебега-Стилтьеса использовать интеграл Римана-Стилтьеса от функции $y=g(x)$ по вероятностной мере $P$.

В конструкции интеграла Римана при составлении сумм Дарбу мерой отрезка разбиения $[x_{i+1}, x_i)$ служит мера Лебега – длина этого отрезка $\Delta x = x_{i+1} - x_i$. В конструкции интеграла Римана-Стилтьеса мерой отрезка разбиения $[x_{i+1}, x_i)$ служит вероятностная мера этого отрезка, то есть $P([x_{i+1}, x_i)) = F(x_{i+1}) - F(x_i)$, называемая мерой Лебега-Стилтьеса.

Определим интеграл Римана-Стилтьеса не в общем виде (это делается в курсе «математического анализа»), а в терминологии курса «теория вероятностей», что облегчит в дальнейшем его применение при рассмотрении теоретических вопросов.

**Теория**

Пусть $<\Omega; \mathcal{B}(\Omega); P>$, вероятностное пространство, где $\Omega$ - множество действительных чисел, $\Omega \subseteq R_1$, $\mathcal{B}(\Omega)$ – борелевская $\sigma -$алгебра подмножеств множества $\Omega$, $P$- вероятностная функция, определяющая распределение вероятностей случайной величины $\xi$. Функция $y = g(x)$ - непрерывная функция, определённая на отрезке $[a,b]$.

Пусть $I = \{x_0,...,x_n\}$, где $a = x_0 < x_1 < ... < x_n = b$, некоторое разбиение отрезка $[a,b]$. Обозначим $S_i = \Omega \bigcap [x_{i-1}, x_i)$ и составим три суммы: $z_n = \sum\limits_{i=1}^{n} m_i \cdot P(S_i)$, $Z_n = \sum\limits_{i=1}^{n} M_i \cdot P(S_i)$ и $\sigma_n = \sum\limits_{i=1} g(\tilde{x}_i) \cdot P(S_i)$, где $m_i = \inf\limits_{[x_{i-1}, x_i)} g(x)$, $M_i = \sup\limits_{[x_{i-1}, x_i)} g(x)$ и $\tilde{x}_i \in [x_{i-1}, x_i)$. Эти суммы аналогичны суммам Дарбу, изучаемым в курсе математического анализа. Ясно, что для любого разбиения $I = \{x_0,...,x_n\}$ будет справедливо: $z_n \le \sigma_n \le Z_n$.

Будем увеличивать число $n$ - число точек разбиения отрезка $[a,b]$, то есть пусть $n \to \infty$, но так что $\max\limits_{1 \le i \le n}(x_{i+1} - x_i) \to 0$. Получим три последовательности интегральных сумм: $\{z_n\}$, $\{Z_n\}$, $\{\sigma_n\}$. Можно показать, что при увеличении числа точек разбиения отрезка $[a,b]$ последовательность $\{z_n\}$ - неубывающая и ограничена сверху, а последовательность $\{Z_n\}$ - невозрастающая и ограничена снизу, то есть $\{z_n\}: \quad z_n \overline{\uparrow}$ и $\{Z_n\}: \quad Z_n \underline{\downarrow}$. Значит, эти последовательности имеют пределы. Покажем, что эти пределы совпадают.

Так как $y = g(x)$ непрерывная функция, то при $n \to \infty$ и $\max\limits_{1 \le i \le n}(x_{i+1} - x_i) \to 0$ для любого сколь угодно малого $\varepsilon > 0$ найдётся $n_0$ такое, что при $n > n_0$ будет выполняться $|M_i - m_i| < \varepsilon$ для всех $i = 1,2,..., n$.

Оценим значение разности интегральных сумм:

$|Z_n - z_n| \le \sum\limits_{i=1}^{n} |M_i - m_i| \cdot P(S_i) < \varepsilon \cdot \sum\limits_{i=1}^{n} P(S_i)$. Так как $\bigcup\limits_{i=1}^{n} S_i \subseteq \Omega$, то $\sum\limits_{i=1}^{n} P(S_i) = P\left(\bigcup\limits_{i=1}^{n} S_i\right) \le P(\Omega) = 1$. Значит, для достаточно большого числа $n$ будет выполняться: $|Z_n - z_n| < \varepsilon$. Следовательно, пределы последовательностей $\{z_n\}$,

$\{\sigma_n\}$и $\{Z_n\}$ совпадают. Это общее значение трёх пределов последовательностей называется *интегралом Римана-Стилтьеса* $(R-S)$ *от функции* $g(x)$ *по вероятностной мере* **P**:

$$\lim_{n\to\infty} z_n = \lim_{n\to\infty}\sigma_n = \lim_{n\to\infty} Z_n \overset{def}{=} (R-S)\int_{\Omega\cap[a,b]} g(x)P(dx).$$

Для любого разбиения $I=\{x_0,\dots,x_n\}$ отрезка $[a,b]$ множество $S_i=\Omega\cap[x_{i-1},x_i)$ является элементом алгебры $\mathcal{B}(\Omega)$, а поэтому мы можем записать $P(S_i)=F(x_i)-F(x_{i-1})$, где $F(x)$ - функция распределения соответствующая вероятностной функции $P$. Переходя к пределу последовательностей интегральных сумм при $n\to\infty$ и $\max\limits_{1\le i\le n}(x_i-x_{i-1})\to 0$, множество $S_i$ мы записываем как $dx$, а поэтому значение вероятности $dx$ будет главной частью приращения значения функции распределения на этом множестве, то есть: $P(dx)=dF(x)$. Значит, интеграл Римана-Стилтьеса можно записать в двух формах: $(R-S)\int\limits_{\Omega\cap[a,b]} g(x)P(dx)=(R-S)\int\limits_{\Omega\cap[a,b]} g(x)dF(x)$.

Если $\Omega=R$ и $\mathcal{B}(\Omega)=\mathcal{B}(R)$, то несобственный интеграл Римана-Стилтьеса определяется обычным способом: $\lim\limits_{\substack{a\to-\infty\\ b\to+\infty}}\int\limits_a^b g(x)P(dx)=\int\limits_{-\infty}^{+\infty} g(x)P(dx)$. Мы будем говорить, что несобственный интеграл Римана-Стилтьеса существует, если он сходится абсолютно, то есть если $\int\limits_{-\infty}^{\infty}|g(x)|P(dx)<\infty$.

## Практика

Пример 1. Пусть функция $g(x)\equiv 1$ для всех $x\in R$.

Тогда $(R-S)\int\limits_{-\infty}^{\infty} P(dx)=(R-S)\int\limits_{-\infty}^{\infty} dF(x)=F(\infty)-F(-\infty)=1$.

Пример 2. Пусть множество $B\subset R$ – элемент $\sigma-$алгебры $\mathcal{B}(R)$. Определим функцию $g(x)=\begin{cases}0, & \text{е с л } x\notin B\\ 1, & \text{е с л } x\in B\end{cases}$. Функция $g(x)$ называется *индикатором* случайного события $B$.

Интеграл Римана-Стилтьеса от этой функции $g(x)$ по вероятностной мере $P$ равен вероятности этого случайного события:

$$(R-S)\int_B 1\cdot P(dx)=(R-S)\int_B 1\cdot dF(x)=P(B).$$

## Теория

Мы изучаем случайные величины двух типов: дискретного и непрерывного.

Областью возможных значений $\Omega$ случайной величины дискретного типа $\xi$ будет не более чем счётное множество точек $\{x_k\}, k=1,2,\dots,n,\dots$. Распределение вероятностей этих значений $\xi$ определяется набором положительных чисел $\{p_k\}, k=1,2,\dots,n,\dots$, где $P(\xi=x_k)=p_k$.

Значит, при любом разбиении $I=\{x_0,...,x_n\}$ отрезка $[a,b]$ вероятностная мера подмножества $S_i=\Omega\bigcap[x_{i-1},x_i)$ будет равна $P(S_i)=\sum\limits_{x_k\in[x_{i-1},x_i)}p_k$.

Ясно, что при $n\to\infty$ и $\max\limits_{1\le i\le n}(x_i-x_{i-1})\to 0$ подмножество $S_i$ будет или пустым $S_i=\varnothing$, или – одноточечным $S_i=x_k$, а тогда для достаточно больших значений $n$ будем иметь или $P(dx)=0$, или $P(dx)=P(x_k)=p_k$.

Значит $\lim\limits_{n\to\infty}\sigma_n=\sum\limits_{x_k\in\Omega\bigcap[a,b]}g(x_k)\cdot p_k$. То есть для случайных величин дискретного типа интеграл $(R-S)$ *будет суммой ряда*:

$$(R-S)\int\limits_{\Omega\bigcap[a,b]}g(x)P(dx)=(R-S)\int\limits_{\Omega\bigcap[a,b]}g(x)dF(x)=\sum\limits_{x_k\in\Omega\bigcap[a,b]}g(x_k)\cdot p_k\,.$$

Областью возможных значений $\Omega$ случайной величины непрерывного типа $\xi$ будет множество мощности континуум, мера Лебега которого – положительное число. Плотность вероятности $p(x)=F'(x)$ для всех $x\in\Omega$, за исключением, быть может, множества нулевой меры. Пусть, например, $\Omega=[\alpha,\beta)$. Тогда при $n\to\infty$ и $\max\limits_{1\le i\le n}(x_i-x_{i-1})\to 0$ подмножество $S_i$ будет или пустым $S_i=\varnothing$, или – полуинтервалом $S_i=[x_{i-1},x_i)$. Значит, для достаточно больших значений $n$ можем записать $P(S_i)=p(\tilde{x}_i)\cdot mesS_i$, где $\tilde{x}_i\in[x_{i-1},x_i)$ и $mesS_i=\begin{cases}0\\ x_i-x_{i-1}\end{cases}$. Следовательно, $\lim\limits_{n\to\infty}\sigma_n=\int\limits_{\Omega\bigcap[a,b]}g(x)\cdot p(x)dx$. То есть для случайных величин непрерывного типа интеграл $(R-S)$ *будет интегралом Римана*:

$$(R-S)\int\limits_{\Omega\bigcap[a,b]}g(x)P(dx)=(R-S)\int\limits_{\Omega\bigcap[a,b]}g(x)dF(x)=\int\limits_{\Omega\bigcap[a,b]}g(x)\cdot p(x)dx\,.$$

Использование интеграла Римана-Стилтьеса позволяет изучать свойства случайных величин и функций случайных величин, не отвлекаясь при этом на тип случайной величины, то есть - на тип закона распределения вероятностей случайной величины.

Аналогично вводится определение кратного интеграла Римана-Стилтьеса в случае, когда рассматривается $<\Omega;\mathcal{B}(\Omega);P>$ - вероятностное пространство, где: $\Omega\subseteq R^n$, $\mathcal{B}(\Omega)$ – борелевская $\sigma-$алгебра подмножеств множества $\Omega$, $P$- вероятностная функция, определяющая распределение вероятностей $n$-мерной случайной величины $\xi$, функция $y=g(x_1,..., x_n): R^n\to R_1$, - непрерывная функция $n$ аргументов.

## §8.III Функции случайных величин

### *Комментарий*

Случайные величины $\xi_1$ и $\xi_2$ могут быть независимыми или зависимыми. Независимость случайных величин мы понимаем, как возможность каждой из случайных величин принимать свои значения с вероятностями, определяемыми её вероятностной функцией, независимо от значений, которые принимает другая случайная величина. Случайные величины $\xi_1$ и $\xi_2$ независимые, если вероятность наступления случайного события $\{\xi_1 \in A \wedge \xi_2 \in B\}$ равна произведению вероятностей наступлений случайных событий $\{\xi_1 \in A\}$ и $\{\xi_2 \in B\}$

$$P(\xi_1 \in A \wedge \xi_2 \in B) = P_1(\xi_1 \in A) \cdot P_2(\xi_2 \in B).$$

Независимые случайные величины рассмотрены в примерах 1.4.IV и 3.4.IV.

Если мы говорим, что случайные величины зависимые, то мы имеем в виду, что значения, которые принимает одна из случайных величин, влияют на диапазон возможных значений, а так же на вероятности наступления этих возможных значений, другой случайной величины. Такие зависимости называются *стохастическими*.

Стохастические зависимости между случайными величинами $\xi_1$ и $\xi_2$ бывают двух видов. К первому виду относятся зависимости подобные тем, с которыми мы встретились в примерах 2.4.III, 4.4.III, 5.4.III. В них значения первой случайной величины *влияют* на варианты возможных значений второй случайной величины. Такие зависимости будем называть *статистическими*.

Ко второму виду зависимостей относятся зависимости, в которых значения первой случайной величины *определяют* значения и вероятности этих значений второй случайной величины. Зависимости такого вида будем называть *функциональными*.

Рассмотрим подробнее этот вид зависимостей.

**Теория**

Рассмотрим два измеримых пространства $\langle R_1; \mathcal{B}(R_1)\rangle$ и $\langle R_2; \mathcal{B}(R_2)\rangle$ и пусть $g(x)$- функция такая, что $g(x): R_1 \to R_2$.

Определение. Функция $g(x)$ называется *борелевской*, если она борелевские множества отображает в борелевские, то есть, если множество $B \in \mathcal{B}(R_2)$, то его прообраз - множество $A = g^{-1}(B) \in \mathcal{B}(R_1)$.

При определении случайной величины мы рассматривали вероятностное и измеримое пространства $\langle \Omega, \mathcal{A}, P\rangle$ и $\langle R, \mathcal{B}(R)\rangle$, соответственно.

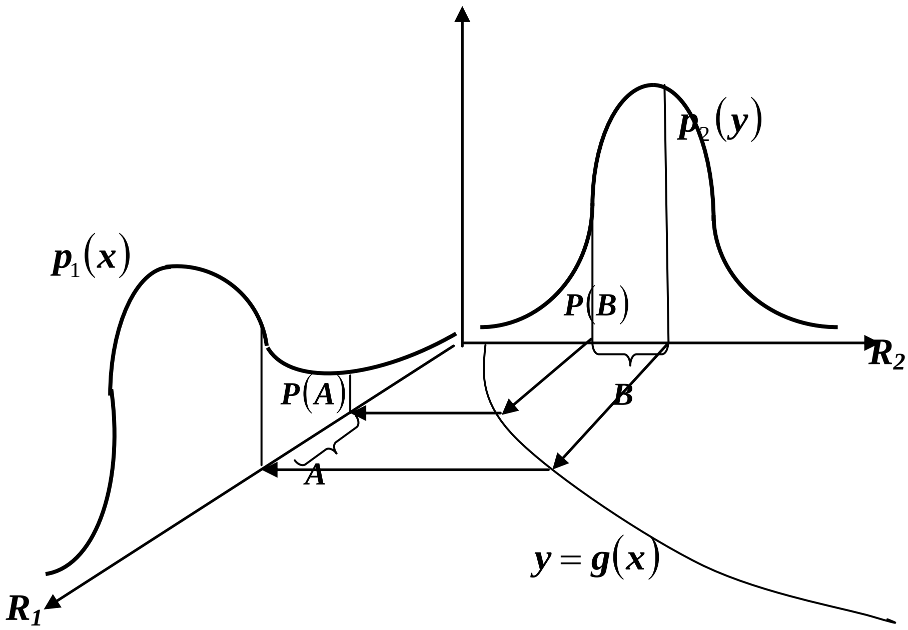
Пусть $\xi(\omega): \Omega \to R_1$ и $g(x): R_1 \to R_2$ - некоторая борелевская функция. Рассмотрим $\eta = g(\xi)$. Так как значения аргумента функции $g(x)$ являются значениями случайной величины $\xi(\omega)$, то значения функции $\eta = g(\xi)$ - случайные, то есть $\eta = g(\xi)$ - случайная величина. Нас интересует вопрос о возможности определения закона распределения вероятностей случайной величины $\eta = g(\xi)$.

Определение. Случайная величина $\eta = g(\xi)$ называется *функцией случайной величины* $\xi(\omega)$, если для любого $B \in \mathcal{B}(R_2)$ будет справедливо: $P(\eta \in B) = P(\xi \in A)$, где $A = g^{-1}(B) \in \mathcal{B}(R_1)$ – полный прообраз множества $B$.

Если для случайных величин $\xi$ и $\eta$ и некоторой борелевской функции $g(x)$ выполняется это определение, то будем говорить, что случайные величины $\xi$ и $\eta$ является *функционально зависимыми*.

Определение функциональной зависимости между случайными величинами иллюстрируется на рисунке, где вероятности наступления событий $\boldsymbol{B}$ и $\boldsymbol{A}$ интерпретируются как площади криволинейных трапеций.

Если $\boldsymbol{\eta} = \boldsymbol{g}(\boldsymbol{\xi})$ - функция случайной величины $\boldsymbol{\xi}$, то для любого множества $\boldsymbol{B} \in \mathcal{B}(\boldsymbol{R_2})$ площадь криволинейной трапеции над множеством $\boldsymbol{B}$, ограниченной плотностью вероятности $\boldsymbol{p}_2(\boldsymbol{y})$, всегда равна площади криволинейной трапеции над прообразом множества $\boldsymbol{B}$ - множеством $\boldsymbol{A} = \boldsymbol{g}^{-1}(\boldsymbol{B})$, которая ограничена плотностью вероятности $\boldsymbol{p}_1(\boldsymbol{x})$.




Покажем, что, при наличии *функциональной* зависимости между случайными величинами $\boldsymbol{\xi}$ и $\boldsymbol{\eta} = \boldsymbol{g}(\boldsymbol{\xi})$, если мы знаем закон распределения вероятностей случайной величины $\boldsymbol{\xi}$, то мы всегда можем найти закон распределения вероятностей случайной величины $\boldsymbol{\eta} = \boldsymbol{g}(\boldsymbol{\xi})$.

Итак, пусть $\boldsymbol{F}_{\xi}(\boldsymbol{x})$ - функция распределения случайной величины $\boldsymbol{\xi}$ и $\boldsymbol{g} : \boldsymbol{R}_1 \to \boldsymbol{R}_2$ - борелевская функция. Случайная величина $\boldsymbol{\eta} = \boldsymbol{g}(\boldsymbol{\xi})$ - функция случайной величины $\boldsymbol{\xi}$. Найдём функцию распределения случайной величины $\boldsymbol{\eta} = \boldsymbol{g}(\boldsymbol{\xi})$.

Вспомним определение функции распределения случайной величины: $\boldsymbol{F}_{\eta}(\boldsymbol{y}) = \boldsymbol{P}(\boldsymbol{\eta} \in \boldsymbol{S}_y)$ и что вероятность наступления любого события равна приращению значения функции распределения на этом событии – элементе алгебры $\mathcal{B}(\boldsymbol{R})$. Дальнейшие преобразования не нуждаются в пояснениях:

$$\boldsymbol{F}_{\eta}(\boldsymbol{y}) = \boldsymbol{P}(\boldsymbol{\eta} \in \boldsymbol{S}_y) = \boldsymbol{P}(\boldsymbol{g}(\boldsymbol{\xi}) \in \boldsymbol{S}_y) = \boldsymbol{P}(\boldsymbol{\xi} \in \boldsymbol{g}^{-1}(\boldsymbol{S}_y)) = \int\limits_{g^{-1}(S_y)} \boldsymbol{dF}_{\xi}(\boldsymbol{x}).$$

Если случайная величина $\boldsymbol{\xi}$ непрерывного типа, то $\boldsymbol{dF}_{\xi}(\boldsymbol{x}) = \boldsymbol{p}_{\xi}(\boldsymbol{x})\boldsymbol{dx}$ и тогда $\boldsymbol{F}_{\eta}(\boldsymbol{y}) = \int\limits_{g^{-1}(S_y)} \boldsymbol{dF}_{\xi}(\boldsymbol{x}) = \int\limits_{g^{-1}(S_y)} \boldsymbol{p}_{\xi}(\boldsymbol{x})\boldsymbol{dx}$. Значит, плотность вероятности $\boldsymbol{p}_{\eta}(\boldsymbol{y})$ случайной величины $\boldsymbol{\eta}$ будет определяться как производная от определённого интеграла, у которого пределы интегрирования – функции переменной $\boldsymbol{y}$:

$$p_\eta(y)=F'_\eta(y)=p_\xi\left(g^{-1}(S_y)\right)\cdot\left|\left(g^{-1}(y)\right)'\right|.$$

**Практика**

Пример 1. Пусть $F_\xi(x)$ - функция распределения случайной величины $\xi$ и $g(x)=ax+b$ - линейная функция. Найдем функцию распределения $F_\eta(y)$ случайной величины $\eta=g(\xi)=a\xi+b$.

$$F_\eta(y)=P(\eta\in S_y)=P(\eta<y)=P(a\xi+b<x)=P\left(\xi<\frac{y-b}{a}\right)=\int_{-\infty}^{\frac{y-b}{a}}dF_\xi(x)=F_\xi\left(\frac{y-b}{a}\right)$$

Пример 2. Пусть $F_\xi(x)$ - функция распределения случайной величины $\xi$ и $g(x)=x^2$ - степенная функция. Найдем функцию распределения $F_\eta(y)$ случайной величины $\eta=g(\xi)=\xi^2$.

$$F_\eta(y)=P(\eta<y)=P(\xi^2<y)=P\left(-\sqrt{y}<\xi<+\sqrt{y}\right)=\text{е}$$

$$=\begin{cases}0, & \text{если } y\le 0\\ F_\xi\left(\sqrt{y}\right)-F_\xi\left(-\sqrt{y}\right), & \text{если } y>0\end{cases}.$$

В частности, если случайная величина имеет нормальное распределение вероятностей, $\xi\in\mathcal{N}(0;1)$, то плотность вероятности случайной величины $\eta=g(\xi)=\xi^2$ для $y>0$ будет равна: $p_\eta(y)=p_\xi\left(\sqrt{y}\right)\cdot\left(\sqrt{y}\right)'+p_\xi\left(-\sqrt{y}\right)\cdot\left|\left(-\sqrt{y}\right)'\right|$, то есть

$$p_\eta(y)=\begin{cases}0, & \text{если } y\le 0\\ \dfrac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{y}{2}}, & \text{если } y>0\end{cases}.$$

**Теория**

Рассмотрим теперь случай, когда рассматривается векторная случайная величина $\xi=(\xi_1,\xi_2,\ldots,\xi_n)$ и $\eta=g(\xi_1,\xi_2,\ldots,\xi_n)$ - функция случайных величин – компонент случайной величины $\xi$.

Функция распределения случайной величины $\eta$ определяется как значение $\boldsymbol{n}$-кратного интеграла от функции распределения векторной случайной величины $\xi$ по области $g^{-1}(S_y)$:

$$F_\eta(y)=P(\eta\in S_y)=P(g(\xi)\in S_y)=P\left(\xi\in g^{-1}(S_y)\right)=\int\limits_{g^{-1}(S_y)}\!\!\ldots\int dF_\xi(x_1,\ldots,x_n).$$

В практических задачах и в предельных теоремах теории вероятностей рассматривается векторная случайная величина $\xi=(\xi_1,\xi_2,\ldots,\xi_n)$, компоненты которой независимые случайные величины, и $y=g(x)$ - линейная функция $\boldsymbol{n}$

аргументов: $y = g(x_1,...,x_n) = x_1 + x_2 + ... + x_n$. Определяется случайная величина $\eta = g(\xi_1,..., \xi_n) = \sum_{k=1}^{n} \xi_k$ и требуется найти функцию распределения случайной величины $\eta$.

Рассмотрим сначала случай $n$=2. Так как компоненты случайной величины $\xi = (\xi_1, \xi_2)$ независимые, то функция распределения вектора равна произведению частных функций распределения компонент $F_\xi(x_1, x_2) = F_1(x_1) \cdot F_2(x_2)$ и, следовательно, $dF_\xi(x_1, x_2) = dF_1(x_1) \cdot dF_2(x_2)$. Функция распределения $F_\eta(y)$ случайной величины $\eta = \xi_1 + \xi_2$ будет определяться из равенства:

$$F_\eta(y) = P(\eta < y) = P(\xi_1 + \xi_2 < y) = \iint_{x_1 + x_2 < y} dF_1(x_1) \cdot dF_2(x_2).$$

Применяя теорему Фубини о переходе от кратного интеграла к повторному интегралу, получаем:

$$F_\eta(y) = \iint_{x_1 + x_2 < y} dF_1(x_1) \cdot dF_2(x_2) = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{y - x_1} dF_2(x_2) \right) dF_1(x_1) = \int_{-\infty}^{\infty} (F_2(y - x_1) - F_2(-\infty)) dF_1(x_1)$$

То есть: $F_\eta(y) = \int_{-\infty}^{\infty} F_2(y - x_1) dF_1(x_1)$.

<u>Определение.</u> Если $F_1(x_1)$ и $F_2(x_2)$ - функции распределения, то функцию $F(y)$, получаемую из равенства $F(y) = \int_{-\infty}^{\infty} F_2(y - x_1) dF_1(x_1)$, называют *свёрткой функций* распределения $F_1(x_1)$ и $F_2(x_2)$, и обозначают $F = F_1 * F_2$.

Очевидно, что аналогично можем получить: $F_\eta(y) = \int_{-\infty}^{\infty} F_1(y - x_2) dF_2(x_2)$. То есть свёртка функций распределения $F_1(x_1)$ и $F_2(x_2)$ - *коммутативна*: $F = F_1 * F_2 = F_2 * F_1$.

Пусть теперь $n$=3. Если компоненты случайной величины $\xi = (\xi_1, \xi_2, \xi_3)$ независимые и $\eta = g(\xi_1, \xi_2, \xi_3) = \xi_1 + \xi_2 + \xi_3$, то для определения функции распределения случайной величины $\eta$ мы воспользуемся свойством ассоциативности операции сложения: $a + b + c = (a + b) + c$. Обозначим $\zeta = \xi_1 + \xi_2$, тогда $\eta = \xi_1 + \xi_2 + \xi_3 = (\xi_1 + \xi_2) + \xi_3$. Для вычисления функции распределения $F_\eta(y)$ сначала вычислим функцию распределения $F_\zeta(z)$ как свёртку функций $F_1(x_1)$ и $F_2(x_2)$: $F_\zeta = F_1 * F_2$. Затем вычислим функцию распределения $F_\eta(y)$ как свёртку функций $F_\zeta(z)$ и $F_3(x_3)$: $F_\eta = F_\xi * F_3$.

Значит, свёртка функций распределения обладает свойством *ассоциативности:* $F = F_1 * F_2 * F_3 = (F_1 * F_2) * F_3 = F_1 * (F_2 * F_3)$.

Свойство ассоциативности позволяет определить функцию распределения случайной величины, являющуюся суммой любого конечного числа независимых случайных величин.

Если случайная величина $\xi=(\xi_1,\xi_2)$ - непрерывного типа и её компоненты независимые, то случайная величина $\eta=\xi_1+\xi_2$ будет непрерывного типа и её плотность вероятности будет равна:

$$p_\eta(y)=F'_\eta(y)=\int_{-\infty}^{\infty}F'_2(y-x_1)\cdot p_1(x_1)dx_1=\int_{-\infty}^{\infty}p_2(y-x_1)\cdot p(x_1)dx_1 .$$

Плотность вероятности, определяемую по формуле:

$p_\eta(y)=\int_{-\infty}^{\infty}p_2(y-x_1)\cdot p(x_1)dx_1$, называется *свёрткой плотностей вероятности* $p_1(x_1)$ и $p_2(x_2)$, и обозначают $p=p_1*p_2$.

**Практика**

Пример 3. Пусть случайные величины $\xi_1$ и $\xi_2$ независимые и подчиняются нормальному распределению вероятностей $\mathcal{N}(0;1)$. Найти плотность вероятности случайной величины $\eta=\xi_1^2+\xi_2^2$.

В примере 2 этого параграфа мы нашли плотности вероятности случайных величин-слагаемых $\xi_1^2$ и $\xi_2^2$:

$$p_{\xi_i^2}(x_i)=\begin{cases}0, & \text{если } x_i\le 0\\ \frac{1}{\sqrt{2\pi}}x_i^{-\frac{1}{2}}\cdot e^{-\frac{x_i}{2}}, & \text{если } x_i>0\end{cases},\quad \text{где } i=1,2.$$

Так как случайные величины $\xi_1$ и $\xi_2$ независимые, то плотность вероятности случайной величины $\eta$ будет равна свёртке плотностей вероятности случайных величин $\xi_1^2$ и $\xi_2^2$:

$$p(y)=p_1*p_2=\int_0^{\infty}p_2(y-x_1)\cdot p_1(x_1)dx_1=\frac{1}{\sqrt{2\pi}}\int_0^y(y-x_1)^{-\frac{1}{2}}e^{-\frac{y-x_1}{2}}\cdot\frac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{x_1}{2}}dx_1=$$

$$=\frac{1}{2\pi}\cdot\frac{2}{y}\cdot e^{-\frac{y}{2}}\int_0^y\frac{dx_1}{\sqrt{1-\left(1-\frac{2x_1}{y}\right)^2}}.$$

Сделаем замену переменной интегрирования: пусть $1-\frac{2x_1}{y}=t$, тогда $dx_1=-\frac{y}{2}dt$ и для $y>0$ получаем:

$$p_{(2)}=p(y)=\frac{1}{2\pi}\cdot e^{-\frac{y}{2}}\int_{-1}^{1}\frac{dt}{\sqrt{1-t^2}}=\frac{1}{2\pi}\cdot e^{-\frac{y}{2}}\cdot \arcsin t\Big|_{-1}^{+1}=\frac{1}{2}e^{-\frac{y}{2}}.$$

Продолжим вычисление плотности вероятности случайной величины $\eta$, полагая $\eta=\xi_1^2+\xi_2^2+\xi_3^2$, где случайные величины $\xi_1,\xi_2$ и $\xi_3$ независимые и подчиняются нормальному распределению вероятностей $\mathcal{N}(0;1)$.

Ясно, что $p_{(3)}=p(z)=p_1*p_2*p_3=(p_1*p_2)*p_3$, где $(p_1*p_2)=\frac{1}{2}e^{-\frac{y}{2}}=p_{(2)}$. Значит:

$$p(z)=\int_0^\infty p_{(2)}(z-x_3)\cdot p_3(x_3)dx_3=\frac{1}{2}\int_0^z e^{-\frac{z-x_3}{2}}\cdot\frac{x_3^{-\frac{1}{2}}\cdot e^{-\frac{x_3}{2}}}{\sqrt{2\pi}}dx_3=\frac{1}{2^{\frac{3}{2}}\sqrt{\pi}}e^{-\frac{z}{2}}\int_0^z\frac{1}{\sqrt{x_3}}dx_3.$$

Окончательно получаем: $p_{(3)}=p(z)=\frac{1}{\sqrt{2\pi}}z^{\frac{1}{2}}\cdot e^{-\frac{z}{2}}$, где $z>0$.

Если мы будем увеличивать число слагаемых, то, учитывая ассоциативность свёртки плотностей вероятности и применяя метод математической индукции, приходим к выводу:

Если $\xi_1,\xi_2,...,\xi_n$ - независимые, одинаково распределённые по закону $\mathcal{N}(0;1)$, случайные величины, то случайную величину, являющуюся суммой квадратов независимых, нормально $\mathcal{N}(0;1)$ распределённых случайных величин, обозначают $\chi_n^2$, то есть $\chi_n^2=\sum_{k=1}^{n}\xi_k^2$, и называют *случайной величиной, подчиняющейся закону Пирсона, или - закону «хи-квадрат» с параметром* $\boldsymbol{n}$.

Случайная величина $\chi_n^2$ имеет плотность вероятности:

$$p_n(x)=\begin{cases}0, & \text{если} \quad x\le 0\\ \dfrac{1}{2^{n/2}\Gamma(n/2)}\cdot x^{n/2-1}\cdot e^{-x/2}, & \text{если} \quad x>0.\end{cases}$$

Полученная плотность вероятности играет важную роль в математической статистике, где изучаются статистические зависимости между случайными величинами.

<u>Пример 4.</u> Пусть $\chi_3^2=\sum_{k=1}^{3}\xi_k^2$ случайная величина подчиняется закону «хи-квадрат» с параметром $n$=3 и функция $y=g(x)=\sqrt{x}$. Найдём плотность вероятности случайной величины $\eta=g(\chi_3^2)=\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}$. Обратная функция $x=g^{-1}(y)=y^2$- возрастающая, а поэтому будем использовать формулу: $p_\eta(y)=p_\xi(g^{-1}(y))\cdot(g^{-1}(y))'$, здесь $\xi=\chi_3^2$. Значит: $p_\eta(y)=\frac{1}{2\pi}\cdot(y^2)^{\frac{1}{2}}\cdot e^{-\frac{y^2}{2}}\cdot(2y)$.

Окончательно получаем плотность вероятности закона, называемого законом Максвелла: $p(y)=\sqrt{\frac{2}{\pi}}\cdot y^2\cdot e^{-\frac{y^2}{2}}$, если $y>0$.

Это распределение вероятностей описывает распределение случайной величины $\eta$ - скорость молекул газа в некотором замкнутом объёме.

## Теория

Если случайная величина $\xi=(\xi_1,\xi_2)$ - дискретного типа, то случайная величина $\eta=\xi_1+\xi_2$ будет случайной величиной дискретного типа.

Если компоненты $\xi_1$ и $\xi_2$ - независимые случайные величины и приняли значения $(x_i)_1$ и $(x_k)_2$ с вероятностями $p_{i*}$ и $p_{*k}$, соответственно, то случайная величина $\eta$ примет значение $y_s = (x_i)_1 + (x_k)_2$ с вероятностью $P(\eta = y_s) = p_s = p_{i*} \cdot p_{*k}$. Очевидно, что для некоторых возможных значений $y_s$ случайной величины $\eta$ может существовать несколько пар значений $(x_i)_1$ и $(x_k)_2$, сумма которых будет равна этому $y_s$. Следовательно, распределение вероятностей $\{p_s\}$ случайной величины $\eta$ будем определять, вычисляя для каждого возможного значения $y_s$ суммы произведений вероятностей $p_{i*} \cdot p_{*k}$ по тем индексам $i$ и $k$, при которых $y_s = (x_i)_1 + (x_k)_2$.

Случайную величину дискретного типа называют *арифметической*, если множеством её возможных значений $\Omega$ является множество целых чисел $\{x_i\}, i = 1,2,...$, образующих арифметическую прогрессию: $x_{i+1} = x_i + d$, разность прогрессии $d$ называется *шагом распределения*.

Например, случайная величина $\xi$, подчиняющаяся биномиальному распределению $\xi \in B_n(p)$, - арифметическая случайная величина. Здесь $\Omega = \{i\}, i = 0,1,...,n$ и $d$=1; $P(\xi = i) = p_i = C_n^i p^i q^{n-i}$.

Пусть компоненты $\xi_1$ и $\xi_2$ случайной величины $\xi = (\xi_1, \xi_2)$ - независимые, арифметические случайные величины. Их возможные значения: $(x_i)_1 = i$ и $(x_k)_2 = k$. Последовательности $\{p_{i*}\}$ и $\{p_{*k}\}$ - распределения вероятностей $\xi_1$ и $\xi_2$. Случайная величина $\eta = \xi_1 + \xi_2$ принимает значения $y_s = s = i + k$. Так как для некоторых значений $s$ может существовать несколько пар $(i,k)$таких что $s = i + k$, то вероятности $p_s$ этих возможных значений $s$ определяются по формуле: $p_s = \sum_i p_{i*} \cdot p_{*s-i}$, где суммирование проводится по тем значениям $i$ первой случайной величины $\xi_1$, при которых разности $s - i$ будут равны возможным значениям $k$ второй случайной величины $\xi_2$. Эта формула будет упрощённой формулой свёртки дискретных распределений.

## Практика

Пример 5. Пусть $\xi_1$ и $\xi_2$ независимые случайные величины, подчиняющиеся бернуллиевскому распределению $B_1(p)$:

| $\xi_1$ | 0 | 1 |
|---|---|---|
| | $q = p_{0*}$ | $p = p_{1*}$ |

| $\xi_2$ | 0 | 1 | |
|---|---|---|---|
| | $q = p_{*0}$ | $p = p_{*1}$ | |

Определим случайную величину $\eta = \xi_1 + \xi_2$ и построим ряд распределения этой случайной величины. Ясно, что случайная величина $\eta$ может принять одно из трёх значений: $s$=0, или $s$=1, или $s$=2. Применяя упрощённую формулу свёртки, определяем вероятности этих значений:

$$P(\eta = 0) = p_0 = \sum_{i=0} p_{i*} \cdot p_{*0-i} = p_{0*} \cdot p_{*0} = q^2;$$

$$P(\eta = 1) = p_1 = \sum_{i=0}^{1} p_{i*} \cdot p_{*1-i} = p_{0*} \cdot p_{*1-0} + p_{1*} \cdot p_{*1-1} = p_{0*} p_{*1} + p_{1*} p_{*0} = 2qp;$$

$$P(\eta = 2) = p_2 = \sum_{i=1} p_{i*} \cdot p_{*2-i} = p_{1*} \cdot p_{*1} = p^2 .$$

Ряд распределения случайной величины $\eta$ будет иметь вид:

| $\eta$ | 0 | 1 | 2 |
|---|---|---|---|
| | $q^2 = p_{0*}$ | $2qp = p_{1*}$ | $p^2 = p_{2*}$ |

Из записи ряда распределения видно, что случайная величина $\eta = \xi_1 + \xi_2$ подчиняется биномиальному распределению, то есть $\eta \in B_2(p)$.

Методом математической индукции можно показать, что если $\xi_1, \xi_2, ..., \xi_n$ независимые случайные величины, подчиняющиеся бернуллиевскому распределению $B_1(p)$, то их сумма - это случайная величина $\eta = \sum_{k=1}^{n} \xi_k$, подчиняющаяся биномиальному распределению $B_n(p)$.

## ГЛАВА ЧЕТВЁРТАЯ
## Числовые характеристики случайных величин и векторов

### §1.IV Математическое ожидание случайной величины
### Теория

Рассматривается вероятностное пространство <$R^n$,$\mathcal{B}(R^n)$,$P$ >, где $P$ – вероятностная функция, определяющая распределение вероятностей $n$-мерной случайной величины $\xi$. Пусть $y = g(x)$ - непрерывная функция $n$ переменных, то есть $g(x_1, ..., x_n): R^n \to R$. Образуем функцию $g(\xi)$ случайной величины $\xi$.

Определение 1.

***Математическим ожиданием*** *функции* $g(\xi)$ *по вероятностной мере* $P$ *называется интеграл Римана-Стилтьеса:*

$$M[g(\xi)] = \int_{R^n} ... \int g(x) P(dx) = \int_{R^n} ... \int g(x) dF(x),$$

*если интеграл сходится абсолютно*.

Ясно, что все сведения, вся информация о возможных значениях случайной величины $\xi$ и о распределении вероятностей этих значений содержатся в конкретной записи вероятностного пространства. Значение определённого интеграла это некоторое число. Из определения видно, что математическое ожидание специально подобранных функций $y = g(x)$ призвано давать некоторую важную, но частичную числовую информацию об изучаемой случайной величине $\xi$.

Чтобы понять: какую информацию о случайной величине $\xi$ даёт нам значение интеграла при конкретно выбранной непрерывной функции $y = g(x)$, рассмотрим несколько примеров наиболее часто встречающихся вероятностных функций $P$.

Пусть функция $y = g(x)$ - линейная, то есть $g(x) = x$.

<u>Определение 2.</u>

***Математическим ожиданием случайной величины*** $\xi$ *называется значение интеграл Римана-Стилтьеса:* $M[\xi] = \int\limits_{R_1} xP(dx) = \int\limits_{R_1} xdF(x)$, *если этот интеграл сходится абсолютно.*

Так как мы рассматриваем случайные величины двух типов, то, соответственно, для дискретного и непрерывного типа запишем вид интеграла Римана-Стилтьеса:

$$M[\xi] = \int\limits_{R_1} xdF(x) = \begin{cases} \sum\limits_{k=1}^{n,\infty} x_k \cdot p_k, \\ \int\limits_{-\infty}^{\infty} x \cdot p(x)dx, \end{cases}$$

если, соответственно, ряд или интеграл сходятся абсолютно.

## Практика

<u>Пример 1.1. IV.</u> Случайная величина $\xi$ дискретного типа, арифметическая, подчиняется равномерному распределению вероятностей. Вычислим значение математического ожидания $M\xi$.

Ряд распределения случайной величины имеет вид:

| $\xi$ | $x_1$ | $x_2$ | $\cdots$ | $x_k$ | $\cdots$ | $x_n$ |
|---|---|---|---|---|---|---|
| $p_k$ | $\frac{1}{n}$ | $\frac{1}{n}$ | $\cdots$ | $\frac{1}{n}$ | $\cdots$ | $\frac{1}{n}$ |

Здесь $x_k = x_1 + d \cdot (k-1),\ k = 1,2,\dots,n$.

Согласно определению математического ожидания для дискретной случайной величины: $M\xi = \sum\limits_{k=1}^{n} x_k \cdot \frac{1}{n} = \frac{1}{n} \cdot \frac{x_1 + x_n}{2} \cdot n = \frac{x_1 + x_n}{2}$.

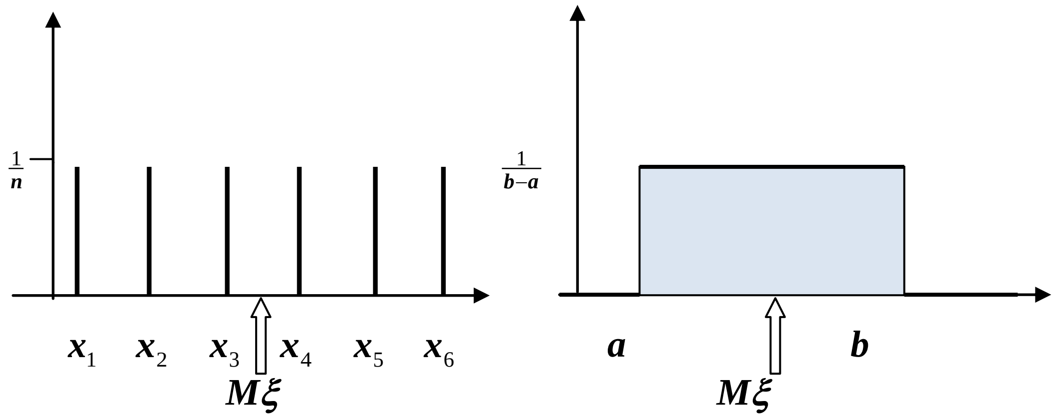
Так как значение математического ожидания это число, то точка $M\xi = \frac{x_1 + x_n}{2}$ на числовой оси находится посередине между точками $x_1$ и $x_n$.

<u>Пример 2.1.IV.</u> Случайная величина $\xi$ непрерывного типа подчиняется равномерному распределению вероятностей на отрезке $[a,b]$. Вычислим значение математического ожидания $M\xi$.

Плотность вероятности равномерного распределения имеет вид: $p(x) = \begin{cases} 0, & x \notin [a,b] \\ \frac{1}{b-a}, & x \in [a,b] \end{cases}$. Согласно определению математического ожидания для непрерывной случайной величины: $M\xi = \int\limits_{-\infty}^{\infty} x \cdot p(x)dx$.

Плотность вероятности $p(x)$ - кусочно-непрерывная функция, поэтому:

$M\xi = \int\limits_{-\infty}^{a} x \cdot 0dx + \int\limits_{a}^{b} x \cdot \frac{1}{b-a}dx + \int\limits_{b}^{\infty} x \cdot 0dx = 0 + \frac{b^2 - a^2}{2 \cdot (b-a)} + 0 = \frac{a+b}{2}$. Полученное значение математического ожидания $M\xi = \frac{a+b}{2}$ интерпретируется точкой на числовой оси, находящейся в центре отрезка $[a,b]$.




***Комментарий***

Рассмотренные два примера позволяют сделать заключение о роли математического ожидания как некоторой числовой характеристики случайной величины. Если распределение вероятностей значений случайной величины интерпретировать как распределение единичной массы на области её возможных значений, то математическое ожидание можно интерпретировать как координату центра тяжести этой единичной массы. То есть, математическое ожидание – это *среднее значение* случайной величины.

Пусть в примере 1.1. IV. испытание состоит в однократном подбрасывании игральной кости, а случайная величина $\xi$ - число «очков», выпавших на верхней грани игральной кости. Ясно, что $x_k = k$, где $k = 1,2,...,6$, а $p_k = \frac{1}{6}$. Математическое ожидание этой случайной величины равно $M\xi = \frac{1+6}{2} = 3{,}5$, или: среднее число «очков», выпадающих на верхней грани при однократном подбрасывании игральной кости, равно 3,5.

Ясно, что подбросив игральную кость, мы никогда на верхней грани не увидим 3,5 «очка». Интуиция, жизненный опыт позволяют нам объяснить значение математического ожидания равное 3,5 следующим образом. Если мы подбросим игральную кость много раз, каждый раз фиксируя число выпавших «очков», а затем сумму всех выпавших «очков» разделим на число сделанных подбрасываний, то полученное число, называемое *средним числом «очков», выпадающих при одном подбрасывании,* будет близко к числу 3,5.

Пусть в примере 2.1. IV. случайная величина $\xi$ - время ожидания автобуса на остановке. Если интервал движения автобусов равен 10 минутам, ($a = 0$ и $b = 10$) то *среднее время* ожидания автобуса на остановке будет равно $M\xi = \frac{0+10}{2} = 5$ минутам.

## §2.IV Примеры значений математических ожиданий
## Практика

Пример 3.1. IV. Определим значение математического ожидания случайной величины, подчиняющейся биномиальному распределению, то есть $\xi \in B_n(p)$. Если $\xi \in B_n(p)$, то множеством возможных значений $\xi$ будут числа $k$=0,1,…,$n$ и $P(\xi = k) = p_k = C_n^k p^k q^{n-k}$. По определению математического ожидания для случайных величин дискретного типа записываем: $M\xi = \sum_{k=0}^{n} k \cdot C_n^k p^k q^{n-k}$. Заметим, что первое слагаемое записанной суммы всегда равно нулю, то есть: $M\xi = \sum_{k=1}^{n} k \cdot C_n^k p^k q^{n-k}$. Значит $M\xi = \sum_{k=1}^{n} \frac{n!}{(k-1)! \cdot (n-k)!} p^k q^{n-k}$. Дальнейшие преобразования не нуждаются в пояснениях:

$$M\xi = np\sum_{k=1}^{n} \frac{(n-1)!}{(k-1)! \cdot ((n-1)-(k-1))!} p^{k-1} q^{(n-1)-(k-1)} = np\sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} q^{(n-1)-(k-1)} = np(q+p)^{n-1}$$

Окончательно получаем: если случайная величина подчиняется биномиальному распределению $\xi \in B_n(p)$, то $M\xi = np$.

Значением математического ожидания распределённой по биномиальному закону случайной величины, которое равно произведению числа проведённых повторно независимых испытаний $n$ и вероятности наступления события в единичном испытании $p$, мы пользуемся в повседневной жизни, исходя из нашего жизненного опыта. Если, например, вероятность выпадения «герба» при одном подбрасывании монеты равна $p = 0{,}5$, то, не утруждая себя подбрасываниями монеты $n = 100$, или $n = 500$ раз, мы заявляем, что *ожидаемое число* выпадений «герба» будет числом мало отличающимся от 50, или 250.

Пусть случайная величина $\xi$ - результат однократного проведения испытания. Такую случайную величину мы называем *бернуллиевской*, и записывать $\xi \in B_1(p)$. Математическое ожидание бернуллиевской случайной величины $\xi$ равно вероятности наступления события в единичном испытании, то есть $M\xi = p$.

Пример 4.1. IV. Определим значение математического ожидания случайной величины, подчиняющейся гипергеометрическому распределению. Такое распределение вероятностей встречается при проведении повторных испытаний, в которых вероятность наступления события зависит от номера испытания. Пусть в урне находятся $n_1$ белых и $n_2$ чёрных шаров и проводится извлечение $m$ шаров без возвращения. Определим случайную величину $\xi$ - число белых шаров в выборке, содержащей $m$ шаров.

Если $k \le min(n_1, n_2)$, то множеством возможных значений случайной величины $\xi$ будет $\Omega = \{0,1,..,k,..,m\}$ и $P(\xi = k) = p_k = \frac{C_{n_1}^k \cdot C_{n_2}^{m-k}}{C_{n_1+n_2}^m}$.

Так как случайная величина $\xi$ - дискретного типа, то её математическое ожидание будет равно: $\boldsymbol{M\xi}=\sum_{k=0}^{m}\boldsymbol{k}\cdot\frac{C_{n_1}^{k}\cdot C_{n_2}^{m-k}}{C_{n_1+n_2}^{m}}$. Учитывая что первое слагаемое всегда равно нулю, запишем:

$$\boldsymbol{M\xi}=\sum_{k=1}^{m}\frac{\frac{n_1(n_1-1)!}{(k-1)!\cdot((n_1-1)-(k-1))!}\cdot\frac{n_2!}{((m-1)-(k-1))!\cdot(n_2-((m-1)-(k-1)))!}}{\frac{n_1+n_2}{m}\cdot\frac{((n_1-1)+n_2)!}{(m-1)!\cdot(((n_1-1)+n_2)-(m-1))!}}.$$

Вынося за знак суммы множители, которые не зависят от индекса суммирования, получим: $\boldsymbol{M\xi}=\frac{n_1\cdot m}{n_1+n_2}\sum_{k=1}^{m}\frac{C_{n_1-1}^{k-1}\cdot C_{n_2}^{(m-1)-(k-1)}}{C_{(n_1-1)+n_2}^{m-1}}$. Полученная сумма равна единице, так как она является суммой вероятностей гипергеометрического распределения с параметрами $n_1-1$, $n_2$ и $m-1$.

Итак, математическое ожидание случайной величины $\xi$, подчиняющейся гипергеометрическому распределению с параметрами $n_1$, $n_2$ и $m$, будет равно $\boldsymbol{M\xi}=\frac{n_1}{n_1+n_2}\cdot m$.

Если случайная величина $\xi$ - число белых шаров в выборке, то её математическое ожидание будет равно *произведению доли белых шаров* $\frac{n_1}{n_1+n_2}$, *находящихся в урне, и количества шаров в выборке* $m$.

Пример 5.1. IV. Определим значение математического ожидания случайной величины, подчиняющейся геометрическому распределению.

Такое распределение вероятностей встречается при проведении независимых испытаний до первого положительного исхода, то есть до первого наступления интересующего нас события. Вероятность наступления этого события в единичном испытании равна $p$, $0<p<1$. Случайная величина $\xi$ - число проведённых испытаний до первого положительного исхода. Областью возможных значений $\xi$ будет множество натуральных чисел $\Omega=N$. Так как $P(\xi=k)=p_k=pq^{k-1}$, то математическое ожидание случайной величины $\xi$ будет равно сумме ряда: $\boldsymbol{M\xi}=\sum_{k=1}^{\infty}k\cdot pq^{k-1}$.

Рассмотрим степенной ряд $S(u)=\sum_{k=1}^{\infty}p\cdot k(qu)^{k-1}$. Этот ряд сходится в области $|u|\le 1$ и $S(1)=\boldsymbol{M\xi}$. Проинтегрируем почленно степенной ряд $S(u)$:

$$\int S(u)du=\sum_{k=1}^{\infty}\left(\int p\cdot k(qu)^{k-1}du\right)=\sum_{k=1}^{\infty}\frac{p}{q}\cdot(qu)^{k}=G(u).$$

Ясно, что: $G'(u)\big|_{u=1}=S(1)=\boldsymbol{M\xi}$. $\qquad (*)$

Получившаяся функция $G(u)$ для $|u|\le 1$ является суммой членов бесконечно убывающей геометрической прогрессии, то есть: $G(u)=p\cdot\frac{u}{1-qu}$.

Вычислим производную этой функции: $G'(u)=\frac{p}{(1-qu)^2}$. Тогда, согласно $(*)$, математическое ожидание случайной величины $\xi$, подчиняющейся геометрическому распределению, будет равно: $M\xi=G'(1)=\frac{p}{(1-q)^2}=\frac{1}{p}$.

***Комментарий***

Рассмотрим испытание: из колоды карт извлекается одна карта, с возвращением её каждый раз обратно, до тех пор, пока не появится карта бубновой масти. Ясно, что здесь случайная величина $\xi$ - *количество сделанных извлечений*. Так как вероятность появления при одном извлечении карты бубновой масти равна $p=\frac{1}{4}$, то можно утверждать, что, при многократном повторении такого испытания, карта бубновой масти в среднем будет появляться при четвёртом извлечении.

**Практика**

<u>Пример 6.1.IV.</u> Проводится серия повторных независимых испытаний, каждом из которых может наступить либо событие $A$, либо событие $\overline{A}$. Если вероятность $P(A)$ наступления события $A$ в каждом испытании – мала, а число проводимых испытаний $n$ – велико, то случайная величина $\xi$ - количество наступлений события $A$ – подчиняется закону распределения вероятностей, называемому законом К.Пуассона (кратко $\xi\in\Pi(\lambda)$) или – «законом редких событий»: $P(\xi=k)=p_k=\frac{\lambda^k\cdot e^{-\lambda}}{k!}$, где $k=0,1,2,\ldots$. Это объясняется тем, что знаменатель дроби - $k!$, с увеличением $k$, принимает такие большие значения, что вероятности $p_k$ становятся пренебрежимо малыми по сравнению со значениями вероятностей для первых значений $\xi$: $k=0,1,2,\ldots$.

Поэтому мы можем записать, что математическое ожидание случайной величины $\xi$ будет равно: $M\xi=\sum_{k=0}^{\infty}k\cdot\frac{\lambda^k\cdot e^{-\lambda}}{k!}$, или $M\xi=\lambda\sum_{k=1}^{\infty}\frac{\lambda^{k-1}\cdot e^{-\lambda}}{(k-1)!}$. Так как набор вероятностей $\{\tilde{p}_k\}=\left\{\frac{\lambda^{k-1}\cdot e^{-\lambda}}{(k-1)!}\right\}$, где $k=2,3,\ldots$., является распределением вероятностей случайной величины $\eta=\xi-1$, то значение суммы равно единице и $M\xi=\lambda\cdot 1=\lambda$.

**Практика**

Определим значения математического ожидания для некоторых законов распределения случайных величин непрерывного типа.

<u>Пример 7.1.IV.</u> Если случайная величина $\xi$ подчиняется экспоненциальному закону распределения, то плотность вероятности

записывается двойным равенством: $\boldsymbol{p}(\boldsymbol{x})=\begin{cases} 0, & \text{если} \quad \boldsymbol{x}\le 0 \\ \dfrac{1}{\boldsymbol{\mu}}\cdot \boldsymbol{e}^{-\frac{x}{\mu}}, & \text{если} \quad \boldsymbol{x}>0 \end{cases}$.

Согласно определению математического ожидания случайной величины непрерывного типа запишем: $\boldsymbol{M\xi}=\int\limits_{-\infty}^{\infty}\boldsymbol{x}\cdot\boldsymbol{p}(\boldsymbol{x})\boldsymbol{dx}$. Плотность вероятности $\boldsymbol{p}(\boldsymbol{x})$ кусочно- непрерывная функция, поэтому $\boldsymbol{M\xi}=\int\limits_{-\infty}^{0}\boldsymbol{x}\cdot 0\boldsymbol{dx}+\dfrac{1}{\boldsymbol{\mu}}\int\limits_{0}^{\infty}\boldsymbol{x}\cdot\boldsymbol{e}^{-\frac{x}{\mu}}\boldsymbol{dx}$. Применяя ко второму слагаемому формулу интегрирования по частям, получаем: $\boldsymbol{M\xi}=\boldsymbol{x}\cdot\boldsymbol{e}^{-\frac{x}{\mu}}\Big|_{0}^{\infty}+\int\limits_{0}^{\infty}\boldsymbol{e}^{-\frac{x}{\mu}}\boldsymbol{dx}=0-\boldsymbol{\mu}\cdot\boldsymbol{e}^{-\frac{x}{\mu}}\Big|_{0}^{\infty}=\boldsymbol{\mu}$.

Если значение математического интеграла мы интерпретируем как среднее значение случайной величины, то создаётся впечатление, что должно быть $\boldsymbol{P}(\xi\le\boldsymbol{M\xi})=\boldsymbol{P}(\xi\ge\boldsymbol{M\xi})=0{,}5$. Однако очевидно, что это равенство будет справедливо только в том случае, если плотность вероятности будет чётной функцией относительно математического ожидания, то есть если будет справедливо: $\boldsymbol{p}(-\boldsymbol{x}+\boldsymbol{M\xi})=\boldsymbol{p}(\boldsymbol{x}+\boldsymbol{M\xi})$.

Время «жизни» электрической лампочки накаливания, случайная величина $\xi$, подчиняется экспоненциальному закону распределения вероятностей. Определим вероятности событий $\{\xi\le\boldsymbol{\mu}\}$ и $\{\xi\ge\boldsymbol{\mu}\}$:

$$\boldsymbol{P}(\xi\le\boldsymbol{\mu})=\boldsymbol{p}_1=\int\limits_{0}^{\mu}\frac{1}{\boldsymbol{\mu}}\boldsymbol{e}^{-\frac{x}{\mu}}\boldsymbol{dx}=1-\frac{1}{\boldsymbol{e}}=\frac{\boldsymbol{e}-1}{\boldsymbol{e}};\ \ \boldsymbol{P}(\xi\ge\boldsymbol{\mu})=\boldsymbol{p}_2=\int\limits_{\mu}^{\infty}\frac{1}{\boldsymbol{\mu}}\boldsymbol{e}^{-\frac{x}{\mu}}\boldsymbol{dx}=\frac{1}{\boldsymbol{e}}=1-\boldsymbol{p}_1.$$

Значения этих вероятностей примерно равны: $\boldsymbol{p}_1\approx 0{,}632$ и $\boldsymbol{p}_2\approx 0{,}368$.

Предположим, что комендант студенческого общежития приобрёл партию из $\boldsymbol{n}$ штук электрических лампочек. Введем случайную величину $\boldsymbol{\eta}$ - число лампочек, время «жизни» которых превысит величину математического ожидания $\boldsymbol{\mu}$. Ясно, что $\boldsymbol{\eta}\in\boldsymbol{B}_{\boldsymbol{n}}(\boldsymbol{p}_2)$. Но мы знаем, что $\boldsymbol{M\eta}=\boldsymbol{np}_2$. Следовательно, через некоторое время комендант общежития может с удивлением обнаружить, что не 50 %, а только 36,8% лампочек «проживут» дольше среднего времени жизни $\boldsymbol{\mu}$, указанного производителем на упаковке лампочек.

<u>Пример 8.1.IV.</u> Пусть случайная величина $\xi$ подчиняется нормальному закону распределения, $\xi\in\mathcal{N}(\boldsymbol{m},\boldsymbol{\sigma})$, то есть $\boldsymbol{p}(\boldsymbol{x})=\dfrac{1}{\sqrt{2\pi}\cdot\boldsymbol{\sigma}}\boldsymbol{e}^{-\frac{(\boldsymbol{x}-\boldsymbol{m})^2}{2\sigma^2}}$ - плотность вероятности случайной величины $\xi$.

Согласно определению математического ожидания запишем:

$$\boldsymbol{M\xi}=\frac{1}{\sqrt{2\pi}\cdot\boldsymbol{\sigma}}\int\limits_{-\infty}^{\infty}\boldsymbol{x}\cdot\boldsymbol{e}^{-\frac{(\boldsymbol{x}-\boldsymbol{m})^2}{2\sigma^2}}\boldsymbol{dx}.$$

Не будем вычислять значение этого интеграла, а рассмотрим интеграл, который равен нулю как интеграл от нечётной функции в симметричных пределах:

$$\frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{\infty}(x-m)\cdot e^{-\frac{(x-m)^2}{2\sigma^2}}dx=0.$$

Интеграл от разности функций равен разности интегралов: $\frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{\infty}x\cdot e^{-\frac{(x-m)^2}{2\sigma^2}}dx-\frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{\infty}m\cdot e^{-\frac{(x-m)^2}{2\sigma^2}}dx=0$. Уменьшаемое в этой разности есть математическое ожидание, а вычитаемое – постоянная $m$, умноженная на интеграл от плотности вероятности, который всегда равен единице, следовательно: $M\xi-m\cdot 1=0$. Значит, числовой параметр $m$ нормального закона распределения вероятностей $\mathcal{N}(m,\sigma)$ значений случайной величины $\xi$ - это её математическое ожидание, то есть $M\xi=m$.

Пример 9.1.IV. Определим случайную величину $\xi$ - величина годового дохода гражданина некоторого государства. Эта случайная величина подчиняется закону Парето, плотность вероятности которого задаётся двойным равенством: $p(x)=\begin{cases}0, & \text{если } x\notin[a_0,\infty)\\ \frac{\alpha}{a_0}\cdot\left(\frac{a_0}{x}\right)^{\alpha+1}, & \text{если } x\in[a_0,\infty)\end{cases}$.

Это распределение вероятностей используются в различных задачах экономической статистики. Например, при изучении распределения величины годового дохода граждан, превышающего некоторый минимальный прожиточный уровень $a_0$.

Определим значение математического ожидания случайной величины $\xi$. Учитывая область возможных значений случайной величины, получаем: $M\xi=\int_{a_0}^{\infty}x\cdot\frac{\alpha}{a_0}\cdot\left(\frac{a_0}{x}\right)^{\alpha+1}dx=\frac{\alpha}{\alpha-1}a_0$. Ясно, что должно быть $\alpha>1$. Точное значение параметра $\alpha$ определяется в каждом конкретном случае отдельно и зависит от конкретной ситуации. Если, например, $\alpha=2$, то $M\xi=2a_0$. Это означает, что в рассматриваемой модели, средняя заработная плата в два раза выше минимального прожиточного уровня. Вероятность случайного события $\{a_0<\xi<M\xi\}$ будет равна: $P(a_0<\xi<M\xi)=\int_{a_0}^{2a_0}\frac{\alpha}{a_0}\cdot\left(\frac{a_0}{x}\right)^{\alpha+1}dx=1-\frac{1}{2^{\alpha}}$. Так как мы назначили значение $\alpha=2$. То результат $P(a_0<\xi<M\xi)=1-\frac{1}{4}=\frac{3}{4}$ мы можем трактовать так: «75 % граждан имеют годовой доход, не превышающий величину среднего годового дохода граждан в этом государстве».

***Комментарий***

Анализ вида плотности распределения Парето и возможный произвол в определении смысла $a_0$ и значений параметров распределения $a_0$ и $\alpha$ приводят нас к выводу, что целенаправленно подбирая значения этих параметров можно обосновать любое субъективное толкование экономической ситуации.

## Практика

Пример 10.1.IV. Если случайная величина $\xi$ подчиняется закону Коши, плотность вероятности которого имеет вид: $p(x)=\dfrac{\vartheta}{\pi\left(\vartheta^2+x^2\right)}$, где $x\in R$. Вычисляя значение математического ожидания, мы приходим к неопределённости вида $[\infty-\infty]$:

$$M\vartheta=\frac{\vartheta}{\pi}\int_{-\infty}^{\infty}\frac{xdx}{\vartheta^2+x^2}=\frac{\vartheta}{2\pi}\int_{-\infty}^{\infty}\frac{d\left(\frac{x}{\vartheta}\right)^2}{1+\left(\frac{x}{\vartheta}\right)^2}=\frac{\vartheta}{2\pi}ln\left(1+\left(\frac{x}{\vartheta}\right)^2\right)\Bigg|_{-\infty}^{\infty}=\infty-\infty.$$

Так как, определяя математическое ожидание, мы требовали абсолютную сходимость интеграла Римана-Стилтьеса, то приходим к выводу, что у распределения Коши математическое ожидание не существует.

### *Комментарий*

Вычисляя значения математического ожидания для некоторых, наиболее часто встречающихся законов распределения вероятностей, мы поняли вероятностный смысл математического ожидания. Ясно, что для вычисления значения математического ожидания, надо обладать техникой суммирования рядов и вычисления определённых интегралов. Очевидно, что и доказательства свойств математического ожидания будут основываться на применении этой техники. Но, прежде чем приступить к изучению свойств математического ожидания случайных величин, рассмотрим задачу, иллюстрирующую важность этого понятия.

Весной 2016 года Центробанк Российской Федерации объявил о своём намерении выпустить в оборот денежные купюры достоинством в 200 и 2000 рублей. Не вникая в суть объяснений Центробанком такого мероприятия, попытаемся с точки зрения теории вероятностей показать его полезность.

У нас сейчас используются купюры достоинством 100, 500 и 1000 рублей. Предположим, что покупатель с равной возможностью может на полках магазина сделать набор покупок суммарной стоимостью 100,200,300, …, 1000 рублей. Оплачивая в кассе свои покупки, он старается обойтись минимально возможным количеством купюр. Введём случайную величину $\xi$ - минимальное количество купюр, предъявленных кассиру для оплаты покупок. Определим $M\xi$ - среднее значение количества купюр, предъявляемых для оплаты покупок.

Случайная величина $\xi$ - случайная величина дискретного типа, а поэтому необходимо составить ряд распределения. Сначала составим таблицу необходимого числа купюр для оплаты покупок.

| Стоимость покупок (руб.) | 100 | 200 | 300 | 400 | 500 | 600 | 700 | 800 | 900 | 1000 |
|---|---|---|---|---|---|---|---|---|---|---|
| Минимальное число купюр для оплаты (штук) | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 5 | 1 |

Так как все стоимости покупок равновозможные, то ряд распределения случайной величины $\xi$ будет иметь вид:

| $\xi$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $p_k$ | $\frac{3}{10}$ | $\frac{2}{10}$ | $\frac{2}{10}$ | $\frac{2}{10}$ | $\frac{1}{10}$ |

Вычислим значение математического ожидания случайной величины $\xi$:

$$M\xi = 1\cdot\frac{3}{10} + 2\cdot\frac{2}{10} + 3\cdot\frac{2}{10} + 4\cdot\frac{2}{10} + 5\cdot\frac{1}{10} = 2{,}6 \text{ штук купюр.}$$

Чтобы оценить будущее нововведение Центрального Банка, введём случайную величину $\eta$ - минимальное количество купюр, предъявленных для оплаты тех же покупок, в предположении, что для оплаты используются купюры достоинством 100,200,500 и 1000 рублей, и определим её математическое ожидание $M\eta$.

Составим таблицу необходимого числа купюр для оплаты покупок.

| Стоимость покупок (руб.) | 100 | 200 | 300 | 400 | 500 | 600 | 700 | 800 | 900 | 1000 |
|---|---|---|---|---|---|---|---|---|---|---|
| Минимальное число купюр для оплаты (штук) | 1 | 1 | 2 | 2 | 3 | 2 | 2 | 3 | 3 | 1 |

Ряд распределения случайной величины $\eta$ будет иметь вид:

| $\eta$ | 1 | 2 | 3 |
|---|---|---|---|
| $p_k$ | $\frac{3}{10}$ | $\frac{4}{10}$ | $\frac{3}{10}$ |

Вычислим значение математического ожидания случайной величины $\eta$:

$$M\eta = 1\cdot\frac{3}{10} + 2\cdot\frac{4}{10} + 3\cdot\frac{3}{10} = 2{,}0 \text{ штук купюр.}$$

Сравнивая значения $M\xi$ и $M\eta$, приходим к выводу, что после введения в оборот новых купюр номиналом 200 и 2000 рублей объём «купюр-бумажек» находящихся в обороте снизится на 23%. Выгода и для Центрального банка, и для населения, пользующегося новым набором «купюр-бумажек», - очевидна.

Кстати заметим, что номиналы долларов США: 10,20,50 и 100. А чтобы американцы не думали, что мы за образец взяли их типы номиналов, мы могли бы ввести в оборот купюры номиналами 300 и 3000 рублей, а не 200 и 2000 руб. Легко проверить, выгода будет такой же - 23%.

## §3.IV Свойства математического ожидания
### Теория

Приступая к изучению свойств математического ожидания случайной величины, напомним, что по определению математическое ожидание это абсолютно сходящийся определённый интеграл Римана-Стилтьеса. А поэтому почти все свойства математического ожидания следуют из свойств определённых интегралов. Будем применять теорему Фубини о возможности сведения кратного интеграла к повторному интегралу, которая в терминах теории вероятностей, может быть сформулирована так:

Пусть $\xi=(\xi_1,\xi_2)$ двумерная случайная величина, $F(x_1,x_2)$ - её функция распределения, $y=g(x_1,x_2)$ - непрерывная функция и $\iint\limits_{R^2}|g(x_1,x_2)|dF(x_1,x_2)<\infty$, тогда интегралы $\int\limits_{R_1}g(x_1,x_2)dF_1(x_1)$ и $\int\limits_{R_2}g(x_1,x_2)dF_2(x_2)$ существуют и

$$\iint\limits_{R^2}|g(x_1,x_2)|dF(x_1,x_2)=\int\limits_{R_1}\left[\int\limits_{R_2}g(x_1,x_2)dF_2(x_2)\right]dF_1(x_1)=\int\limits_{R_2}\left[\int\limits_{R_1}g(x_1,x_2)dF_1(x_1)\right]dF_2(x_2).$$

<u>Свойство 1.</u> Математическое ожидание - линейный функционал.

Если $\xi=(\xi_1,\xi_2)$ - двумерная случайная величина, $y=g(x_1,x_2)=ax_1+bx_2$ - линейная функция и $\iint\limits_{R^2}|g(x_1,x_2)|dF(x_1,x_2)<\infty$, то $M[a\xi_1+b\xi_2]=aM\xi_1+bM\xi_2$ - *математическое ожидание линейной комбинации случайных величин равно линейной комбинации их математических ожиданий*.

Доказательство. Согласно общему определению математического ожидания и линейным свойствам интеграла запишем :

$$M[a\xi_1+b\xi_2]=\iint\limits_{R^2}(ax_1+bx_2)dF(x_1,x_2)=a\cdot\iint\limits_{R^2}x_1dF(x_1,x_2)+b\cdot\iint\limits_{R^2}x_2dF(x_1,x_2).$$

Применяя теорему Фубини, перейдём к повторным интегралам:

$$M[a\xi_1+b\xi_2]=a\cdot\int\limits_{R_1}x_1\left[\int\limits_{R_2}dF_2(x_2)\right]dF_1(x_1)+b\cdot\int\limits_{R_2}x_2\left[\int\limits_{R_1}dF_1(x_1)\right]dF_2(x_2).$$

Так как $\int\limits_{R_2}dF_2(x_2)=1$ и $\int\limits_{R_1}dF_1(x_1)=1$, то:

$$M[a\xi_1+b\xi_2]=a\cdot\int\limits_{R_1}x_1dF_1(x_1)+b\cdot\int\limits_{R_2}x_2dF_2(x_2)=aM\xi_1+bM\xi_2,$$

что и требовалось доказать.

Из свойства 1 следует:

*а)* $M[a]=a$. *Математическое ожидание постоянной величины равно этой постоянной* .Здесь аргумент математического ожидания $\xi=a$ - постоянная «случайная величина», то есть $P(\xi=a)=1$. Такую случайную величину будем называть «*выроjденной в точке* $a$».

*б)* $M[a\xi]=aM\xi$. *Постоянная величина выносится за знак математического ожидания.*

*в)* $M[\xi_1+\xi_2]=M\xi_1+M\xi_2$. *Математическое ожидание суммы случайных* *величин равно сумме их математических ожиданий.*

Свойство 2. Математическое ожидание произведения случайных величин равно произведению математических ожиданий случайных величин, если случайные величины независимые: $M[\xi_1 \cdot \xi_2] = M\xi_1 \cdot M\xi_2$.

Доказательство. Если $\xi = (\xi_1, \xi_2)$ - двумерная случайная величина и её компоненты независимые, то $F(x_1, x_2) = F_1(x_1) \cdot F_2(x_2)$. Пусть $y = g(x_1, x_2) = x_1 \cdot x_2$, тогда по общему определению математического ожидания: $M[\xi_1 \cdot \xi_2] = \iint\limits_{R^2} x_1 \cdot x_2 dF(x_1, x_2) = \iint\limits_{R^2} x_1 \cdot x_2 dF_1(x_1) \cdot dF_2(x_2)$. Воспользуемся теоремой Фубини и перейдём к повторному интегралу. Применяя определение математического ожидания и свойство 1*б)*, завершим доказательство:

$$M[\xi_1 \cdot \xi_2] = \int\limits_{R_1} x_1 \left[ \int\limits_{R_2} x_2 dF_2(x_2) \right] dF_1(x_1) = \int\limits_{R_1} x_1 M\xi_2 dF_1(x_1) = M\xi_2 \int\limits_{R_1} x_1 dF_1(x_1) = M\xi_1 \cdot M\xi_2$$

Свойство 3. Если случайная величина $\xi$ принимает неотрицательные значения $(\xi \geq 0)$ и её математическое ожидание равно нулю $(M\xi = 0)$, то случайная величина $\xi$ принимает значение ноль с вероятностью равной единице $(P(\xi = 0) = 1)$.

Доказательство. Так как $\xi \geq 0$, то для $x \leq 0$ функция распределения $F(x)$ случайной величины $\xi$ принимает значение ноль: $F(x) = 0$.

Так как $\xi$ принимает неотрицательные значения, то $M\xi = \int\limits_0^\infty x dF(x)$. По условию $M\xi = 0$, следовательно, $dF(x) = 0$. То есть, для всех $x > 0$ функция распределения $F(x)$ принимает постоянное значение: $F(x) = const$. Из свойств функции распределения мы знаем, что $F(\infty) = 1$, значит $const = 1$. Итак, получаем:

$$F(x) = \begin{cases} 0, & \text{если} \quad x \leq 0 \\ 1, & \text{если} \quad x > 0 \end{cases}.$$

Следовательно, случайная величина $\xi$ принимает значение ноль с вероятностью равной единице, то есть $P(\xi = 0) = 1$. То есть: $\xi$ - вырожденная в точке ноль случайная величина.

Свойство 4. Неравенство Коши-Буняковского. Для любых случайных величин $\xi$ и $\eta$ будет справедливо: $M^2[\xi \cdot \eta] \leq M[\xi^2] \cdot M[\eta^2]$.

Доказательство. Случайная величина $\zeta = (a\xi + b\eta)^2$ при любых $a$ и $b$ принимает неотрицательные значения. Так как $\zeta \geq 0$, то $M\zeta \geq 0$ - это легко проверить по определению математического ожидания. Применяя первое свойство математического ожидания, получаем $M\zeta = a^2 M[\xi^2] + 2abM[\xi \cdot \eta] + b^2 M[\eta^2]$. Полученная квадратичная форма принимает неотрицательные значения, следовательно, её дискриминант будет неотрицательным числом, то есть: $M^2[\xi \cdot \eta] - M[\xi^2] \cdot M[\eta^2] \leq 0$. Откуда следует требуемое неравенство.

Следующие свойства математического ожидания легко доказываются.

<u>Свойство 5.</u> Если $\xi \geq \eta$, то $M\xi \geq M\eta$.

<u>Свойство 6.</u> $|M\xi| \leq M|\xi|$.

***Комментарий***

Математическое ожидание $M\xi$ случайной величины $\xi$ мы понимаем как среднее значение случайной величины, и интерпретируем как координату центра тяжести единичной массы распределённой на числовой оси по закону, описываемому функцией распределения $F(x)$.

Рассмотрим случайную величину $\eta = \xi - m$, где $m$ – значение $M\xi$, то есть $m = const$. Ясно, значениями случайной величины $\eta$ будут величины отклонений значений случайной величины $\xi$ от значения $M\xi$. Для любой случайной величины $\xi$ математическое ожидание случайной величины $\eta$ будет равно: $M\eta = M[\xi - m] = M\xi - Mm = m - m = 0$. То есть среднее значение отклонений значений $\xi$ от значения её математического ожидания $M\xi = m$ всегда равно нулю. Случайную величину $\eta = \xi - m$ будем называть *центрированной* случайной величиной.

Но, в то же время мы понимаем, что, если $\xi_1$ и $\xi_2$ - две случайные подчиняющиеся одному закону распределения вероятностей и их математические ожидания равны, картины разбросов значений этих случайных величин около математического ожидания могут довольно значительно различаться. Мы пришли к выводу, что помимо математического ожидания надо определить числовую характеристику, которая будет мерой разброса значений случайной величины около её математического ожидания.

## §4.IV Дисперсия случайной величины

### Теория

Пусть $\xi$ - случайная величина, математическое ожидание $M\xi$ которой равно $m$, и непрерывная функция $y = g(x) = (x - m)^2$.

<u>Определение 3.</u>

***Дисперсией*** *случайной величины* $\xi$ *называется математическое ожидание квадрата отклонения случайной величины от её математического ожидания:*

$$D\xi = M[\xi - M\xi]^2.$$

То есть, дисперсия это значение интеграла Римана-Стилтьеса от функции $g(x) = (x - m)^2$ по вероятностной мере $P$:

$$D[\xi] = M[\xi - m]^2 = \int_{R_1} (x - m)^2 P(dx) = \int_{R_1} (x - m)^2 dF(x),$$

если этот интеграл сходится абсолютно.

Так как мы изучаем случайные величины двух типов, то, для вычислений значений дисперсии, соответственно в дискретном и непрерывном случае, запишем вид интеграла Римана-Стилтьеса от функции $g(x) = (x - m)^2$:

$$D[\xi]=\int\limits_{R_1}(x-m)^2dF(x)=\begin{cases}\sum\limits_{k=1}^{n,\infty}(x_k-m)^2\cdot p_k,\\ \int\limits_{-\infty}^{\infty}(x-m)^2\cdot p(x)dx,\end{cases}$$ если, соответственно, ряд или интеграл сходятся абсолютно.

Определив математическое ожидание $M\xi$ случайной величины $\xi$, мы, чтобы выяснить смысл этой числовой характеристики и приёмы вычисления её значения, сначала рассмотрели примеры определения $M\xi$ для наиболее часто встречающихся законов распределения вероятностей случайных величин, а потом рассмотрели свойства математического ожидания $M\xi$.

Дисперсия случайной величины $D\xi$ – это *мера разброса* значений $\xi$ около её математического ожидания $M\xi$, а так как мы определили дисперсию как математическое ожидание функции $g(\xi)$, то её свойства вытекают из свойств математического ожидания.

<u>Свойство 1.</u> У любой случайной величины $\xi$ дисперсия $D\xi\ge 0$ и $D\xi=0$ тогда и только тогда, когда $\xi$ - вырожденная, то есть $\xi=const$ с вероятностью равной единице.

Первое утверждение, о том, что значениями дисперсии могут быть только неотрицательные числа очевидно из определения дисперсии.

Пусть дано: $D\xi=0$. Докажем, что $\xi$ - вырожденная, то есть $\xi=c$. Определим случайную величину $\eta$ - функцию случайной величины $\xi$: $\eta=(\xi-m)^2$. Ясно, что случайная величина $\eta$ принимает неотрицательные значения: $\eta=(\xi-m)^2\ge 0$. Но $M\eta=M[\xi-M\xi]^2$, а по условию $M[\xi-M\xi]^2=D\xi=0$, то есть $M\eta=0$. Так как значения $\eta\ge 0$ и $M\eta=0$, то по третьему свойству математического ожидания: $P(\eta=0)=1$. Значит, с вероятностью равной единице мы можем утверждать, что $(\xi-m)^2=0$. Следовательно, $P(\xi=m)=1$, и $\xi$ - вырожденная в точке $m$ случайная величина.

Обратно, пусть дано: $\xi$ - вырожденная, то есть $\xi=const$. Доказать, что $D\xi=0$.

По первому свойству математического ожидания, если $\xi=const=c$, то $Mc=c$. Следовательно, $Dc=M[c-c]^2=0$.

<u>Свойство 2.</u> $D[c\xi]=c^2D\xi$.

*Постоянная за знак дисперсии выносится с квадратом.*

Для доказательства этого свойства достаточно вычислить дисперсию случайной величины $\eta=c\xi$:

$$D\eta=D[c\xi]=M[c\xi-M[c\xi]]^2=M[c(\xi-M\xi)]^2=c^2M[\xi-M\xi]^2=c^2D\xi.$$

<u>Свойство 3.</u> Если случайные величины $\xi_1$ и $\xi_2$ - независимые, то дисперсия суммы случайных величин равна сумме дисперсий этих случайных величин: $D[\xi_1+\xi_2]=D\xi_1+D\xi_2$.

Вычислим дисперсию случайной величины $\eta = \xi_1 + \xi_2$, используя определение дисперсии:

$$D\eta = D[\xi_1 + \xi_2] = M[(\xi_1 + \xi_2) - M(\xi_1 + \xi_2)]^2 = M[(\xi_1 - M\xi_1) - (\xi_2 - M\xi_2)]^2$$

Выражение в квадратных скобках возведём в квадрат и воспользуемся первым свойством математического ожидания:

$$D[\xi_1 + \xi_2] = M\left[(\xi_1 - M\xi_1)^2 - 2(\xi_1 - M\xi_1)\cdot(\xi_2 - M\xi_2) + (\xi_2 - M\xi_2)^2\right] =$$
$$= M[\xi_1 - M\xi_1]^2 - 2M[(\xi_1 - M\xi_1)\cdot(\xi_2 - M\xi_2)] + M[\xi_2 - M\xi_2]^2$$

Если случайные величины $\xi_1$ и $\xi_2$ независимые, то и центрированные случайные величины $\xi_1 - M\xi_1$ и $\xi_2 - M\xi_2$ будут независимыми. Применяя второе свойство математического ожидания к среднему члену последнего выражения, получим:

$$M[(\xi_1 - M\xi_1)\cdot(\xi_2 - M\xi_2)] = M[\xi_1 - M\xi_1]\cdot M[\xi_2 - M\xi_2] = 0\cdot 0 = 0.$$

Следовательно: $D[\xi_1 + \xi_2] = M[\xi_1 - M\xi_1]^2 + M[\xi_2 - M\xi_2]^2 = D\xi_1 + D\xi_2$, что и требовалось доказать.

Следующие два следствия из свойства 3 – очевидны.

Следствие 1. $D[\xi + c] = D\xi$, где $c = const$.

Следствие 2. $D[\xi_1 - \xi_2] = D\xi_1 + D\xi_2$, если случайные величины $\xi_1$ и $\xi_2$ независимые.

Размерность математического ожидания $M\xi$ совпадает с размерностью случайной величины $\xi$. Размерность дисперсии $D\xi$, в силу её определения, равна квадрату размерности случайной величины $\xi$, что не совсем удобно при решении практических задач. Поэтому в практических задачах используют в качестве меры разброса величину $\sigma = \sqrt{D\xi}$, которая называется ***средним квадратическим отклонением*** случайной величины $\xi$.

Если случайная величина $\xi$ имеет дисперсию $D\xi$, то у случайной величины $\eta = \dfrac{\xi}{\sigma}$ дисперсия будет равна $D\eta = D\left[\dfrac{\xi}{\sigma}\right] = \dfrac{1}{\sigma^2}\cdot D\xi = 1$. Случайную величину $\eta = \dfrac{\xi}{\sigma}$ называют *нормированной* случайной величиной.

Если у случайной величины $\xi$ математическое ожидание равно $M\xi = m$, а дисперсия равна $D\xi = \sigma^2$, то функция $\overset{0}{\xi} = \dfrac{\xi - m}{\sigma}$ случайной величины $\xi$ имеет математическое ожидание равное нулю и дисперсию равную единице. Центрированная и нормированная случайная величина $\overset{0}{\xi}$ имеет функцию распределения $\overset{0}{F}(x) = F(\sigma\cdot x + m)$, где $F(x)$ - функция распределения случайной величины $\xi$.

По определению $D\xi = M[\xi - M\xi]^2$. Возведём в квадрат выражение в квадратных скобках и воспользуемся свойствами математического ожидания:

$$D\xi = M[\xi - M\xi]^2 = M\left[\xi^2 - 2\xi M\xi + M^2\xi\right] = M\xi^2 - 2M\xi \cdot M\xi + M^2\xi = M\xi^2 - M^2\xi$$

Получаем ***формулу***: $D\xi = M\xi^2 - M^2\xi$, удобную для вычисления значения дисперсии. Здесь $M\xi^2 = \begin{cases} \sum\limits_{k=1}^{n,\infty} x_k^2 \cdot p_k, & \text{если } \xi \text{ дискретного типа} \\ \int\limits_R x^2 \cdot p(x)dx, & \text{если } \xi \text{ непрерывного типа} \end{cases}$.

Ясно, что если ряд или интеграл $M\xi^2$ сходятся абсолютно, то у случайной величины $\xi$ существуют и математическое ожидание, и дисперсия. Но если ряд или интеграл $M\xi$ сходятся абсолютно, то у случайной величины $\xi$ существует математическое ожидание, а дисперсия может не существовать.

## §5.IV Примеры дисперсий случайных величин

## Практика

Ясно, что как и в примерах вычисления значений математического ожидания случайных величин, распределённых по наиболее часто встречающимся на практике законам распределения вероятностей, так и при вычислении значений дисперсии для этих же законов необходимо обладать техникой суммирования рядов и вычисления определённых интегралов.

Рассмотрим несколько примеров.

Пример 1.5.IV Случайная величина $\xi$ дискретного типа, арифметическая, подчиняется равномерному распределению вероятностей. Вычислим значение дисперсии, применяя формулу $D\xi = M\xi^2 - M^2\xi$.

В примере 1.1. IV было вычислено значение математического ожидания $M\xi = \dfrac{x_1 + x_n}{2}$. Так как $x_n = x_1 + d\cdot(n-1)$, то $M^2\xi = x_1^2 + x_1 d(n-1) + d^2 \cdot \dfrac{(n-1)^2}{4}$.

Вычисляя $M\xi^2 = \sum\limits_{k=1}^{n} x_k^2 \cdot p_k = \dfrac{1}{n}\sum\limits_{k=1}^{n}(x_1 + d\cdot(k-1))^2$, будем использовать формулу: $\sum\limits_{k=1}^{n} k^2 = \dfrac{n(n+1)(2n+1)}{6}$. Получаем $M\xi^2 = x_1^2 + x_1 d(n-1) + d^2 \cdot \dfrac{(n-1)(2n-1)}{6}$.

Значит $D\xi = M\xi^2 - M^2\xi = \dfrac{d^2\cdot(n^2-1)}{12}$.

Пример 2.5.IV. Случайная величина $\xi$ непрерывного типа подчиняется равномерному распределению вероятностей на отрезке $[a,b]$. Вычислим значение дисперсии, применяя формулу $D\xi = M\xi^2 - M^2\xi$.

В примере 2.1. IV было вычислено значение математического ожидания: $M\xi = \dfrac{a+b}{2}$. Значит $M^2\xi = \dfrac{(a+b)^2}{4}$. Вычислим значение $M\xi^2$:

$$M\xi^2 = \int\limits_{-\infty}^{a} x^2 \cdot 0dx + \int\limits_{a}^{b} x^2 \cdot \frac{1}{b-a}dx + \int\limits_{b}^{\infty} x^2 \cdot 0dx = 0 + \frac{b^3 - a^3}{3\cdot(b-a)} + 0 = \frac{a^2 + ab + b^2}{3}.$$

Окончательно получаем: $\boldsymbol{D\xi} = \dfrac{\boldsymbol{a}^2 + \boldsymbol{ab} + \boldsymbol{b}^2}{3} - \dfrac{(\boldsymbol{a}+\boldsymbol{b})^2}{4} = \dfrac{(\boldsymbol{a}-\boldsymbol{b})^2}{12}$. Значит среднее квадратическое отклонение случайной величины $\boldsymbol{\xi}$, которая подчиняется равномерному распределению, будет равно $\boldsymbol{\sigma} = \dfrac{\boldsymbol{b}-\boldsymbol{a}}{2\sqrt{3}}$.

Пример 3.4.IV. Определим значение дисперсии случайной величины, подчиняющейся биномиальному распределению, то есть: $\boldsymbol{\xi} \in \boldsymbol{B_n}(\boldsymbol{p})$.

Математическое ожидание этой случайной величины $\boldsymbol{\xi}$ равно $\boldsymbol{M\xi} = \boldsymbol{np}$. Для определения значения $\boldsymbol{D\xi}$ воспользуемся тем, что $\boldsymbol{\xi} = \sum\limits_{\boldsymbol{k}=1}^{\boldsymbol{n}} \boldsymbol{\xi_k}$, где $\boldsymbol{\xi_k} \in \boldsymbol{B}_1(\boldsymbol{p})$ - бернулиевские и независимые случайные величины.

$\boldsymbol{M\xi_k^2} = 0^2 \cdot \boldsymbol{q} + 1^2 \cdot \boldsymbol{p} = \boldsymbol{p}$ и тогда $\boldsymbol{D\xi_k^2} = \boldsymbol{p} - \boldsymbol{p}^2 = \boldsymbol{pq}$.

Следовательно: $\boldsymbol{D\xi} = \sum\limits_{\boldsymbol{k}=1}^{\boldsymbol{n}} \boldsymbol{D\xi_k} = \boldsymbol{npq}$ и $\boldsymbol{\sigma} = \sqrt{\boldsymbol{npq}}$ .

Случайную величину $\boldsymbol{\tau_n} = \dfrac{\boldsymbol{\xi} - \boldsymbol{np}}{\sqrt{\boldsymbol{npq}}}$ будем называть *центрированной и нормированной суммой* бернуллиевских случайных величин.

***Комментарий***

Мы познакомились с методикой и техникой вычисления значений математического ожидания и дисперсии случайных величин. Значения этих числовых характеристик дают нам упрощённую, но важнейшую для практических задач информацию о распределении значений рассматриваемой случайной величины. Следующие примеры, где рассматриваются наиболее часто встречающиеся законы распределения вероятностей, имеют своей целью, как ранее это было сделано для математических ожиданий, выяснить значения дисперсий, для запоминания и использования их в практических задачах без предварительных вычислений.

**Практика**

Пример 4.5.IV. Мы определили значение математического ожидания $\boldsymbol{M\xi} = \sum\limits_{\boldsymbol{k}=0}^{\boldsymbol{m}} \boldsymbol{k} \cdot \dfrac{\boldsymbol{C_{n_1}^k} \cdot \boldsymbol{C_{n_2}^{m-k}}}{\boldsymbol{C_{n_1+n_2}^m}}$ случайной величины $\boldsymbol{\xi}$, подчиняющейся гипергеометрическому распределению, используя формулу $\boldsymbol{C_n^m} = \dfrac{\boldsymbol{n!}}{\boldsymbol{m!} \cdot (\boldsymbol{n}-\boldsymbol{m})\boldsymbol{!}}$.

Проведя аналогичные преобразования в выражении $\boldsymbol{M\xi}^2 = \sum\limits_{\boldsymbol{k}=0}^{\boldsymbol{m}} \boldsymbol{k}^2 \cdot \dfrac{\boldsymbol{C_{n_1}^k} \cdot \boldsymbol{C_{n_2}^{m-k}}}{\boldsymbol{C_{n_1+n_2}^m}}$ и воспользовавшись формулой $\boldsymbol{D\xi} = \boldsymbol{M\xi}^2 - \boldsymbol{M}^2\boldsymbol{\xi}$, получим значение дисперсии гипергеометрического распределения: $\boldsymbol{D\xi} = \dfrac{\boldsymbol{n}_1 + \boldsymbol{n}_2 - \boldsymbol{m}}{\boldsymbol{n}_1 + \boldsymbol{n}_2 - 1} \cdot \dfrac{\boldsymbol{n}_1\boldsymbol{n}_2}{(\boldsymbol{n}_1 + \boldsymbol{n}_2)^2} \cdot \boldsymbol{m}$.

Пример 5.5.IV. Математическое ожидание случайной величины $\boldsymbol{\xi}$, подчиняющейся геометрическому распределению вероятностей, было определено как значение производной функции $G(u)=p\cdot\dfrac{u}{1-qu}$ при $u=1$, то есть мы получили: $M\xi=G'(1)=\dfrac{1}{p}$.

Функцию $G(u)$ будем использовать и при определении $M\xi^2=\sum\limits_{k=1}^{\infty}k^2\cdot pq^{k-1}$.

Последнюю сумму представим в виде двух сумм-слагаемых: $M\xi^2=\sum\limits_{k=1}^{\infty}k^2\cdot pq^{k-1}=pq\sum\limits_{k=1}^{\infty}k(k-1)\cdot q^{k-2}+\sum\limits_{k=1}^{\infty}k\cdot pq^{k-1}$. Легко проверить, что первая сумма является значением второй производной функции $G(u)=p\cdot\dfrac{u}{1-qu}$ при $u=1$, то есть $pq\sum\limits_{k=1}^{\infty}k(k-1)\cdot q^{k-2}=G''(1)=\dfrac{2qp}{(1-q)^3}=\dfrac{2q}{p^2}$. Вторая сумма – это математическое ожидание случайной величины $\xi$, то есть $\sum\limits_{k=1}^{\infty}k\cdot pq^2=\dfrac{1}{p}$. Итак, дисперсия $D\xi$ случайной величины $\xi$ будет равна:

$$D\xi=M\xi^2-M^2\xi=\frac{2q}{p^2}+\frac{1}{p}-\left(\frac{1}{p}\right)^2=\frac{q}{p^2}.$$

<u>Пример 6.5.IV.</u> Определим значение дисперсии случайной величины $\xi$, подчиняющейся закону Пуассона. Сначала, как и в предыдущих примерах, определим значение суммы $M\xi^2=\sum\limits_{k=0}^{\infty}k^2\cdot\dfrac{\lambda^k e^{-\lambda}}{k!}pq^{k-1}$. Следующие преобразования не нуждаются в пояснениях:

$$M\xi^2=\sum_{k=0}^{\infty}k^2\cdot\frac{\lambda^k e^{-\lambda}}{k!}pq^{k-1}=\sum_{k=1}^{\infty}k\cdot\frac{\lambda^k e^{-\lambda}}{(k-1)!}=\sum_{k=1}^{\infty}(k-1)\cdot\frac{\lambda^k e^{-\lambda}}{(k-1)!}+\lambda\sum_{k=1}^{\infty}\frac{\lambda^{k-1}e^{-\lambda}}{(k-1)!}=$$
$$=\lambda^2\sum_{k=2}^{\infty}\frac{\lambda^{k-2}e^{-\lambda}}{(k-2)!}+\lambda\cdot 1=\lambda^2\cdot 1+\lambda=\lambda^2+\lambda.$$

Следовательно, дисперсия $D\xi$ случайной величины $\xi$ будет равна:

$$D\xi=M\xi^2-M^2\xi=\lambda^2+\lambda-(\lambda)^2=\lambda.$$

<u>Пример 7.5.IV.</u> Определим значение дисперсии случайной величины $\xi$, подчиняющейся экспоненциальному закону. Сначала вычислим значение интеграла $M\xi^2=\dfrac{1}{\mu}\int\limits_0^{\infty}x^2e^{-\frac{x}{\mu}}dx$. Сделав замену переменной интегрирования $\dfrac{x}{\mu}=t$, получаем $M\xi^2=\mu^2\int\limits_0^{\infty}t^2e^{-t}dt=\mu^2\cdot\Gamma(3)=\mu^2\cdot 2!=2\mu^2$. Так как $M\xi=\mu$, то $D\xi=M\xi^2-M^2\xi=2\mu^2-(\mu)^2=\mu^2$.

<u>Пример 8.5.IV.</u> Определим значение дисперсии случайной величины $\xi$,

подчиняющейся нормальному закону с параметрами $m$ и $\sigma$, то есть $\xi \in \mathcal{N}(m,\sigma)$. Ранее мы вычислили: $M\xi = m$. Значения дисперсии случайной величины $\xi$ и центрированной случайной величины $\overset{0}{\xi} = \xi - m$ - совпадают, то есть $D\xi = D[\xi - m] = M\overset{0}{\xi} = \frac{1}{\sqrt{2\pi}\sigma}\int\limits_{-\infty}^{\infty} x^2 \cdot e^{-\frac{x^2}{2\sigma^2}}dx$. После замены переменной интегрирования $\frac{x^2}{2\sigma^2} = t$, используя $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, получаем ответ, который объясняет смысл параметра $\sigma$, используемой при записи плотности вероятности:

$D\xi = \frac{2}{\sqrt{\pi}}\sigma^2\int\limits_0^{\infty} t^{\frac{1}{2}} \cdot e^{-t}dt = \frac{2\sigma^2}{\sqrt{\pi}}\Gamma\left(\frac{3}{2}\right) = \frac{\sigma^2}{\sqrt{\pi}}\Gamma\left(\frac{1}{2}\right) = \sigma^2$. Значит $\sigma = \sqrt{D\xi}$ - среднее квадратическое отклонение.

Пример 9.5.IV. Рассматривая распределение Парето, мы вычислили значение математического ожидания $M\xi = \frac{\alpha}{\alpha - 1}a_0$. Легко вычислить, что $M\xi^2 = \frac{\alpha \cdot a_0^2}{\alpha - 2}$. Тогда $D\xi = \frac{\alpha \cdot a_0^2}{\alpha - 2} - \left(\frac{\alpha}{\alpha - 1}a_0\right)^2 = \frac{\alpha \cdot a_0^2}{(\alpha - 2)(\alpha - 1)^2}$.

Значения параметров $\alpha$ и $a_0$ зависят от конкретной ситуации, в которой применяется это распределение, они могут быть назначены исследователем, желающим проанализировать различные экономические модели.

Но мы, определив значение $M\xi = \frac{\alpha}{\alpha - 1}a_0$, установили, что параметр $\alpha$ должен быть больше единицы. Так как значение дисперсии случайной величины $\xi$ не может быть отрицательным числом, то, при $1 < \alpha \le 2$, у случайной величины $\xi$ существует математическое ожидание, но не существует дисперсия. Если $\alpha > 2$, то у случайной величины $\xi$ существуют и математическое ожидание, и дисперсия.

Пример 10.5.IV. Так как у случайной величины $\xi$, подчиняющейся закону Коши, не существует математическое ожидание, то о существовании дисперсии даже ставить вопрос не стоит.

## §6.IV Начальные и центральные моменты случайной величины
## Теория

Продолжим знакомство с числовыми характеристиками случайной величины $\xi$, которые определяются как математические ожидания степенных функций $\eta = g(\xi)$.

Определение 1. *Начальным моментом* ***k****-того порядка* $\alpha_k$ случайной величины $\xi$ называется математическое ожидание функции $g(\xi) = \xi^k$:

$$\alpha_k = M\left[\xi^k\right] = \int\limits_R x^k \cdot P(dx) = \int\limits_R x^k \cdot dF(x),$$

если интеграл Римана–Стилтьеса сходится абсолютно.
С некоторыми начальными моментами мы же знакомы:

$\alpha_0 = 1$ - у любой случайной величины;

$\alpha_1 = M\xi = m$ - математическое ожидание, или среднее значение случайной величины $\xi$;

$\alpha_2 = M\xi^2$- начальный момент второго порядка, который мы использовали при вычислении дисперсии случайной величины.

Ясно, что если у случайной величины $\xi$ существует начальный момент $\alpha_s$ порядка $s$, то существуют все начальные моменты $\alpha_i$ порядка $i = 1,2,...,s-1$.

Определение 2. *Центральным моментом* $k$*-того порядка* $\mu_k$ называется математическое ожидание функции $g(\xi) = (\xi - M\xi)^k$:

$$\mu_k = M\left[(\xi - M\xi)^k\right] = \int\limits_R (x-m)^k \cdot P(dx) = \int\limits_R (x-m)^k \cdot dF(x),$$

если интеграл Римана–Стилтьеса сходится абсолютно.
Ясно, что:

$\mu_0 = 1$ - у любой случайной величины;

$\mu_1 = 0$ - у любой случайной величины;

$\mu_2 = D\xi = \sigma^2$ - дисперсия случайной величины $\xi$.

Справедлива лемма:

*Центральный момент* $k$*-того порядка* $\mu_k$*, если он существует, может быть представлен в виде линейной комбинации начальных моментов* $\alpha_i$ *порядка* $i$ *не выше, чем* $k$*, то есть* $i = 0,1,...,k$.

Чтобы доказать это утверждение надо с помощью формулы бинома Ньютона представить функцию $g(\xi) = (\xi - M\xi)^k$ в виде:

$$(\xi - M\xi)^k = (\xi - \alpha_1)^k = \sum_{i=0}^{k} (-1)^i C_k^i \xi^{k-i} \alpha_1^i$$

и воспользоваться первым свойством математического ожидания случайной величины: математическое ожидание – линейный функционал.
Из леммы следует, что существование конечного центрального момента $\mu_s$ порядка $s$ зависит от существования начального момента $\alpha_s$ порядка $s$.

В частности, если $k$=2, получаем знакомую формулу, которую мы использовали при вычислении примеров дисперсии случайных величин $\xi$:

$$\mu_2 = D\xi = \sigma^2 = \alpha_2 - \alpha_1^2 = \alpha_2 - m^2.$$

Некоторые числовые характеристики, дающие некоторую упрощённую информацию о случайной величине, используются для оценки отличия закона распределения случайной величины $\xi$ от нормального распределения.

Числовая характеристика $\gamma_1 = \dfrac{\mu_3}{\sigma^3}$, называемая *коэффициентом асимметрии*, оценивает величину асимметричности закона распределения случайной величины $\xi$. Закон распределения вероятностей будет симметричен относительно $M\xi = m$, если $F(x+m) = 1 - F(-x+m)$. У симметричного закона коэффициент асимметрии $\gamma_1 = 0$.

Числовая характеристика $\gamma_2 = \dfrac{\mu_4}{\sigma^4} - 3$ называется *коэффициентом эксцесса.* С помощью коэффициента эксцесса сравнивают «крутизну склонов» холмообразной кривой, являющейся графиком плотности вероятности случайной величины, с «крутизной склонов» графика плотности вероятности нормального закона, у которого $\gamma_2 = 0$.

## §7.IV Числовые характеристики многомерной случайной величины

### *Комментарий*

Математическое ожидание одномерной случайной величины мы интерпретировали как координату центра тяжести единичной массы распределённой на числовой оси. Распределение вероятностей многомерной случайной величины мы интерпретируем как распределение единичной массы на $n$-мерном евклидовом пространстве $R^n = R_1 \times R_2 \times ... \times R_n$. Следовательно, по аналогии, логично рассмотреть вопрос и о точке в пространстве $R^n$, которая будет центром тяжести этой единичной массы.

### Теория

Рассматривается вероятностное пространство $<R^n, \mathcal{B}(R^n), P>$, где $P$ – вероятностная функция, определяющая распределение вероятностей $n$-мерной случайной величины $\xi$.

Пусть $y = g(x)$ - непрерывная функция $n$ переменных и $g(\xi) = \xi_i$, $i = 1,2,...,n$, здесь $\xi_i$ - $i$-тая компонента $n$-мерной случайной величины $\xi$. Ясно, что значение интеграла $\int\limits_{R_1} ... \int\limits_{R_i} ... \int\limits_{R_n} x_i dF(x_1, x_2, ..., x_n) = M[\xi_i]$ - будет значением математического ожидания $i$-той компоненты случайной величины $\xi$.

Определение 1. ***Математическим ожиданием*** $n$-мерной случайной величины $\xi$ называется матрица-столбец $(n \times 1)$:

$$M\xi = \begin{pmatrix} M\xi_1 \\ ... \\ M\xi_i \\ ... \\ M\xi_n \end{pmatrix}, \text{ если все интегралы } \int\limits_{R_1} ... \int\limits_{R_n} x_i dF(x_1, ..., x_n) = M[\xi_i], \ i = 1,2,...,n,$$

сходятся абсолютно.

Ясно, что все свойства математического ожидания одномерной случайной величины будут справедливы (с соответствующей корректировкой записи) для математического ожидания многомерной случайной величины.

Будем рассматривать двумерный случай $(\boldsymbol{n}=2)$, имея в виду, что определение числовых характеристик многомерных случайных величин и свойства этих характеристик, рассмотренные при $\boldsymbol{n}=2$, легко обобщаются на случай $\boldsymbol{n}>2$. Двумерную случайную величину будем записывать так: $\boldsymbol{\zeta}=(\boldsymbol{\xi},\boldsymbol{\eta})$, где случайные величины $\boldsymbol{\xi}$ и $\boldsymbol{\eta}$ - компоненты случайной величины $\boldsymbol{\zeta}$. Как и в одномерном случае рассматривается $\boldsymbol{z}=\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y})$ непрерывная степенная функция двух переменных.

Определение 2. *Начальным моментом* $\boldsymbol{\alpha}_{ik}$ порядка $(\boldsymbol{i},\boldsymbol{k})$ двумерной случайной величины называется:

$\boldsymbol{\alpha}_{ik}=\boldsymbol{M}\left[\boldsymbol{\xi}^{i}\cdot\boldsymbol{\eta}^{k}\right]=\iint\limits_{\boldsymbol{R}^2}\boldsymbol{x}^{i}\boldsymbol{y}^{k}\boldsymbol{dF}(\boldsymbol{x},\boldsymbol{y})$, если интеграл сходится абсолютно.

В частности: $\boldsymbol{\alpha}_{00}=1$; $\boldsymbol{\alpha}_{10}=\boldsymbol{M\xi}=\boldsymbol{m}_1$; $\boldsymbol{\alpha}_{01}=\boldsymbol{M\eta}=\boldsymbol{m}_2$;

$\boldsymbol{\alpha}_{20}=\boldsymbol{M\xi}^2$; $\boldsymbol{\alpha}_{02}=\boldsymbol{M\eta}^2$ и $\boldsymbol{\alpha}_{11}=\boldsymbol{M}[\boldsymbol{\xi}\cdot\boldsymbol{\eta}]$.

Ясно, что если случайные величины $\boldsymbol{\xi}$ и $\boldsymbol{\eta}$ - независимые, то $\boldsymbol{\alpha}_{11}=\boldsymbol{m}_1\cdot\boldsymbol{m}_2$.

Определение 2. *Центральным моментом* $\boldsymbol{\mu}_{ik}$ порядка $(\boldsymbol{i},\boldsymbol{k})$ называется

$$\boldsymbol{\mu}_{ik}=\boldsymbol{M}\left[(\boldsymbol{\xi}-\boldsymbol{M\xi})^{i}\cdot(\boldsymbol{\eta}-\boldsymbol{M\eta})^{k}\right]=\iint\limits_{\boldsymbol{R}^2}(\boldsymbol{x}-\boldsymbol{m}_1)^{i}\cdot(\boldsymbol{y}-\boldsymbol{m}_2)^{k}\boldsymbol{dF}(\boldsymbol{x},\boldsymbol{y}),$$

если интеграл сходится абсолютно.

В частности: $\boldsymbol{\mu}_{00}=1$; $\boldsymbol{\mu}_{10}=\boldsymbol{M}[\boldsymbol{\xi}-\boldsymbol{M\xi}]=0$; $\boldsymbol{\mu}_{01}=\boldsymbol{M}[\boldsymbol{\eta}-\boldsymbol{M\eta}]=0$. В практических задачах используются центральные моменты, у которых $\boldsymbol{i}+\boldsymbol{k}=2$:

$\boldsymbol{\mu}_{20}=\boldsymbol{M}[\boldsymbol{\xi}-\boldsymbol{M\xi}]^2=\boldsymbol{D\xi}=\boldsymbol{\sigma}_1^2$; $\boldsymbol{\mu}_{02}=\boldsymbol{M}[\boldsymbol{\eta}-\boldsymbol{M\eta}]^2=\boldsymbol{D\eta}=\boldsymbol{\sigma}_2^2$ и

$\boldsymbol{\mu}_{11}=\boldsymbol{M}[(\boldsymbol{\xi}-\boldsymbol{M\xi})\cdot(\boldsymbol{\eta}-\boldsymbol{M\eta})]$. Ясно, что: $\boldsymbol{\mu}_{11}=\boldsymbol{\alpha}_{11}-\boldsymbol{m}_1\cdot\boldsymbol{m}_2$.

Если случайные величины $\boldsymbol{\xi}$ и $\boldsymbol{\eta}$ независимые, то, очевидно, $\boldsymbol{\mu}_{11}=0$.

Обратное не всегда верно, то есть равенство $\boldsymbol{\mu}_{11}=0$ ещё не означает, что случайные величины $\boldsymbol{\xi}$ и $\boldsymbol{\eta}$ независимые. Что иллюстрирует следующий пример.

## Практика

Пример. Пусть двумерная случайная величина $\boldsymbol{\zeta}=(\boldsymbol{\xi},\boldsymbol{\eta})$ подчиняется равномерному распределению на круге $\boldsymbol{\Omega}=\left\{(\boldsymbol{x},\boldsymbol{y})\in\boldsymbol{R}^2 : \boldsymbol{x}^2+\boldsymbol{y}^2\le\boldsymbol{r}^2\right\}$, то есть

$\boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})=\begin{cases}0, & \text{если } (\boldsymbol{x},\boldsymbol{y})\notin\boldsymbol{\Omega}\\ \dfrac{1}{\boldsymbol{\pi}\cdot\boldsymbol{r}^2}, & \text{если } (\boldsymbol{x},\boldsymbol{y})\in\boldsymbol{\Omega}\end{cases}$. Очевидно, что случайные величины $\boldsymbol{\xi}$ и $\boldsymbol{\eta}$ при таком распределении вероятностей *не будут независимыми*, так как при любом возможном значении $\boldsymbol{x}_0$, $(|\boldsymbol{x}_0|\le\boldsymbol{r})$ случайной величины $\boldsymbol{\xi}$ случайная величина $\boldsymbol{\eta}$ может принять значения $\boldsymbol{y}$, удовлетворяющие неравенству $|\boldsymbol{y}|\le\sqrt{\boldsymbol{r}^2-\boldsymbol{x}_0^2}$.

В то же время значение $\boldsymbol{\mu}_{11}=\boldsymbol{\alpha}_{11}-\boldsymbol{m}_1\cdot\boldsymbol{m}_2$ будет равно нулю, так как:

$m_1 = M[\xi^1 \cdot \eta^0] = 0$, $m_2 = M[\xi^0 \cdot \eta^1] = 0$ и $\alpha_{11} = M[\xi^1 \cdot \eta^1] = \iint\limits_{\Omega} xy \cdot \dfrac{1}{\pi \cdot r} dxdy = 0$.

Центральный момент $\mu_{11}$ называют ***ковариационным моментом*** случайных величин $\xi$ и $\eta$, кратко: $\mu_{11} = cov(\xi, \eta)$.

Определение 3. Симметричная матрица $\Sigma$, составленная из центральных моментов $\mu_{ik}$, у которых $i + k = 2$, называется *ковариационной матрицей* двумерной случайной величины $\zeta = (\xi, \eta)$:

$$\Sigma = \begin{pmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \mu_{11} \\ \mu_{11} & \sigma_2^2 \end{pmatrix}.$$

***Комментарий***

Если рассматривать дисперсию суммы случайных величин в общем случае, то получаем:

$$D[\xi + \eta] = M[(\xi + \eta) - M(\xi + \eta)]^2 = M[(\xi - M\xi) + (\eta - M\eta)]^2 = D\xi + 2\mu_{11} + D\eta.$$

В частности, если случайные величины $\xi$ и $\eta$ связаны линейной функциональной зависимостью $\eta = a\xi + b$, то, $D[\xi + \eta] = D\xi + 2aD\xi + D\eta$, то есть $\mu_{11} = cov(\xi, \eta) = a\sigma_1^2$.

Возникает предположение: нельзя ли использовать величину $|\mu_{11}|$ в качестве меры силы стохастической зависимости случайных величин?

Однако, ковариационный момент $\mu_{11} = cov(\xi, \eta)$ как мера силы стохастической зависимости неудобен по следующим причинам.

Во-первых, размерность ковариационного момента равна произведению размерностей случайных величин $\xi$ и $\eta$.

Во-вторых, малое значение величины $|\mu_{11}|$ может вызываться не слабой зависимостью случайных величин $\xi$ и $\eta$, а малостью разброса значений одной из случайных величин около её математического ожидания.

Этих недостатков лишён *коэффициент линейной корреляции*.

**Теория**

Определение 3. *Коэффициентом линейной корреляции* $\rho$ случайных величин $\xi$ и $\eta$ называется математическое ожидание произведения центрированных и нормированных случайных величин $\dfrac{\xi - M\xi}{\sqrt{D\xi}}$ и $\dfrac{\eta - M\eta}{\sqrt{D\eta}}$:

$$\rho = M\left[\frac{\xi - M\xi}{\sqrt{D\xi}} \cdot \frac{\eta - M\eta}{\sqrt{D\eta}}\right] = \frac{\mu_{11}}{\sigma_1 \cdot \sigma_2}.$$

Из определения видно, что коэффициент линейной корреляции $\rho$ безразмерная характеристика, а произведение средних квадратических отклонений, находящееся в знаменателе, «погашает» влияние значений разбросов случайных величин около их математических ожиданий на величину коэффициента линейной корреляции $\rho$.

Теорема. (Свойства коэффициента линейной корреляции).

*Для любых случайных величин* $\xi$ *и* $\eta$ *модуль значения коэффициента линейной корреляции не превосходит единицы* ($|\rho| \le 1$)*, и модуль значения коэффициента линейной корреляции равен единице* ($|\rho| = 1$) *тогда и только тогда, когда случайные величины* $\xi$ *и* $\eta$ *связаны линейной функциональной зависимостью:* $\eta = a\xi + b$.

Для доказательства теоремы рассмотрим сначала случайную величину $\zeta$ – линейную функцию $\zeta = \dfrac{\xi - m_1}{\sigma_1} \pm \dfrac{\eta - m_2}{\sigma_2}$ произвольных случайных величин $\xi$ и $\eta$, имеющих математические ожидания $m_1$, $m_2$ и дисперсии $\sigma_1^2$, $\sigma_2^2$.

Вычислим значение дисперсии случайной величины $\zeta$, используя формулу $D\zeta = M\zeta^2 - M^2\zeta$. Ясно, что $M\zeta = 0$, а для вычисления $M\zeta^2$ сначала запишем вид $\zeta^2$, а потом воспользуемся свойствами математического ожидания:

$$\zeta^2 = \frac{(\xi - m_1)^2}{\sigma_1^2} \pm 2 \cdot \frac{(\xi - m_1)(\eta - m_2)}{\sigma_1\sigma_2} + \frac{(\eta - m_2)^2}{\sigma_2^2};$$

$$M\zeta^2 = \frac{1}{\sigma_1^2} M[\xi - m_1]^2 \pm \frac{2}{\sigma_1\sigma_2} M[(\xi - m_1)(\eta - m_2)] + \frac{1}{\sigma_2^2} M[\eta - m_2]^2;$$

Итак: $$D\zeta = 1 \pm 2 \cdot \frac{\mu_{11}}{\sigma_1\sigma_2} + 1 = 2(1 \pm \rho). \qquad (*)$$

Доказательство теоремы.

1) Убедимся сначала, что всегда $|\rho| \le 1$.

Значение дисперсии любой случайной величины – неотрицательное число, то есть $D\zeta \ge 0$. Из полученного выражения $(*)$ для значения дисперсии $D\zeta$ следует, что всегда выполняются неравенства $\begin{cases} 1 + \rho \ge 0 \\ 1 - \rho \ge 0 \end{cases}$. $\qquad (**)$

Из этих неравеств заключаем, что всегда $|\rho| \le 1$.

2) Докажем, что, для того чтобы выполнялось $|\rho| = 1$, необходимо выполнение условия линейной зависимости между случайными величинами $\xi$ и $\eta$.

Если $|\rho| = 1$, то, независимо от знака коэффициента линейной корреляции $\rho$, из системы $(**)$ заключаем, что всегда: $D\zeta = 0$. Из свойств дисперсии следует, что если $D\zeta = 0$, то $P(\zeta = c) = 1$. То есть, с вероятностью равной единице, случайная величина $\zeta$ – вырожденная в точке $c$, или $\dfrac{\xi - m_1}{\sigma_1} \pm \dfrac{\eta - m_2}{\sigma_2} = c$. Линейная функциональная зависимость между случайными величинами $\xi$ и $\eta$ - очевидна.

3) Докажем теперь достаточность требования линейной функциональной зависимости между случайными величинами $\xi$ и $\eta$ для выполнения равенства

$|\rho| = 1$.

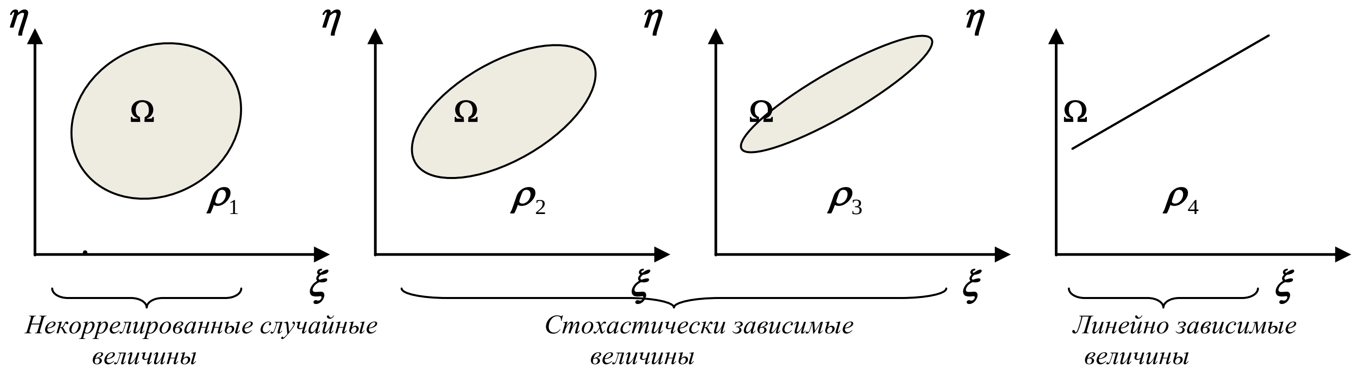
Для доказательстваь этого равенства, вычислим значение коэффициента линейной корреляции, считая, что $\eta = a\xi + b$.

Так как: $m_2 = am_1 + b$ и $\sigma_2^2 = a^2\sigma_1^2$, то $\rho = M\left[\frac{\xi - m_1}{\sigma_1} \cdot \frac{(a\xi + b) - (am_1 + b)}{|a|\sigma_1}\right] =$

$= \frac{a}{|a|\sigma_1^2} \cdot M[\xi - m_1]^2 = \frac{a}{|a|}$. Следовательно: $|\rho| = 1$.

***Комментарий***

Коэффициент линейной корреляции $\rho$ оценивает силу стохастической связи между случайными величинами $\xi$ и $\eta$, в том случае, когда эта связь имеет линейный характер. Чем сильнее эта связь, то есть, чем меньше, при каждом возможном значении одной из случайных величин, диапазон возможных значений другой случайной величины, тем ближе значение $|\rho|$ к единице, и - наоборот. Случайные величины $\xi$ и $\eta$, у которых коэффициент линейной корреляции равен нулю, называются *некоррелированными случайными величинами.*




Здесь: $0 \approx \rho_1 < \rho_2 < \rho_3 < \rho_4 = 1$. Ясно, что стохастическая зависимость между случайными величинами $\xi$ и $\eta$ на втором рисунке слабее, чем стохастическая зависимость между случайными величинами $\xi$ и $\eta$ на третьем рисунке.

**Теория**

Пусть $\xi = (\xi_1, \xi_2, ..., \xi_n)$ - $n$-мерная случайная величина и $p_n(x) = p(x_1, ..., x_n)$ - её плотность вероятности. Дисперсия $l$-той компоненты $\xi_l$ равна $\sigma_l^2 = M[\xi_l - M\xi_l]^2$ и для любой пары компонент $\xi_l$ и $\xi_j$ ковариационным моментом $cov(\xi_l, \xi_j)$ будет $\mu_{lj} = M[(\xi_l - M\xi_l)\cdot(\xi_j - M\xi_j)]$. Ясно, что $\mu_{lj} = \mu_{jl}$.

<u>Определение.</u> Симметричная, положительно определённая, порядка $n \times n$ матрица: $\Sigma = \begin{pmatrix} \sigma_1^2 & \mu_{12} & ... & \mu_{1n} \\ \mu_{21} & \sigma_2^2 & ... & \mu_{2n} \\ ... & ... & ... & ... \\ \mu_{n1} & \mu_{n2} & ... & \sigma_n^2 \end{pmatrix}$ называется *ковариационной матрицей*.

Обозначим: $\left(\boldsymbol{x}-\boldsymbol{m}\right)=\begin{pmatrix}\boldsymbol{x}_1-\boldsymbol{m}_1\\ \dots\\ \boldsymbol{x}_n-\boldsymbol{m}_2\end{pmatrix}$ , $\left(\boldsymbol{x}-\boldsymbol{m}\right)'$- транспонированную матрицу $\left(\boldsymbol{x}-\boldsymbol{m}\right)$, и $\left|\boldsymbol{\Sigma}^{-1}\right|=\boldsymbol{det}\boldsymbol{\Sigma}^{-1}$ - определитель обратной матрицы $\boldsymbol{\Sigma}^{-1}$.

Определение. Будем говорить, что случайная величина $\xi=(\xi_1,\xi_2,...,\xi_n)$ подчиняется ***n**-мерному нормальному закону,* $\xi\in\mathcal{N}(\boldsymbol{m},\boldsymbol{\Sigma})$, если её плотность вероятности имеет вид: $\boldsymbol{p}_n\left(\boldsymbol{x}\right)=\boldsymbol{p}\left(\boldsymbol{x}_1,...,\boldsymbol{x}_n\right)=\dfrac{\left|\boldsymbol{\Sigma}^{-1}\right|^{1/2}}{(2\pi)^{n/2}}\cdot\boldsymbol{e}^{-\frac{1}{2}\left[(\boldsymbol{x}-\boldsymbol{m})'\cdot\Sigma^{-1}\cdot(\boldsymbol{x}-\boldsymbol{m})\right]}$.

В частности, если $\boldsymbol{n}$=2, то $\mu_{12}=\rho\cdot\sigma_1\sigma_2$, $\boldsymbol{\Sigma}^{-1}=\dfrac{1}{\left(1-\rho^2\right)^2}\begin{pmatrix}\dfrac{1}{\sigma_2^2} & -\dfrac{\rho}{\sigma_1\sigma_2}\\ -\dfrac{\rho}{\sigma_1\sigma_2} & \dfrac{1}{\sigma_1^2}\end{pmatrix}$ и

$\left|\boldsymbol{\Sigma}^{-1}\right|=\dfrac{1}{\sigma_1^2\sigma_2^2\left(1-\rho^2\right)}$. В результате получаем уже знакомую нам плотность вероятности двумерной случайной величины, распределённой по нормальному закону:

$$\boldsymbol{p}_2\left(\boldsymbol{x}\right)=\boldsymbol{p}\left(\boldsymbol{x}_1,\boldsymbol{x}_2\right)=\frac{1}{2\pi\cdot\sigma_1\sigma_2\cdot\sqrt{1-\rho^2}}\boldsymbol{e}^{-\frac{1}{2\left(1-\rho^2\right)}\left[\frac{(x_1-m_1)^2}{\sigma_1^2}-2\rho\frac{(x_1-m_1)(x_2-m_2)}{\sigma_1\sigma_2}+\frac{(x_2-m_2)^2}{\sigma_2^2}\right]}.$$

## §8.IV Гильбертово пространство случайных величин

Рассматривается множество $\Xi=\{\xi\}$ случайных величин, имеющих конечный начальный момент второго порядка, то есть величин, для которых интеграл $\alpha_2=\boldsymbol{M}\xi^2=\int\limits_{\boldsymbol{R}}\boldsymbol{x}^2\boldsymbol{dF}\left(\boldsymbol{x}\right)$ сходится абсолютно. Для любых $\xi$ и $\eta$, являющихся элементами $\Xi$, определим операцию:

$$(\xi,\eta)\equiv\boldsymbol{M}[\xi\cdot\eta]=\iint\limits_{\boldsymbol{R}^2}\boldsymbol{x}\cdot\boldsymbol{y}\boldsymbol{dF}\left(\boldsymbol{x},\boldsymbol{y}\right).$$

Ясно, что введённая операция обладает свойствами, вытекающими из свойств интеграла Римана-Стилтьеса:

1) $(\boldsymbol{a}\xi_1+\boldsymbol{b}\xi_2,\eta)=\boldsymbol{a}(\xi_1,\eta)+\boldsymbol{b}(\xi_2,\eta)$, $\boldsymbol{a},\boldsymbol{b}\in\boldsymbol{R}$;

2) всегда $(\xi,\xi)\geq 0$;

3) $(\xi,\xi)=0\Leftrightarrow\boldsymbol{P}(\xi=0)=1$.

Значит, введённая операция $(\xi,\eta)$ является скалярным произведением случайных величин $\xi$ и $\eta$.

Определим норму элементов $\Xi$: $\|\xi\|=(\xi,\xi)^{1/2}=\sqrt{\boldsymbol{M}\left[\xi^2\right]}$ и введём метрику - расстояние между случайными величинами $\xi$ и $\eta$:

$$\boldsymbol{d}(\xi,\eta)=\|\xi-\eta\|=(\xi-\eta,\xi-\eta)^{1/2}.$$

Если $(\xi,\eta)=0$, то будем говорить, что случайные величины $\xi$ и $\eta$ ортогональны, то есть $\xi \perp \eta$.

Будем говорить, что последовательность $\{\xi_n\}$ сходится в среднем квадратическом к случайной величине $\xi$, если при $n \to \infty$ будет выполняться $M[\xi_n - \xi]^2 = (\xi_n - \xi, \xi_n - \xi) \to 0$.

Следовательно, относительно нормы $\|\xi\| = (\xi,\xi)^{1/2}$, индуцируемой скалярным произведением, множество $\Xi = \{\xi\}$ становится полным, метрическим пространством, называемым *гильбертовым пространством* случайных величин с конечным начальным моментом второго порядка.

Гильбертово пространство является обобщением понятия евклидова пространства на бесконечномерный случай.

Обозначим $l_0$ - множество вырожденных случайных величин, то есть таких, что $P(\xi = c)=1$. Ясно, что $l_0$ - подпространство $\Xi$.

Пусть $\xi \in \Xi \ / \ l_0$. Среди элементов $l_0$ найдём такой элемент $m_\xi$, что расстояние $d(\xi, m_\xi) = \min\limits_{c\in l_0} d(\xi, c)$.

Для произвольного элемента $c \in l_0$ вычислим:

$$d^2(\xi,c) = \|\xi - c\|^2 = (\xi - c, \xi - c) = M[\xi - c]^2 = M[(\xi - M\xi) - (M\xi - c)]^2 =$$
$$= M[\xi - M\xi]^2 + 2(M\xi - c)\cdot M[\xi - M\xi] + (M\xi - c)^2 = D\xi + (M\xi - c)^2.$$

Ясно, что $\min\limits_{c\in l_0} d(\xi,c) = \min\limits_{c\in l_0}\|\xi - c\|$ достигается если $c = M\xi = m_\xi$ и $\|\xi - m_\xi\|^2 = D\xi$

А так как $(\xi - m_\xi, m_\xi) = M[(\xi - m_\xi)\cdot m_\xi] = 0$, то $\xi - m_\xi \perp m_\xi$ и приходим к выводу, что $m_\xi = proj_{l_0}\xi$, то есть математическое ожидание случайной величины $\xi$ это её проекция на подпространство $l_0$, а квадрат длины перпендикуляра $\xi - m_\xi$ к этому подпространству равен дисперсии случайной величины $\xi$.

Если случайные величины $\xi$ и $\eta$ рассматривать как два произвольных вектора пространства $\Xi$, то угол между этими векторами равен углу между векторами $\overset{0}{\xi} = \xi - M\xi$ и $\overset{0}{\eta} = \eta - M\eta$. В метрическом пространстве можно определить косинус угла между двумя векторами:

$$cos(\xi,\eta) = cos\left(\overset{0}{\xi},\overset{0}{\eta}\right) = \frac{\left(\overset{0}{\xi},\overset{0}{\eta}\right)}{\left\|\overset{0}{\xi}\right\|\cdot\left\|\overset{0}{\eta}\right\|} = \frac{M[(\xi - M\xi)\cdot M(\eta - M\eta)]}{\sqrt{D\xi}\cdot\sqrt{D\eta}} = \frac{cov(\xi,\eta)}{\sigma_\xi \cdot \sigma_\eta} = \rho.$$

То есть в пространстве случайных величин $\Xi$ с конечным начальным моментом второго порядка коэффициент линейной корреляции – это косинус угла между двумя случайными величинами – векторами: $\rho = cos(\xi,\eta)$.

Мы получили возможность ещё раз доказать свойства коэффициента линейной корреляции $\rho$.

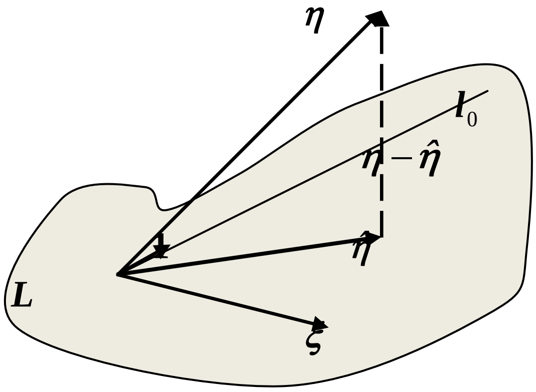
Так как всегда $|cos(\xi,\eta)| \le 1$, то $|\rho| = |cos(\xi,\eta)| \le 1$, для любых случайных величин $\xi$ и $\eta$.

Если $|\rho| = 1$, то $|cos(\xi,\eta)| = 1$. Значит угол между случайными величинами-векторами $\xi$ и $\eta$ равен или 0, или $\pi$, то есть векторы $\xi$ и $\eta$ - коллинеарные, или $\eta = a\xi + b$.

И, наоборот, если $\eta = a\xi + b$, то векторы $\xi$ и $\eta$ - коллинеарные, а это значит, что угол между ними равен или 0, или $\pi$, то есть $|cos(\xi,\eta)| = 1$, следовательно: $|\rho| = 1$.

Пусть случайная величина $\xi \notin l_0$ и вырожденная случайная величина $c = 1 \in l_0$. Рассмотрим подпространство $L \subset \Xi$ - множество случайных величин-векторов вида $\zeta = a \cdot \xi + b \cdot \mathbf{1}$, где $a, b \in R$. Подпространство $L$ в функциональном анализе называется *линейной оболочкой векторов* $\xi$ *и* **1**: $L = Span(\xi, \mathbf{1})$.

Пусть случайная величина $\eta \notin L$. Найдём случайную величину $\hat{\eta} = proj_L \eta$. Ясно, что должно быть: $\hat{\eta} = a \cdot \xi + b \cdot 1$. Так как $\hat{\eta} = proj_L \eta$, то вектор $\eta - \hat{\eta}$ будет перпендикулярен двум любым неколлинеарным векторам подпространства $L$.




Значит, коэффициенты $a$ и $b$ можно найти из условий: $\begin{cases} \eta - \hat{\eta} \perp \xi \\ \eta - \hat{\eta} \perp 1 \end{cases}$, которые приводят нас к системе уравнений: $\begin{cases} (\eta - \hat{\eta}, \xi) = 0 \\ (\eta - \hat{\eta}, 1) = 0 \end{cases}$, или: $\begin{cases} M[(\eta - \hat{\eta}) \cdot \xi] = 0 \\ M[(\eta - \hat{\eta}) \cdot 1] = 0 \end{cases}$, то есть: $\begin{cases} M[\xi \cdot \eta] - M[\xi \cdot \hat{\eta}] = 0 \\ M\eta - M\hat{\eta} = 0 \end{cases}$. Так как $\hat{\eta} = a \cdot \xi + b \cdot 1$, то $\begin{cases} \alpha_{11} - M[\xi \cdot (a\xi + b)] = 0 \\ m_2 - M[a\xi + b] = 0 \end{cases}$,

и мы получаем систему двух линейных уравнений с двумя неизвестными:

$\begin{cases} a \cdot M[\xi^2] + b \cdot m_1 = \alpha_{11} \\ a \cdot m_1 + b = m_2 \end{cases}$. Решениями этой системы будут: $a = \rho \cdot \dfrac{\sigma_2}{\sigma_1}$ и $b = m_2 - \rho \cdot \dfrac{\sigma_2}{\sigma_1} \cdot m_1$. Следовательно, $\hat{\eta} = \rho \dfrac{\sigma_2}{\sigma_1} \cdot (\xi - m_1) + m_2$.

Мы ещё вернёмся к объяснению вероятностного смысла полученной линейной зависимости между элементами $\xi$ и $\hat{\eta}$.

***Комментарий***

Если $0 < |\rho| < 1$, то между случайными величинами (если $0 < |\rho| < 1$) существует зависимость, называемая стохастической зависимостью. Стохастическая зависимость – это зависимость случайной величины $\eta$ относительно функции распределения $F_1(x)$ случайной величины $\xi$. Если случайную величину $\eta$ представить в виде суммы $\eta=(\eta-\hat{\eta})+\hat{\eta}$, то можно сказать, что значения случайной величины $\eta$ определяются: *а)* её внутренними особенностями – это значения случайной величины $(\eta-\hat{\eta})$, и *б)* влиянием на неё случайной величиной $\xi$ - это значения случайной величины $\hat{\eta}$.

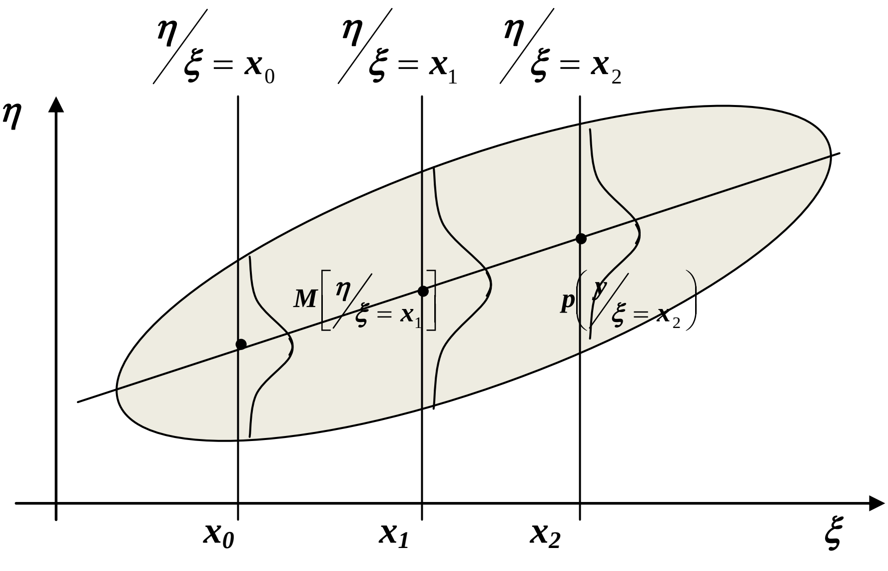
## §9.IV Условные распределения и условные математические ожидания

### Теория

Пусть $\xi$ и $\eta$ компоненты двумерной случайной величины $\zeta=(\xi,\eta)$, которые мы не можем назвать независимыми, и $F(x,y)$ - функция распределения $\zeta$.

В отличие от функциональной зависимости $\eta=g(\xi)$, когда каждому значению случайной величины $\xi$ соответствует одно определённое значение случайной величины $\eta$, при стохастической зависимости одному значению $\xi$ могут соответствовать различные значения случайной величины $\eta$, зависящие от этого значения $\xi$.

То есть, в случае не функциональной, а стохастической зависимости между компонентами $\xi$ и $\eta$ каждое возможное значение $x_i$ случайной величины $\xi$ определяет некоторую область возможных значений и распределение вероятностей случайной величины $\eta/_{\xi=x_i}$, которую будем называть *условной случайной величиной.* Распределение вероятностей условной случайной величины $\eta/_{\xi=x_i}$, очевидно, зависит от распределения вероятностей случайной величины $\xi$.




Ясно, что, при $\xi=x_i$ значение интеграла $F(x_i,y)=\int\limits_{R_2} dF(x_i,y)=F_1(x_i)$

равно вероятности наступления случайного события $\{x_i \times R_2\}$. Если это значение взять в качестве нормирующего множителя, то функция $\dfrac{F(x_i, y)}{F_1(x_i)} = F\left(y / \xi = x_i\right)$ будет функцией распределения условной случайной величины $\eta / \xi = x_i$.

Функцию $F\left(y / \xi = x_i\right)$ будем называть *условной функцией распределения* условной случайной величины $\eta / \xi = x_i$.

Если двумерная случайная величина $\zeta = (\xi, \eta)$ дискретного типа, то для каждого возможного значения $x_{i_o}$ мы можем рассмотреть условную случайную величину $\eta / \xi = x_{i_0}$.

Из множества возможных значений $\{(x_i, y_k)\}$, $i = 1,2,..., n,...; k = 1,2,....m...$, случайной величины $\zeta = (\xi, \eta)$ для выбранного значения $\xi = x_{i_0}$ выбираем все пары $(x_{i_0}, y_k)$. Полученное подмножество $\{(x_{i_0}, y_k)\}$, $k = 1,2,...,m_0,...$, будет множеством возможных значений условной случайной величины $\eta / \xi = x_{i_0}$. Для каждого значения $y_k$ условной случайной величины, по аналогии с определением условной вероятности для случайных событий, определим вероятность: $P\left(\eta = y_k / \xi = x_{i_0}\right) = \dfrac{p_{i_0 k}}{p_{i_0 *}} = p_{k / i_0}$. Набор вероятностей $\left\{p_{k / i_o}\right\}$ будем называть *условным распределением вероятностей* дискретной случайной величины $\eta / \xi = x_{i_0}$.

Если двумерная случайная величина $\zeta = (\xi, \eta)$ непрерывного типа, то для каждого возможного значения $x_0$ мы можем рассмотреть условную случайную величину $\eta / \xi = x_0$. Функцию $p\left(y / x_0\right) = \dfrac{p(x_0, y)}{p_1(x_0)}$ будет *условной плотностью вероятности* непрерывной случайной величины $\eta / \xi = x_0$.

Аналогично определяется условная случайная величина $\xi / \eta = y_0$ и её условная функция распределения: $F\left(x / \eta = y_0\right) = \dfrac{F(x, y_0)}{F_2(y_0)}$.

Согласно общему определению математического ожидания функции случайной величины, мы можем определить условное математическое ожидание условной случайной величины $\eta / \xi = x_0$:

$$M\left[{}^{\eta}\!\big/_{\xi = x_0}\right] = \int\limits_{R_2} y \cdot dF\left({}^{y}\!\big/_{\xi = x_0}\right).$$

Так как определённое в предыдущем параграфе подпространстао $L$ – линейное, то мы получили, что случайная величина $\hat{\eta}$, являясь условной случайной величиной, будет линейной функцией случайной величины $\xi$. Поэтому мы имеем право записать, что при $\xi = x_0$ среднее значение условной случайной величины будет равно:

$$M\left[{}^{\eta}\!\big/_{\xi = x_0}\right] = \hat{\eta}(x_0) = ax_0 + b = \rho\frac{\sigma_2}{\sigma_1}(x_0 - m_1) + m_2.$$

Следовательно, условное математическое ожидание является функцией возможных значений случайной величины $\xi$:

$$M\left[{}^{\eta}\!\big/_{\xi = x}\right] = \rho\frac{\sigma_2}{\sigma_1}(x - m_1) + m_2.$$

Эта функция называется ***функцией регрессии*** *случайной величины* $\eta$ *на случайную величину* $\xi$.

Значит, рассматривая график функции регрессии, мы анализируем изменение среднего значения условной случайной величины ${}^{\eta}\!\big/_{\xi = x}$, вызываемое влиянием случайной величины $\xi$.

Слово *регрессия* переводится на русский язык как *отклик.* Поэтому мы можем сказать, что функция регрессии – это функция, описывающая *отклик средних значений условной случайной величины на изменения значений случайной величины* $\xi$.

# ГЛАВА ПЯТАЯ
# Классическая предельная проблема теории вероятностей

## §1.V Теоремы Муавра-Лапласа
### *Комментарий*

Проводится серия $n$ повторных независимых испытаний, в каждом из которых вероятность наступления события $A$ равна $p$. Случайная величина $\xi$ - число наступлений события $A$ - имеет биномиальное распределение вероятностей, которое мы обозначаем $B_n(p)$.

Обычно при рассмотрении повторных независимых испытаний решаются задачи двух типов, в которых надо:

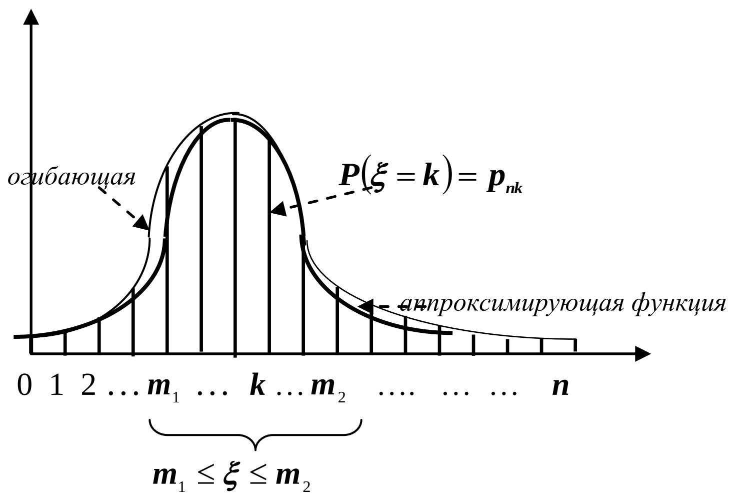
1) определить вероятность того, что в результате проведения $n$ испытаний событие $A$ наступит ровно $k$ раз, то есть определить вероятность того, что случайная величина $\xi$ примет значение $k$;

2) определить вероятность того, что в результате проведения $n$ испытаний число $k$ - число наступлений события $A$ - будет не меньше, чем $m_1$, но и не

больше, чем $m_2$, то есть определить вероятность того, что значение $\xi$ будет удовлетворять неравенству $m_1 \le \xi \le m_2$.

Задача первого типа решается применением формулы Бернулли: $P(\xi = k) = p_{nk} = C_n^k p^k q^{n-k}$. Для решения задачи второго типа надо вычислить значение суммы: $P(m_1 \le \xi \le m_2) = \sum_{k=m_1}^{m_2} p_{nk} = \sum_{k=m_1}^{m_2} C_n^k p^k q^{n-k}$.

При геометрической интерпретации биномиального распределения решение задачи первого типа иллюстрируется определением длины $k$-той вертикали, а решение задачи второго типа иллюстрируется определением суммы длин вертикалей, основаниями которых будут значения $m_1, m_1+1, \ldots, m_2$ случайной величины $\xi$.




## Практика

Пример 1. Будем считать, что вероятности появления на свет мальчика и девочки – одинаковые ($p = \frac{1}{2}$). В семье шесть детей ($n = 6$). Определим случайную величину $\xi$ - число мальчиков в семье.

Вероятность того, что в этой семье четыре мальчика ($k = 4$) будет равна:

$$P(\xi = 4) = C_6^4 \cdot \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^{6-4} = \frac{C_6^4}{2^6} = \frac{15}{64} \approx 0{,}2344.$$

Вероятность того, что число мальчиков в семье находится в диапазоне $[2;4]$, будет равна:

$$P(2 \le \xi \le 4) = \sum_{k=2}^{4} C_6^k \cdot \left(\frac{1}{2}\right)^k \cdot \left(\frac{1}{2}\right)^{6-k} = \frac{C_6^2 + C_6^3 + C_6^4}{2^6} = \frac{50}{64} \approx 0{,}8438.$$

Пример 2. В старшей группе детского сада двадцать пять детей ($n = 25$). Определить вероятность того, что в этой группе пятнадцать мальчиков ($k = 15$) и вероятность того, что в этой группе мальчиков не менее чем одиннадцать, но и не более чем четырнадцать ($11 \le k \le 14$).

Практический смысл этого примера не отличается от практического смысла первого примера, но возникают некоторые трудности вычислительного характера.

$$P(\xi=15)=C_{25}^{15}\cdot\left(\frac{1}{2}\right)^{15}\cdot\left(\frac{1}{2}\right)^{25-15}=\frac{C_{25}^{15}}{2^{25}}\approx 0,0974;$$

$$P(11\le\xi\le 14)=\sum_{k=11}^{14}C_{25}^{k}\cdot\left(\frac{1}{2}\right)^{k}\cdot\left(\frac{1}{2}\right)^{25-k}=\frac{C_{25}^{11}+C_{25}^{12}+C_{25}^{13}+C_{25}^{14}}{2^{25}}=\frac{19315400}{2^{25}}\approx 0,5756.$$

Пример 3. В ЗАГСе города зарегистрировано пятьсот младенцев ($n=500$), появившихся на свет в течение некоторого периода. Определить вероятность того, что среди зарегистрированных младенцев будет двести шестьдесят пять мальчиков ($k=265$). Чему равна вероятность того, что в списках зарегистрированных младенцев число мальчиков будет не меньше чем двести тридцать восемь, но и не больше чем двести семьдесят два?

Ясно, что для вычисления значений вероятностей событий $\{\xi=265\}$и $\{238\le\xi\le 272\}$ надо использовать формулы $P(\xi=265)=C_{500}^{265}\cdot p^{265}\cdot q^{500-265}$ и $P(238\le\xi\le 272)=\sum\limits_{k=238}^{272}C_{500}^{k}\cdot p^{k}\cdot q^{500-k}$. Но при вычислении значений этих вероятностей возникают значительные трудности арифметического характера. Полученные результаты, являясь числами между нулём и единицей, могут вызвать сомнение в их точности из за округлений и приближений, которые придётся делать при вычислениях.

### *Комментарий*

Схема повторных независимых испытаний, определяющая случайную величину $\xi\in B_n(p)$, является вероятностной моделью многих практических задач, в которых необходимо определить вероятности событий $\{\xi=k\}$ и $\{m_1\le\xi\le m_2\}$. Поэтому возникает задача: для серий повторных независимых испытаний большой длительности, то есть когда $n$ велико, найти непрерывную функцию $f(x)$ аппроксимирующую *огибающую кривую* $n+1$ штук вертикалей $p_{nk}$, являющихся вероятностями биномиального распределения $B_n(p)$. Тогда при решении задач первого типа для больших значений $n$ и $k$ можно будет вычислить значение $f(k)$ и это значение $f(k)$ принять в качестве оценки значения вероятности $P(\xi=k)$. При решении задач второго типа $P(m_1\le\xi\le m_2)$, будет достаточно, вычислив значение интеграла $\int\limits_{m_1}^{m_2}f(x)dx$, принять это значение в качестве оценки значения вероятности $P(m_1\le\xi\le m_2)$. Техническая сторона решения этих задач следует из локальной и интегральной теорем Муавра-Лапласа.

### Теория

**Локальная теорема Муавра-Лапласа.** (1733г.,1809г.)

*Если $\xi\in B_n(p)$, то равномерно по всем $k$ таким, что $|k-np|=o(npq)^{2/3}$ для больших значений $n$, будет справедливо: $P(\xi=k)\approx\frac{1}{\sqrt{npq}}\cdot\varphi(x)$, где*

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\ где\ x=\frac{k-np}{\sqrt{npq}}.$$

*И это приближённое равенство будет тем точнее, чем больше* $n$.

Значит, для вычисления приближённого значения вероятности $P(\xi=k)$ надо: вычислить значение $x$, определить для этого значения $x$ значение функции $\varphi(x)$ и полученное значение функции $\varphi(x)$ умножить на число $\frac{1}{\sqrt{npq}}$.

Функция $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ является плотностью вероятности нормального закона $\mathcal{N}(0,1)$, значения этой функции – табулированы.

Можно сделать вывод, что нормальный закон является аппроксимирующим законом для биномиального распределения при больших значениях $n$. Условие $|k-np|=o(npq)^{2/3}$ показывает, что хорошая аппроксимация вероятности $P(\xi=k)$ будет тогда, когда значение $k$ случайной величины $\xi$ мало отличается от значения математического ожидания $M\xi=np$.

***Комментарий***

В книге «Доктрина шансов»(1738) А.Муавр приводит результат своих исследований биномиального распределения , сделанных в 1733 году, в котором вероятность $p$ наступления события в единичном испытании равна $p=\frac{1}{2}$.

П.С.Лаплас усовершенствовал методы доказательств теорем Я.Бернулли и А.Муавра, сделал их менее громоздкими и рассмотрел повторные независимые испытания для любых значений вероятности $p$ наступления события в единичном испытании («Аналитическая теория вероятностей», 1812).

Не рассматривая доказательство теоремы Муавра-Лапласа, проанализируем последовательность преобразований функции $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ при построении аппроксимации *огибающей концов вертикалей* биномиального распределения.

Преобразование $x=\frac{k-np}{\sqrt{npq}}$ переносит график функции $\varphi(x)$ вправо на $np$ единиц и «растягивает» его в $\sqrt{npq}$ раз вдоль оси абсцисс. Умножение значений функции $\varphi\left(\frac{k-np}{\sqrt{npq}}\right)$ на число $\frac{1}{\sqrt{npq}}$ «сжимает» вдоль оси ординат график этой функции в $\sqrt{npq}$ раз. График построенной функции $y=\frac{1}{\sqrt{npq}}\cdot\varphi(x)$, согласно теореме Муавра-Лапласа, аппроксимирует для $k$ таких, что $|k-np|=o(npq)^{2/3}$, функцию, огибающую на геометрической иллюстрации концы вертикалей, длины которых равны вероятностям $\{C_n^k p^k q^{n-k}\}$, где $k=0,1,2,...,n$.

## Практика

В примере 3, определяя значение вероятности $P(\xi = 265) = C_{500}^{265} \cdot p^{265} \cdot q^{500-265}$ с помощью локальной теоремы, мы сначала вычисляем значение аргумента $x = \frac{k - np}{\sqrt{npq}} = \frac{265 - 250}{\sqrt{125}} \approx 1{,}342$. По таблицам значений функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ определяем $\varphi(1{,}342) \approx 0{,}16216$. Так как $\frac{1}{\sqrt{npq}} \approx 0{,}08944$, то получаем значение аппроксимирующей функции $y = \frac{1}{\sqrt{npq}} \cdot \varphi(1{,}342) \approx 0{,}0145$. Это значение принимаем в качестве значения вероятности наступления события $\{\xi = 265\}$, то есть $P(\xi = 265) \approx 0{,}0145$.

## Теория

**Интегральная теорема Муавра-Лапласа.**

*Если* $\xi \in B_n(p)$, *где* $0 < p < 1$, *тогда* $P(m_1 \le \xi \le m_2) \approx \Phi(b) - \Phi(a)$, *где* $b = \frac{m_2 - np}{\sqrt{npq}}$, $a = \frac{m_1 - np}{\sqrt{npq}}$ *и* $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz$ *- функция Лапласа.*

*И это приближённое равенство будет тем точнее, чем больше* $n$.

Значит, для вычисления приближённого значения вероятности $P(m_1 \le \xi \le m_2)$ надо: вычислить значения аргументов $a$ и $b$, из таблиц значений функции Лапласа выписать значения $\Phi(a)$ и $\Phi(b)$, воспользоваться формулой Ньютона-Лейбница, то есть определить значение разности $\Phi(b) - \Phi(a)$.

### *Комментарий*

Очевидно, что $F_n(x) = P\left(\frac{\xi - np}{\sqrt{npq}} < x\right)$ - функция распределения центрированной и нормированной случайной величины $\frac{\xi - np}{\sqrt{npq}}$.

$F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} dx$ - функция распределения случайной величины, имеющей нормальное распределение $\mathcal{N}(0{,}1)$. Интегральная теорема Муавра-Лапласа утверждает, что, при $n \to \infty$, будет справедливо $\sup\limits_{-\infty < x < +\infty} |F_n(x) - F_0(x)| \to 0$. Точность аппроксимации функции $F_n(x)$ функцией $F_0(x)$ оценивается теоремой Берри-Эссеена, частный случай которой утверждает, что $\sup\limits_{-\infty < x < +\infty} |F_n(x) - F_0(x)| \le \frac{p^2 + q^2}{\sqrt{npq}}$.

## Практика

В примере 3 мы можем теперь вычислить приближённое значение вероятности случайного события $\{238 \le \xi \le 272\}$. Вычисляем: $a = \frac{238-250}{\sqrt{125}} = -1{,}073$ и $b = \frac{272-250}{\sqrt{125}} = 1{,}968$. Из таблиц значений функции Лапласа выписываем: $\Phi(a) = \Phi(-1{,}073) = -0{,}3584$ и $\Phi(b) = \Phi(1{,}968) = 0{,}4752$. Применяя формулу Ньютона-Лейбница, получаем:

$$P(238 \le \xi \le 272) \approx \Phi(1{,}968) - \Phi(-1{,}073) = 0{,}4752 + 0{,}3584 = 0{,}8336.$$

<u>*Замечание.*</u> Иллюстрация биномиального распределения вероятностей при $n$=500 состоит из 501 вероятностей $P_{500}(k)$. Определяя вероятность случайного события $\{238 \le \xi \le 272\}$, мы с помощью теоремы Муавра-Лапласа подсчитали сумму только 35 вероятностей $P_{500}(k)$. Эта сумма составляет более 84% суммы всех вероятностей $P_{500}(k)$, которая равна единице. Так как $M\xi = np = 250$ и $\sigma = \sqrt{npq} \approx 11{,}2$, то получается, что вероятность события $\{M\xi - \sigma \le \xi \le M\xi + 2\sigma\}$ близка к единице, то есть в качестве первого вывода можем сказать, что значимые значения вероятностей $P_{500}(\xi = k)$ будут для значений $\xi = k$ близких к значению математического ожидания $np = 250$. Если точнее, то: $P(M\xi - \sigma \le \xi \le M\xi + \sigma) = 2\Phi(1) = 0{,}6426$,

$$P(M\xi - 2\sigma \le \xi \le M\xi + 2\sigma) = 2\Phi(2) = 0{,}9544,$$
$$P(M\xi - 3\sigma \le \xi \le M\xi + 3\sigma) = 2\Phi(3) = 0{,}9973.$$

## §2.V Теорема Бернулли
### Теория

Пусть $\xi \in B_n(p)$ - число наступлений события при проведении $n$ повторных независимых испытаний. Случайную величину $\eta = \frac{1}{n}\xi$ называют *относительной частотой* наступлений события при проведении этих испытаний. Ясно, что всегда $0 < \frac{1}{n}\xi < 1$. (Случаи $p = 0$ или $p = 1$ не представляют никакого интереса). Для больших значений $n$, используя интегральную теорему Муавра-Лапласа, вычислим вероятность случайного события: $\left\{\left|\frac{1}{n}\xi - p\right| < \varepsilon\right\}$, то есть вероятность того что при проведении $n$ испытаний относительная частота наступлений события $\frac{1}{n}\xi$ отклонится от вероятности наступления события в единичном испытании $p$ меньше, чем на сколь угодно малое положительное число $\varepsilon$.

Ясно, что события $\left\{\left|\frac{1}{n}\xi - p\right| < \varepsilon\right\}$, $\{|\xi - np| < \varepsilon\}$ и $\{np - \varepsilon n < \xi < np + n\varepsilon\}$ имеют равные вероятности наступления, то есть:

$$P\left(\left|\frac{1}{n}\xi - p\right| < \varepsilon\right) = P(|\xi - np| < \varepsilon) = P(np - \varepsilon n < \xi < np + \varepsilon n).$$

Для достаточно больших значений $n$, считая, что $np - \varepsilon n = m_1$ и $np + \varepsilon n = m_2$, согласно теореме Муавра-Лапласа получаем:

$P\left(\left|\frac{1}{n}\xi - p\right| < \varepsilon\right) = 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right)$. В полученном равенстве перейдём к пределу при $n \to \infty$: $\lim\limits_{n\to\infty} P\left(\left|\frac{1}{n}\xi - p\right| < \varepsilon\right) = \lim\limits_{n\to\infty} 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right)$. Так как $\Phi(\infty) = \frac{1}{2}$, то окончательно получаем результат, который формулируется как теорема Я.Бернулли.

**Теорема Я.Бернулли** (1713г.)

*При очень большом числе повторных независимых испытаний с вероятностью близкой к единице, то есть практически достоверно, можно утверждать, что относительная частота наступлений события будет сколь угодно мало отличаться от вероятности наступления этого события в единичном испытании:*

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

***Комментарий***

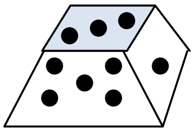
Теорема Я.Бернулли (1654-1705) была сформулирована и доказана в книге «Искусство предположений», изданной его племянником Н.Бернулли в 1713г. Сам Я.Бернулли говорил, что идею этой теоремы он вынашивал двадцать лет.

Рассмотрим несколько примеров, в которых рассматриваются повторные независимые испытания, иллюстрирующих смысл и практическую ценность теоремы Бернулли.

1) Подбрасывается наудачу один раз игральная кость. Пусть случайное событие $A$ – «на верхней грани выпало число очков кратное трём». Ясно, что вероятность наступления этого события при одном подбрасывании равна $P(A) = p = \frac{1}{3}$. Определим случайную величину $\xi$ - число наступлений события $A$ при подбрасывании игральной кости $n$ раз. Случайная величина $\xi$ подчиняется биномиальному распределению вероятностей, $\xi \in B_n(p)$. Тогда случайная величина $\frac{1}{n}\xi$ - это относительная частота наступлений события $A$ при проведении $n$ повторных независимых испытаний, то есть - при подбрасывании игральной кости $n$ раз.

Из теоремы Бернулли следует, что с большой степенью уверенности можно сказать при большом числе подбрасываний игральной кости относительное число наступлений события $A$ будет мало отличаться от вероятности наступления события $A$ при одном подбрасывании, то есть от числа $p = \frac{1}{3}$.

2) Сделана «игральная кость» в виде усечённой четырёхугольной пирамиды.



Эта «кость» подбрасывается наудачу один раз. Определим случайное событие $\boldsymbol{A}$ – «на верхней грани выпало число очков равное трём». Так как элементарные исходы подбрасывания «кости» не будут равновозможными, то мы не можем утверждать, что $\boldsymbol{P}(\boldsymbol{A}) = \boldsymbol{p} = \frac{1}{6}$. Объективно вероятность события $\boldsymbol{A}$ существует и нам хочется знать хотя бы приблизительное значение её.

Будем подбрасывать эту «кость» много раз и фиксировать количество появлений на верхней грани числа очков равное трём. Пусть было сделано $\boldsymbol{n}$ подбрасываний, в которых число очков равное трём появилось $\boldsymbol{m}$ раз. Значит значение $\frac{\boldsymbol{m}}{\boldsymbol{n}}$ - это значение относительной частоты наступлений события $\boldsymbol{A}$. Согласно теореме Бернулли, при большом числе подбрасываний $(\boldsymbol{n} \to \infty)$ с вероятностью $\boldsymbol{\gamma}$ близкой к единице мы можем утверждать, что полученное значение относительной частоты будет очень мало ($\boldsymbol{\varepsilon} > 0$ - мало) отличаться от вероятности $\boldsymbol{p}$ случайного события $\boldsymbol{A}$, то есть $\boldsymbol{P}\left(\left|\frac{\boldsymbol{m}}{\boldsymbol{n}} - \boldsymbol{p}\right| < \boldsymbol{\varepsilon}\right) > \boldsymbol{\gamma}$.

Такую сходимость значений относительной частоты к вероятности наступления события в единичном испытании мы будем называть *сходимостью по вероятности.*

Значит, с уверенностью $\boldsymbol{\gamma}$ мы можем считать, что $\boldsymbol{P}(\boldsymbol{A}) = \boldsymbol{p} \approx \frac{\boldsymbol{m}}{\boldsymbol{n}}$.

При этом мы должны иметь в виду, что неизвестное значение $\boldsymbol{p}$ может значительно отличаться от полученного в результате эксперимента значения относительной частоты $\frac{\boldsymbol{m}}{\boldsymbol{n}}$. Вероятность случайного события $\left\{\left|\frac{\boldsymbol{m}}{\boldsymbol{n}} - \boldsymbol{p}\right| > \boldsymbol{\varepsilon}\right\}$ равна $1-\boldsymbol{\gamma}$.

3) Хорошей иллюстрацией сходимости по вероятности относительной частоты $\frac{\boldsymbol{m}}{\boldsymbol{n}}$ к вероятности $\boldsymbol{p}$ могут служить испытания, которые провели Ж.Бюффон (XVIII век) и К.Пирсон (XIX-XX век).

Ж.Бюффон подбросил монету $\boldsymbol{n}$=4040 раз и зафиксировал $\boldsymbol{m}$=2048 появлений герба. К.Пирсон сначала подбросил монету $\boldsymbol{n}$=12000 раз и зафиксировал $\boldsymbol{m}$=6019 появлений герба. Затем он увеличил число подбрасываний монеты до $\boldsymbol{n}$=24000раз, и зафиксировал $\boldsymbol{m}$=12012 появлений герба.

Относительные частоты выпадений герба равны, соответственно:

$\frac{m}{n}=5{,}00693$; $\frac{m}{n}=5{,}00158$; $\frac{m}{n}=5{,}0005$. Чтобы оценить приближение, с увеличением числа испытаний $n$, значений относительной частоты $\frac{m}{n}$ к вероятности $p=0{,}5$, проанализируем изменение отношения отклонения $\varepsilon=\frac{m}{n}-p$ к стандартной мере отклонения – среднему квадратическому отклонению $\sigma=\frac{1}{2\sqrt{n}}$. В опыте Ж.Бюффона $\frac{\varepsilon}{\sigma}=0{,}8806$; в первом опыте К.Пирсона $\frac{\varepsilon}{\sigma}=0{,}3472$, во втором опыте К.Пирсона $\frac{\varepsilon}{\sigma}=0{,}1548$.

Эти результаты хорошо иллюстрируют утверждение теоремы Бернулли: $\lim\limits_{n\to\infty}P\left(\left|\frac{1}{n}\xi-p\right|<\varepsilon\right)=1$. Однако мы должны иметь в виду, что при многократных повторениях таких опытов не исключено, что мы можем примерно в $30\%$ случаях получить отклонения значения относительной частоты $\frac{m}{n}$ от вероятности $p=0{,}5$ большие, чем на величину среднего квадратического отклонения $\sigma=\frac{1}{2\sqrt{n}}$.

4) В книге Б.В.Гнеденко «Курс теории вероятностей» говорится, что однажды был зафиксирован факт, когда при раздаче тридцати шести карт между четырьмя партнёрами каждый получил девять карт только одной масти. Вероятность наступления случайного события $A$ – «*каждый из четырёх партнёров при раздаче карт получил девять карт одной масти»,* равна $P(A)=\frac{(9!)^4\cdot 4!}{36!}\approx 1{,}1187\cdot 10^{-18}$. То есть вероятность события $A$ очень мала, но она не равна нулю и случайное событие $A$ наступило. Событие $A$ называется *практически невозможным.*

Дадим объяснение этому факту, используя теорему Бернулли.

Пусть случайная величина $\xi$ - число наступлений события $A$ при $n$ раздачах тридцати шести карт между четырьмя партнёрами. Ясно, что случайная величина $\xi$ подчиняется биномиальному распределению, то есть $\xi\in B_n(p)$, где $p=P(A)$. Математическое ожидание $\xi$ равно $M\xi=np$.

Утверждение теоремы Бернулли можно переписать так: $\lim\limits_{n\to\infty}P(|\xi-np|<\varepsilon n)=1$, где $\varepsilon$ - сколь угодно малое положительное число, например можно взять $\varepsilon=1\cdot 10^{-20}$, то есть, подбирая значение $\varepsilon$, мы всегда можем сделать число $\varepsilon n$ достаточно малым.

Вероятность $p$ наступления события $A$ очень мала, но при увеличении числа $n$ - числа раздач карт между четырьмя партнёрами, величина $np$ принимает значения, мало отличающиеся от единицы. Значениями случайной величины $\xi$ являются целые неотрицательные числа: 0,1,2,…. Значит, для того чтобы выполнилось неравенство $|\xi - np| < \varepsilon$, случайная величина $\xi$ с вероятностью близкой к единице должна принять значение равное единице, то есть должно наступить случайное событие $A$.

Анализируя этот пример, мы приходим к выводу: «*как бы ни была мала вероятность* $p$ *наступления события* $A$*, при очень большом числе испытаний* $n$*, с вероятностью* $P$ *близкой к единице, событие* $A$ *наступит хотя бы один раз*».

5) Нас интересует вопросы: «Есть ли жизнь на планетах вращающихся вокруг звёзд, составляющих нашу Галактику?», «Есть ли во Вселенной разумные существа подобные нам?»

Для того чтобы на планете возникла жизнь надо чтобы в состав среды на планете входили углерод, водород, кислород и азот – основные элементы органических веществ, составляющих всё живое. В первый миллиард лет существования Земли природа осуществляла бесчисленное множество раз эксперимент по созданию органических соединений. Необходимые для него реактивы в изобилии имелись в атмосфере и в воде. Химические реакции создали на Земле вещества, которые лежат в основе жизни, в том числе молекулу ДНК, обладающую способностью к самовоспроизводству и являющуюся носителем наследственности всего живого.

Пусть случайное событие $A$ – «на исследуемой планете есть молекула ДНК».

Вероятность $p = P(A)$ создания наудачу молекулы ДНК из элементов таблицы Д.И.Менделеева очень мала. Она сравнима с вероятностью создания точной копии здания МГУ на Воробьёвых горах в результате бросания наудачу кирпичей в большую кучу. И, тем не менее, если случайная величина $\xi$ – число планет, на которых есть жизнь, то, исследовав большое число планет - $n$, мы с уверенностью близкой к единице можем утверждать, что случайная величина $\xi$ примет значение 1, так как $\lim\limits_{n\to\infty} P(|\xi - np| < \varepsilon n) = 1$.

Количество галактик во Вселенной – велико. Количество звёзд даже в нашей Галактике ещё никто не смог подсчитать, а вокруг звёзд вращаются планеты, на некоторых из них могут быть условия для создания и существования органических соединений. Наши средства наблюдения планет и идентификации условий существования жизни на них пока ещё очень слабы и несовершенны. Но, как бы ни была мала вероятность $p$ наступления события $A$, мы можем утверждать с вероятностью $P \approx 1$, что, при увеличении числа исследованных планет $n$, случайная величина $\xi$ примет значение равное единице, то есть наступит событие $\{\xi \ge 1\}$, а, значит, мы будем наблюдать наступление случайного события $A$.

**Теория**

В теоремах Муавра-Лапласа и Бернулли рассматривалась случайная величина $\xi$, которая подчиняется биномиальному распределению $\xi \in B_n(p)$. Ранее было показано, что если $\xi_k, k = 1,2,\dots,n$, независимые бернуллиевские случайные величины, $\xi_k \in B_1(p)$, то $\xi = \sum_{k=1}^{n} \xi_k$ .

Ясно, что вероятности наступления событий $\{m_1 \le \xi < m_2\}$ и $\left\{\frac{m_1 - np}{\sqrt{npq}} \le \frac{\xi - np}{\sqrt{npq}} < \frac{m_2 - np}{\sqrt{npq}}\right\}$ равны. Так как $M\xi_k = p$ и $D\xi_k = pq$, то утверждение интегральной теоремы Муавра-Лапласа можно записать так:

$$P\left(a \le \frac{\sum_{k=1}^{n} \xi_k - \sum_{k=1}^{n} M\xi_k}{\sqrt{\sum_{k=1}^{n} D\xi_k}} < b\right) \approx \Phi(b) - \Phi(a).$$ Обозначим $$\tau_n = \frac{\sum_{k=1}^{n} \xi_k - \sum_{k=1}^{n} M\xi_k}{\sqrt{\sum_{k=1}^{n} D\xi_k}}.$$

Ясно, что $M\tau_n = 0$ и $D\tau_n = 1$, поэтому случайную величину $\tau_n$ будем называть *центрированной и нормированной суммой* независимых бернуллиевских случайных величин.

Согласно теореме Муавра-Лапласа *распределение вероятностей центрированной и нормированной суммы* $\tau_n$ *при больших значениях* $n$ *мало отличается от нормального* $\mathcal{N}(0,1)$. То есть, если $F_n(x) = P(\tau_n < x)$ - функция распределения случайной величины $\tau_n$, то, при $n \to \infty$, будет справедливо $\sup\limits_{-\infty < x < +\infty} |F_n(x) - F_0(x)| \to 0$ .

Если в утверждении теоремы Бернулли: $\lim\limits_{n\to\infty} P\left(\left|\frac{1}{n}\xi - p\right| < \varepsilon\right) = 1$ заменить $\frac{1}{n}\xi$ на $\frac{1}{n}\sum_{k=1}^{n} \xi_k$ , а $p$ на $\frac{1}{n}\sum_{k=1}^{n} M\xi_k$ , то это утверждение будет примет вид:

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} \xi_k - \frac{1}{n}\sum_{k=1}^{n} M\xi_k\right| < \varepsilon\right) = 1,$$

и мы можем утверждать, что при очень большом числе испытаний, с вероятностью близкой к единице, среднее арифметическое значений бернуллиевских случайных величин будет мало отличаться от среднего арифметического их математических ожиданий.

Обозначим $\tau'_n = \frac{1}{n}\sum_{k=1}^{n} \xi_k - \frac{1}{n}\sum_{k=1}^{n} M\xi_k$ . У случайной величины $\tau'_n$ математическое ожидание $M\tau'_n = 0$ и дисперсия $D\tau'_n = \frac{pq}{n}$. Теперь мы теорему Бернулли можем сформулировать так:

$\lim\limits_{n\to\infty} P(|\tau'_m| < \varepsilon) = 1$ - *случайная величина* $\tau'_n$, *являющаяся центрированной суммой независимых бернуллиевских случайных величин, при больших значениях* $n$ *с вероятностью близкой к единице принимает значения, мало отличающиеся от значения вырожденной в нуле случайной величины* $\tau_0$, *то есть – от нуля.*

## §3.V Теорема Пуассона

### *Комментарий*

В теоремах Муавра-Лапласа распределение вероятностей случайной величины $\xi = \sum_{k=1}^{n} \xi_k \in B_n(p)$ аппроксимируется симметричной относительно прямой $x = np$ функцией, получающейся путём преобразований чётной функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

Точность этой аппроксимации оценивается теоремой Берри-Эссеена: $\Delta_n(p) = \frac{p^2 + q^2}{\sqrt{npq}}$. Из формулы, определяющей точность оценки видно, что чем больше будет значение $|p - 0{,}5|$, то есть, чем больше значение $p$ отличается в ту, или иную сторону от 0,5, тем хуже будет аппроксимация.

Например, если $p = 0{,}5$, то $\Delta_n(0{,}5) = \frac{1}{\sqrt{n}}$; если $p = 0{,}1$, то $\Delta_n(0{,}1) \approx \frac{2{,}73}{\sqrt{n}}$; если $p = 0{,}01$, то $\Delta_n(0{,}01) \approx \frac{9{,}85}{\sqrt{n}}$; если $p = 0{,}001$, то $\Delta_n(0{,}001) \approx \frac{31{,}58}{\sqrt{n}}$.

Хорошую аппроксимацию биномиальных распределений $B_n(p)$ при малых значениях $p$ даёт распределение вероятностей Пуассона $\Pi(\lambda)$ с параметром $\lambda$, которое мы назвали «законом редких событий». Напомним: если $\xi \in \Pi(\lambda)$, то $P(\xi = m) = \frac{\lambda^m \cdot e^{-\lambda}}{m!}$.

### Теория

Рассматривается «схема серий» $\{\xi_{nk}\}$, $k = 1,2,\dots,n$, $n = 1,2,\dots$, случайных величин $\xi_{nk} \in B_1(p_n)$. «Схема серий» имеет треугольный вид, в которой все случайные величины $\xi_{nk}$ $n$-той строки имеют одинаковое распределение вероятностей. То есть для всех $k = 1,2,\dots,n$ распределение вероятностей $\xi_{nk}$ зависит от $n$:

| $\xi_{nk}$ | 0 | 1 |
|---|---|---|
| | $q_n$ | $p_n$ |

Обозначим $\sum_{k=1}^{n} \xi_{nk} = \tau_n''$. Ясно, что $\tau_n'' \in B_n(p_n)$.

**Теорема Пуассона** (1837г.)

*Если* $p_n \to 0$, *но так, что* $\lim_{n\to\infty} p_n \cdot n = \lambda$, *то для любого числа* $m$ *будет справедливо* $\lim_{n\to\infty} P(\tau_n'' = m) = \frac{\lambda^m \cdot e^{-\lambda}}{m!}$.

<u>Доказательство</u>. Так как $\tau_n'' \in B_n(p_n)$, то

$$P\left(\tau_n''=m\right)=C_n^m\cdot p_n^m\cdot q_n^{n-m}=\frac{n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot(n-(m-1))}{m!}\cdot p_n^m\cdot(1-p_n)^{n-m}.$$

При больших значениях $n$ мы можем считать, что $p_n\approx\frac{\lambda}{n}$, тогда $P\left(\tau_n''=m\right)\approx\frac{1}{m!}\cdot\frac{n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot(n-(m-1))}{n^m}\cdot\lambda^m\cdot\left(1-\frac{\lambda}{n}\right)^{n-m}$. Перейдём к пределу при $n\to\infty$, так как $m$ – фиксированное число, то получаем:

$$\lim_{n\to\infty}P\left(\tau_n''=m\right)=\frac{\lambda^m}{m!}\lim_{n\to\infty}\left(1-\frac{1}{n}\right)\cdot\ldots\cdot\left(1-\frac{m-1}{n}\right)\cdot\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n\cdot\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-m}=\frac{\lambda^m\cdot e^{-\lambda}}{m!},$$

что и требовалось доказать.

## §4.V Постановка классической предельной проблемы теории вероятностей

### ***Комментарий***

В предыдущих параграфах этой главы мы рассматривали случайную величину $\xi$ - число наступлений события при проведении повторных независимых испытаний, то есть $\xi\in B_n(p)$. Эта случайная величина является суммой независимых бернуллиевских случайных величин $\xi_k\in B_1(p)$, $k=1,2,\ldots,n$, то есть $\xi=\sum_{k=1}^{n}\xi_k$. Из этих случайных величин мы «построили» три случайные величины: $\tau_n=\frac{\sum_{k=1}^{n}\xi_k-\sum_{k=1}^{n}M\xi_k}{\sqrt{\sum_{k=1}^{n}D\xi_k}}$, $\tau_n'=\frac{\sum_{k=1}^{n}\xi_k-\sum_{k=1}^{n}M\xi_k}{n}$, $\tau_n''=\sum_{k=1}^{n}\xi_{nk}$ и рассматривали предельные распределения вероятностей этих случайных величин при неограниченном увеличении числа $n$, то есть при $n\to\infty$.

Теоремы Муавра-Лапласа утверждают, что если $n\to\infty$, то распределение вероятностей случайной величины $\tau_n$ мало отличается от нормального распределения $\mathcal{N}(0,1)$.

Теорема Бернулли утверждает, что если $n\to\infty$, то распределение вероятностей случайной величины $\tau_n'$ мало отличается от вырожденного в нуле распределения $\mathrm{E}(0)$.

Теорема Пуассона утверждает, что если $n\to\infty$, то распределение вероятностей случайной величины $\tau_n''$ мало отличается от распределения вероятностей Пуассона $\Pi(\lambda)$.

Если обозначить, соответственно, $\xi_{nk}=\frac{\xi_k-M\xi_k}{\sqrt{\sum_{k=1}^{n}D\xi_k}}$ и $\xi_{nk}=\frac{\xi_k-M\xi_k}{n}$, то можно говорить, что для каждого $n$ рассматриваемые случайные величины $\tau_n$, $\tau_n'$ и $\tau_n''$ будут суммами элементов $n$-той строки треугольной «схемы серий».

Все случайные величины $\xi_{nk}$ «строятся» из одних и тех же независимых бернуллиевских случайных величин $\{\xi_1, \xi_2, \ldots, \xi_k, \ldots, \xi_n\}$, а предельными распределениями для сумм $\tau_n$, $\tau'_n$ и $\tau''_n$, при $n \to \infty$, будут, соответственно, нормальное, вырожденное и пуассоновское распределения.

Возникает вопрос: «А если снять ограничение, состоящее в том, что случайные величины $\{\xi_k\}$, $k = 1,2,\ldots,n$, - бернуллиевские, то какие ограничения на случайные величины $\{\xi_k\}$, $k = 1,2,\ldots,n$ надо наложить, чтобы предельными распределениями для сумм $\tau_n$, $\tau'_n$ и $\tau''_n$, при $n \to \infty$, были, соответственно, нормальное, вырожденное и пуассоновское распределения»?

Этот вопрос и ответы на него составляют содержание *«Классической предельной проблемы»* теории вероятностей.

Рассматриваются последовательности случайных величин $\{\xi_k\}$, $k = 1,2,\ldots,n$, у которых существует начальный момент второго порядка, то есть для любого номера $k$ выполняется $\alpha_k^2 = M\xi_k^2 < \infty$. Диапазон рассматриваемых последовательностей $\{\xi_k\}$, $k = 1,2,\ldots,n$, - расширяется.

Сначала, что мы уже сделали, рассматриваются последовательности бернуллиевских случайных величин. Затем рассматриваются последовательности одинаково распределенных случайных величин, причём вид закона распределения этих случайных величин нас не интересует. И, наконец, рассматриваются последовательности разнораспределённых случайных величин.

Случайные величины $\{\xi_1, \xi_2, \ldots, \xi_k, \ldots, \xi_n\}$, составляющие эти последовательности, – независимые. При довольно общих ограничениях, налагаемых на эти случайные величины, строятся суммы $\tau_n$, $\tau'_n$ и $\tau''_n$ этих случайных величин и показывается, что при $n \to \infty$ распределения вероятностей этих сумм мало отличаются от нормального, вырожденного и пуассоновского распределения вероятностей.

То есть последовательности случайных величин $\tau_n$, $\tau'_n$ и $\tau''_n$ «сходятся» к случайным величинам $\tau_N$, $\tau_0$ и $\tau_\lambda$, имеющим нормальное, вырожденное и пуассоновское распределение вероятностей.

Нам надо определить: как мы будем понимать сходимость последовательностей $\tau_n$, $\tau'_n$ и $\tau''_n$ к этим предельным случайным величинам.

## Теория

Определение 1. Будем говорить, что последовательность функций распределения $\{F_n(x)\}$ *слабо сходится* к функции распределения $F(x)$ (обозначается $F_n(x) \Rightarrow F(x)$), если имеет место обычная сходимость последовательности функций $\{F_n(x)\}$ в точках непрерывности функции $F(x)$.

Определение 2. Будем говорить, что последовательность случайных величин $\{\eta_1, \eta_2, \ldots, \eta_n\}$ *сходится по распределению* к случайной величине $\eta$ (обозначается $\eta_n \xrightarrow{w} \eta$), если соответствующая ей последовательность функций распределения слабо сходится к функции распределения случайной величины $\eta$.

Определение 3. Будем говорить, что последовательность случайных величин $\{\eta_1, \eta_2, ..., \eta_n\}$ *сходится по вероятности* к случайной величине $\eta$ (обозначается $\eta_n \xrightarrow{P} \eta$), если $\lim\limits_{n\to\infty} P(|\eta_n - \eta| < \varepsilon) = 1$.

Определение 4. Будем говорить, что последовательность случайных величин $\{\eta_1, \eta_2, ..., \eta_n\}$ *сходится почти наверное* к случайной величине $\eta$ (обозначается $\eta_n \xrightarrow{п.н.} \eta$), если $P(\lim\limits_{n\to\infty} \eta_n = \eta) = 1$.

***Комментарий***

«Сходимость *по вероятности*» и «сходимость *почти наверное*» – это формулируемые в терминах теории вероятностей «сходимость *по мере*» и «сходимость *почти всюду*», рассматриваемые в курсе математического анализа.

Каждый следующий вид сходимости «более сильный», чем вид сходимости данный в предыдущем определении. То есть:

$$\boxed{\text{«почти всюду»}} \Rightarrow \boxed{\text{«по вероятности»}} \Rightarrow \boxed{\text{« по распределению»}}$$

Обратное не всегда верно, но, если предел последовательности случайных величин $\{\eta_n\}$, сходящейся по распределению, является вырожденной случайной величиной, то есть, если $\eta_n \xrightarrow{w} c$, и $P(\eta = c) = 1$, то эта последовательность $\{\eta_n\}$ сходится по вероятности к этой вырожденной случайной величине: $\eta_n \xrightarrow{P} c$.

**Теория**

Мы рассматриваем в качестве предельных распределений последовательностей сумм случайных величин три распределения: вырожденное, нормальное и пуассоновское. В соответствии с этим дадим три определения, разделяющие суть классической предельной проблемы на три части.

Определение 1. ***Законом больших чисел*** «**ЗБЧ**» *называется совокупность теорем, в которых на последовательность случайных величин* $\{\xi_k\}$, $k = 1, ..., n$, *налагаются условия, при которых их «сумма»* $\tau'_n = \frac{1}{n}\sum_{k=1}^{n} \xi_k - \frac{1}{n}\sum_{k=1}^{n} M\xi_k$ *сходится по вероятности к вырожденной в нуле случайной величине:* $\lim\limits_{n\to\infty} P(|\tau'_n| < \varepsilon) = 1$, *или* $\tau'_n \xrightarrow{P} \tau_0$.

Другими словами можно сказать так: термин «закон больших чисел» (ввёл этот термин С.Д.Пуассон) означает, что при больших значениях $n$ с вероятностью близкой к единице можно утверждать, что среднее арифметическое случайных величин $\bar{\xi} = \frac{1}{n}\sum_{k=1}^{n} \xi_k$ имеет распределение вероятностей, мало отличающееся распределения вырожденной случайной величины $\tau_c = M\bar{\xi} = \frac{1}{n}\sum_{k=1}^{n} M\xi_k = const$. Или, ещё проще, при больших $n$ с вероятностью близкой к единице среднее арифметическое случайных величин ведёт себя как постоянная величина.

Определение 2. ***Центральной предельной теоремой*** **«ЦПТ»** *называется совокупность теорем, в которых на последовательность случайных величин* $\{\xi_k\}$, $k=1,...,n$, *налагаются условия, при которых их «центрированная и нормированная сумма»* $\tau_n = \dfrac{\sum_{k=1}^{n}\xi_k - \sum_{k=1}^{n}M\xi_k}{\sqrt{\sum_{k=1}^{n}D\xi_k}}$ *сходится по распределению к случайной величине, имеющей нормальное распределение* $\mathcal{N}(0,1)$: $\tau_n \xrightarrow{w} \tau_N$.

Термин «центральная предельная теорема» (ввёл этот термин А.М.Ляпунов) означает, что при больших значениях $n$ функция распределения $F_n(x)$ случайной величины $\tau_n$ мало отличается от функции распределения $F(x)$ нормального закона $\mathcal{N}(0,1)$: $\lim\limits_{n\to\infty} P(\tau_n < x) = \dfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{z^2}{2}}dz$.

***Комментарий***

Закон Пуассона мы назвали «законом редких событий» и познакомились с ним как с законом, по которому распределяется число наступлений события при проведении большой серии повторных независимых испытаний, в каждом из которых вероятность наступления этого события – мала.

Определяя аппроксимацию асимметричного биномиального распределения $B_n(p)$, у которого $M\xi = np$ - мало по сравнению с числом испытаний $n$, мы рассматривали «схему серий» $\{\xi_{nk}\}$, $k=1,2,...,n$, $n=1,2,...$, случайных величин $\xi_{nk} \in B_1(p_n)$, где $p_n \xrightarrow[n\to\infty]{} 0$. Эту «схему серий» можно записать как треугольную матрицу с неограниченно увеличивающимся числом строк. В каждой $n$-той строке случайные величины $\xi_{nk}$ имеют одинаковое распределение $B_1(p_n)$.

В каждой следующей строке треугольной «схемы серий» вероятность $p_n$ наступления события в одном испытании меньше, чем вероятность $p_{n-1}$ в предыдущей строке. С увеличением числа $n$ доля числа событий $\{\tau''_n = k\}$, где $k$ не превосходит некоторого фиксированного числа $m$, уменьшается и становится очень «малой» (например, количество случайных величин $\xi_{nk}$, которые примут значение «единица»). Поэтому мы можем использовать термин, предложенный Г.Корн и Т.Корн - «Закон малых чисел».

**Теория**

Определение 3. ***Законом малых чисел*** «**ЗМЧ**» *называется совокупность теорем, в которых на последовательность случайных величин* $\{\xi_{nk}\}$, $k=1,2,...,n$, $n=1,2,...$, *налагаются условия, при которых их «сумма»* $\tau''_n = \sum_{k=1}^{n}\xi_{nk}$ *сходится по распределению к случайной величине* $\tau_\lambda$, *подчиняющейся распределению* *вероятностей Пуассона* $\Pi(\lambda)$*:* $\tau''_n \xrightarrow{w} \tau_\lambda$.

*То есть:* $\lim\limits_{n\to\infty} P(\tau''_n < x) = \sum\limits_{m<x} \dfrac{\lambda^m \cdot e^{-\lambda}}{m!}$.

***Комментарий***

Мы познакомимся с тремя группами теорем, рассматриваемых в классической предельной проблеме теории вероятностей, зависящих от ограничений, налагаемых на последовательности случайных величин $\{\xi_k\}$, $k = 1,...,n$, и от способа образования из этих последовательностей сумм $\tau_n$, $\tau'_n$ и $\tau''_n$. Эти группы теорем можно записать в виде таблицы.

Таблица не претендует на представление всего списка известных теорем, входящих в ЗБЧ, ЦПТ или ЗМЧ. В этой таблице записаны наиболее важные, часто цитируемые и используемые на практике теоремы.

| | **ЗБЧ** | **ЦПТ** | **ЗМЧ** |
|---|---|---|---|
| Последовательности бернуллиевских случайных величин | Теорема Бернулли (1713г.) | Теорема Муавра-Лапласа (1733г.;1812г.) | Теорема Пуассона (1837г.) |
| Последовательности одинаково распределённых случайных величин | Теорема Хинчина А.Я. (1935г.) | Теорема П.Леви-А.Я.Хинчина (1935г.) | |
| Последовательности разно распределённых случайных величин | Теорема П.Л.Чебышёва (1867-1868г.) Теорема Пуассона (1837г.) | Теорема А.М.Ляпунова (1900-1901г.) Теорема Линдеберга-Феллера (1922г.;1934г.) | |

В теоремах каждой из трёх групп перечисляются требования, налагаемые на последовательность случайных величин $\{\xi_k\}$, $k = 1,...,n$, и формулируется утверждение о виде предельного распределения вероятностей последовательности сумм, соответственно $\tau_n$, $\tau'_n$ или $\tau''_n$, рассматриваемых случайных величин $\{\xi_k\}$.

Ход (суть) доказательства утверждений каждой теоремы можно схематически представить следующим образом.

Рассматривается последовательность случайных величин $\{\xi_k\}$, $k = 1,...,n$, на которые наложены некоторые ограничения. Строится сумма этих случайных величин $\sum\limits_{k=1}^{n} \xi_k = \eta_n$. Число случайных величин-слагаемых $\{\xi_k\}$ неограниченно увеличивается и нас интересует закон распределения предельной для последовательности $\{\eta_n\}$ случайной величины $\eta$, которую условно запишем так

$\lim\limits_{n\to\infty}\eta_n=\eta$.

$$\{\xi_k\},\ k=1,...,n, \Longrightarrow \eta_n=\sum_{k=1}^{n}\xi_k \Longrightarrow \lim_{n\to\infty}\eta_n=\eta.$$

Рассматриваемой последовательности $\{\xi_k\}$, $k=1,...,n$, соответствует последовательность функций распределения $\{F_k(x)\}$, $k=1,...,n$. Так как во всех теоремах случайные величины, составляющие последовательность, независимые, то функция распределения $F_{\eta_n}(x)$ случайной величины $\eta_n$, являющейся суммой независимых случайных величин, будет свёрткой функций распределения $\{F_k(x)\}$, то есть: $F_{\eta_n}(x)=F_1(x)*F_2(x)*...*F_n(x)$.

Ясно, что при неограниченном увеличении числа слагаемых последовательность $\{F_{\eta_n}(x)\}$ слабо сходится к некоторой функции $F(x)$. Если $F(x)$ - будет функцией распределения, то это будет функция распределения предельной случайной величины $\eta$.

$$\{F_k(x)\},k=1,...,n, \Longrightarrow F_{\eta_n}(x)=F_1(x)*F_2(x)*...*F_n(x) \Longrightarrow \lim_{n\to\infty}F_{\eta_n}=F_\eta(x).$$

Случайные величины $\eta_n$, для упрощения формы записи мы центрируем и нормируем тремя способами и записываем три случайных величины $\tau_n$, $\tau'_n$ и $\tau''_n$.

При неограниченном увеличении числа слагаемых, получается, что предельными случайными величинами могут быть $\tau_N\in\mathcal{N}(0,1)$, или $\tau_0\in\mathrm{E}(0)$, или $\tau_\lambda\in\Pi(\lambda)$.

В предложенной схеме перечисляется последовательность действий с функциями распределения случайных величин – слагаемых для выяснения вида закона распределения предельной случайной величины. Технически наиболее трудными являются действия по определению вида функции распределения $F_{\eta_n}(x)$, являющейся свёрткой функций распределения

$$F_{\eta_n}(x)=F_1(x)*F_2(x)*...*F_n(x).$$

Существенно упрощает технические сложности действий по определению законов распределения вероятностей случайных величин $\eta_n$ и $\eta$ применение обратного преобразования Фурье к функциям распределения $F(x)$ случайных величин. Функции, которые получаются при применении обратного преобразования Фурье к функциям распределения, называются *характеристическими функциями.*

## §5.V Характеристические функции случайных величин

***Комментарий***

Применение характеристических функций и использование их свойств являются «инструментом», существенно облегчающим доказательство предельных теорем. Метод доказательства предельных теорем теории вероятностей с помощью характеристических функций был применён А.М.Ляпуновым и носит название «метод характеристических функций».

## Теория

Пусть $\xi$ некоторая случайная величина и $F(x)$ - её функция распределения.

<u>Определение.</u> ***Характеристической функцией*** *случайной величины* $\xi$ *называется комплекснозначная функция* $\varphi_\xi(t)$ *вещественного аргумента* $t \in R$:

$$\varphi_\xi(t) = M\left[e^{it\xi}\right] = \int_R e^{itx} dF(x).$$

Определяя математическое ожидание функции $g(\xi)$ случайной величины, мы требовали, чтобы интеграл Римана-Стилтьеса сходился абсолютно. Покажем, что характеристическая функция существует у любой случайной величины, то есть несобственный интеграл Римана-Стилтьеса от функции $g(\xi) = e^{it\xi}$ по функции распределения $F(x)$ сходится абсолютно.

Так как $\left|e^{itx}\right| = |\cos tx + i \sin tx| = 1$, то $\int_R \left|e^{itx}\right| dF(x) = \int_R dF(x) = 1$.

Если случайная величина $\xi$ дискретного типа, то $\varphi_\xi(t) = \sum_{k=1}^{n,\infty} e^{itx_k} \cdot p_k$.

Если случайная величина $\xi$ непрерывного типа, то $\varphi_\xi(t) = \int_R e^{itx} \cdot p(x) dx$.

Прежде, чем приступить к изучению свойств характеристических функций, определим вид характеристических функций для некоторых, наиболее часто встречающихся на практике законов распределения вероятностей.

## Практика

<u>Пример 1.</u> Случайная величина $\xi$ - вырожденная в точке ноль, то есть $\xi \in \mathrm{E}(0)$, что означает: $P(\xi = 0) = 1$.

$\varphi_\xi(t) = e^{it0} \cdot 1 \equiv 1$.

<u>Пример 2.</u> Случайная величина $\xi$ распределена по биномиальному закону распределения вероятностей, то есть $\xi \in B_n(p)$:

$$\varphi_\xi(t) = \sum_{k=0}^{n} e^{itk} \cdot C_n^k p^k q^{n-k} = \sum_{k=0}^{n} C_n^k \left(e^{it} p\right)^k q^{n-k} = \left(q + pe^{it}\right)^n.$$

В частности, если $\xi$ - бернуллиевская, то есть $\xi \in B_1(p)$, то $\varphi_\xi(t) = \left(q + pe^{it}\right)$.

<u>Пример 3.</u> Случайная величина $\xi$ распределена равномерно на отрезке $[-c;c]$: $\varphi_\xi(t) = \int_{-c}^{c} e^{itx} \cdot \frac{1}{2c} dx = \frac{e^{itc} - e^{-itc}}{2cit} = \frac{\sin tc}{tc}$.

<u>Пример 4.</u> Случайная величина $\xi$ распределена по нормальному закону с параметрами $m = 0$ и $\sigma = 1$, то есть $\xi \in \mathcal{N}(0,1)$:

$\varphi_\xi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} \cdot e^{-\frac{x^2}{2}} dx$. Применяя формулу Эйлера $e^{itx} = \cos tx + i \sin tx$ и учитывая, что интеграл то нечётной функции в симметричных пределах равен нулю, получаем: $\varphi_\xi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \cdot \cos tx dx$. Полученное выражение продифференцируем по переменной $t$:

$\frac{d\varphi_\xi(t)}{dt}=\frac{-1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}x\cdot e^{-\frac{x^2}{2}}\cdot \sin txdx$. Применим формулу интегрирования по частям:

$\frac{d\varphi_\xi(t)}{dt}=-\frac{-1}{\sqrt{2\pi}}\left[-e^{-\frac{x^2}{2}}\cdot \sin tx\Big|_{-\infty}^{\infty}+t\int\limits_{-\infty}^{\infty}e^{-\frac{x^2}{2}}\cdot \cos txdx\right]$. Так как первое слагаемое равно нулю, то приходим к дифференциальному уравнению: $\frac{d\varphi_\xi(t)}{dt}=-t\cdot\varphi_\xi(t)$, решением которого будет: $ln\varphi_\xi(t)=-\frac{t^2}{2}+lnc$ или $\varphi_\xi(t)=c\cdot e^{-\frac{t^2}{2}}$. Полагая $t$=0, определяем значение постоянной $c=1$.

Следовательно, у случайной величины $\xi$, распределённой по нормальному закону $\mathcal{N}(0,1)$, характеристическая функция имеет вид $\varphi_\xi(t)=e^{-\frac{t^2}{2}}$.

Пример 5. Случайная величина $\xi$ распределена по закону Пуассона с параметром $\lambda$, то есть $\xi\in\Pi(\lambda)$: $\varphi_\xi(t)=\sum\limits_{k=0}^{\infty}e^{itk}\cdot\frac{\lambda^k e^{-\lambda}}{k!}=e^{-\lambda}\sum\limits_{k=0}\frac{\left(\lambda\cdot e^{it}\right)^k}{k!}=e^{\lambda\left(e^{it}-1\right)}$.

Анализируя процесс определения характеристической функции в рассмотренных примерах, приходим к выводу, что если $F(x)$ - функция распределения случайной величины $\xi$, то, обладая навыками суммирования рядов или вычисления несобственных интегралов, характеристическую функцию этой случайной величины всегда можно найти.

**Теория**

Рассмотрим свойства характеристических функций.

Свойство 1. Для характеристической функция $\varphi_\xi(t)$ любой случайной величины $\xi$ справедливо: $|\varphi_\xi(t)|\le 1$, где $t\in R$, и $\varphi_\xi(0)=1$.

Справедливость этого свойства – очевидна.

Свойство 2. Характеристическая функция $\varphi_\xi(t)$ равномерно непрерывна на всей числовой оси.

Для доказательства этого свойства надо убедиться, что по любому, сколь угодно малому $\varepsilon>0$ можно найти $h>0$, что будет справедливо: $|\varphi_\xi(t+h)-\varphi_\xi(t)|<\varepsilon$ для любого $t\in R$.

Очевидно: $|\varphi_\xi(t+h)-\varphi_\xi(t)|=\left|\int\limits_{-\infty}^{\infty}e^{i(t+h)x}dF(x)-\int\limits_{-\infty}^{\infty}e^{itx}dF(x)\right|\le\int\limits_{-\infty}^{\infty}\left|e^{ihx}-1\right|dF(x)$.

Область интегрирования разобьём на две части:

$$|\varphi_\xi(t+h)-\varphi_\xi(t)|\le\int\limits_{|x|<a}\left|e^{ihx}-1\right|dF(x)+\int\limits_{|x|\ge a}\left|e^{ihx}-1\right|dF(x).$$

По выбранному $\varepsilon>0$ всегда можно подобрать достаточно большое $a$, чтобы выполнялось $\int\limits_{|x|\ge a} dF(x) < \frac{\varepsilon}{3}$. Кроме того, для $|x| \ge a$ и любого $h$ всегда будет справедливо: $\left|e^{ihx} - 1\right| \le \left|e^{ihx}\right| + 1 = 2$.

По выбранному $a$, которое зависит от $\varepsilon > 0$, можно для $|x| < a$ найти достаточно малое $h > 0$ такое, что $\left|e^{ihx} - 1\right| < \frac{\varepsilon}{3}$. А так как всегда $\int\limits_{|x|<a} dF(x) \le 1$, то получаем необходимую оценку: $\left|\varphi_\xi(t+h) - \varphi_\xi(t)\right| < \frac{\varepsilon}{3} \cdot 1 + 2 \cdot \frac{\varepsilon}{3} = \varepsilon$.

<u>Свойство 3.</u> Характеристическая функция обладает свойством эрмитовости, то есть: $\overline{\varphi_\xi(-t)} = \varphi_\xi(t)$.

Доказательство этого свойства можно провести самостоятельно.

<u>Свойство 4.</u> Если случайная величина $\xi$ имеет начальные моменты $\alpha_l$ до $k$-того порядка включительно $(l = 1,2,\ldots,k)$, то характеристическая функция $\varphi_\xi(t)$ имеет непрерывные производные $\varphi_\xi^{(l)}(t)$ до $k$-того порядка включительно и справедливо равенство: $\varphi_\xi^{(l)}(0) = i^l \cdot \alpha_l$ для $l = 1,2,\ldots,k$.

Для доказательства этого свойства продифференцируем формально выражение характеристической функции $l$ раз:

$\varphi_\xi^{(l)}(0) = i^l \cdot \int\limits_{-\infty}^{\infty} e^{itx} \cdot x^l dF(x)$. Дифференцирование будет правомочно, если получающийся несобственный интеграл будет сходиться абсолютно. Такая сходимость интеграла для каждого $l$ вытекает из существования начального момента $\alpha_l$, то есть $\left|\varphi_\xi^{(l)}(0)\right| \le \int\limits_{-\infty}^{\infty} |x|^l dF(x) < \infty$.

Непрерывность каждой производной порядка $l = 1,2,\ldots,k-1$ следует из существования производной следующего порядка. Непрерывность производной $k$ –того порядка доказывается способом аналогичным тому, как это было сделано при доказательстве свойства 2. Справедливость последнего равенства свойства проверяется непосредственной подстановкой $t = 0$:

$$\varphi_\xi^{(l)}(0) = i^l \int\limits_{-\infty}^{\infty} x^l dF(x) = i^l \cdot M\xi^l = i^l \cdot \alpha_l.$$

<u>Свойство 5.</u> Если случайные величины связаны линейной функциональной зависимостью $\eta = a\xi + b$, то для их характеристических функций справедливо:

$$\varphi_\eta(t) = e^{itb} \cdot \varphi_\xi(at).$$

Справедливость этого равенства вытекает из свойств математического ожидания: $\varphi_\eta(t) = M\left[e^{it\eta}\right] = M\left[e^{it(a\xi+b)}\right] = M\left[e^{iat\xi} \cdot e^{itb}\right] = e^{itb} M\left[e^{i(at)\xi}\right] = e^{itb} \cdot \varphi_\xi(at)$.

**Практика**

В примере 3 была определена характеристическая функция случайной величины $\xi$ распределённой равномерно на отрезке $[-c;c]$. Используя свойство

5, мы можем определить характеристическую функцию случайной величины $\eta$ распределённой равномерно на отрезке $[a;b]$.

Линейная зависимость между случайными величинами $\xi$ и $\eta$ имеет вид: $\eta = \frac{b-a}{2c}\xi + \frac{a+b}{2}$. Следовательно, $\varphi_\eta(t) = e^{it\frac{a+b}{2}}\varphi_\xi\left(\frac{b-a}{2c}t\right) = e^{it\frac{a+b}{2}} \cdot \frac{\sin\frac{b-a}{2}t}{\frac{b-a}{2}}$.

В примере 4 была определена характеристическая функция случайной величины $\xi$ распределённой по нормальному закону с параметрами 0 и 1, то есть $\xi \in \mathcal{N}(0,1)$. Если случайная величина $\eta$ распределена по нормальному закону с параметрами $m$ и $\sigma$, то есть $\eta \in \mathcal{N}(m,\sigma)$, то эти случайные величины связаны линейной зависимостью $\eta = \sigma\xi + m$. Следовательно, $\varphi_\eta(t) = e^{itm} \cdot \varphi_\xi(\sigma t) = e^{itm - \frac{\sigma^2 t^2}{2}}$.

**Теория**

Свойство 6. Если случайные величины $\xi_1$ и $\xi_2$ независимые, то характеристическая функция их суммы равна произведению их характеристических функций.

Если случайные величины $\xi_1$ и $\xi_2$ независимые, то независимыми будут и функции этих случайных величин: $e^{it\xi_1}$ и $e^{it\xi_2}$, а так как математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий, то:

$$\varphi_{\xi_1+\xi_2}(t) = M\left[e^{it(\xi_1+\xi_2)}\right] = M\left[e^{it\xi_1} \cdot e^{it\xi_2}\right] = M\left[e^{it\xi_1}\right] \cdot M\left[e^{it\xi_2}\right] = \varphi_{\xi_1}(t) \cdot \varphi_{\xi_2}(t).$$

Следствие. Если случайные величины $\{\xi_k\}$, $k = 1,\dots,n$, - независимые, то характеристическая функция их суммы равна произведению их характеристических функций: $\varphi_{\sum_{k=1}^{n}\xi_k}(t) = \prod_{k=1}^{n}\varphi_{\xi_k}(t)$.

***Комментарий***

Это свойство характеристических функций подсказывает нам, как можно использовать характеристические функции при доказательстве теорем классической предельной проблемы. Функция распределения суммы независимых случайных величин равна свёртке функций распределения случайных величин-слагаемых: $F_{\eta_n}(x) = F_1(x) * F_2(x) * \dots * F_n(x)$. Процесс вычисления свёртки большого числа функций распределения с использованием ассоциативности операции свёртки – весьма трудоёмкий и длительный. Значительно легче найти характеристическую функцию случайной величины $\eta_n$ как произведение характеристических функций слагаемых: $\varphi_{\eta_n}(t) = \prod_{k=1}^{n}\varphi_k(t)$, а потом по характеристической функции $\varphi_{\eta_n}(t)$ определить функцию распределения $F_{\eta_n}(x)$.

Свойство 6 позволяет проверить свойство *устойчивости* некоторых законов распределения вероятностей. Если случайные величины $\xi_1$ и $\xi_2$ независимые и распределены по нормальному закону $\xi_1 \in \mathcal{N}(m_1, \sigma_1)$ и $\xi_2 \in \mathcal{N}(m_2, \sigma_2)$, то случайная величина $\xi_1 + \xi_2$ будет также распределена по нормальному закону $\xi_1 + \xi_2 \in \mathcal{N}(m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2})$, так как $\varphi_{\xi_1+\xi_2}(t) = \varphi_{\xi_1}(t) \cdot \varphi_{\xi_2}(t) = e^{it(m_1+m_2) - \frac{(\sigma_1^2+\sigma_2^2)t^2}{2}}$.

Легко проверить, что свойством устойчивости обладает закон Пуассона.

У случайных величин $\xi_1$ и $\xi_2$, которые имеют законом распределения вероятностей закон Пуассона ($\xi_i \in \Pi(\lambda_i), i = 1,2$), характеристические функции равны $\varphi_{\xi_i}(t) = e^{\lambda_i(e^{it}-1)}$. Если они независимые, характеристическая функция их суммы $\xi_1 + \xi_2$ будет равна $\varphi_{\xi_1+\xi_2}(t) = \varphi_{\xi_1}(t) \cdot \varphi_{\xi_2}(t) = e^{(\lambda_1+\lambda_2)(e^{it}-1)}$. Значит $\xi_1 + \xi_2 \in \Pi(\lambda_1 + \lambda_2)$.

То есть в обоих случаях вид закона распределения при суммировании случайных величин остался тем же («устоял»).

Пусть случайные величины $\xi_1$ и $\xi_2$ распределены равномерно на отрезке $[-1;1]$. Их характеристические функции равны $\varphi_{\xi_1}(t) = \varphi_{\xi_2}(t) = \frac{\sin t}{t}$. В том случае, когда $\xi_1$ и $\xi_2$ независимые, характеристическая функция их суммы $\xi_1 + \xi_2$ будет равна $\varphi_{\xi_1+\xi_2}(t) = \varphi_{\xi_1}(t) \cdot \varphi_{\xi_2}(t) = \frac{\sin^2 t}{t^2}$. Мы получили характеристическую функцию случайной величины, распределение вероятностей которой называется законом Симпсона: $p(x) = \begin{cases} 0, & если \quad x \notin [-2;2] \\ \frac{1}{2} - \frac{1}{4}|x|, & если \quad x \in [-2;2] \end{cases}$. Это распределение вероятностей не будет равномерным. Значит, равномерное распределение вероятностей не будет устойчивым распределением.

Анализируя вид получающегося распределения вероятностей суммы независимых случайных величин, мы называли вид закона распределения вероятностей по виду характеристической функции. То есть мы предполагали, что если по функции распределения всегда можно единственным образом определить характеристическую функцию, то справедливо и обратное: по характеристической функции всегда можно единственным образом определить функцию распределения. Справедливость такого предположения подтверждается свойством 7 характеристических функций.

**Теория**

Свойство 7. Если $x_1$ и $x_2$ - точки непрерывности функции распределения $F(x)$, то имеет место *формула обращения*:

$$F(x_2) - F(x_1) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx_1} - e^{-itx_2}}{it} \cdot \varphi(t) dt.$$

В частности, если $\varphi(t)$ интегрируема на числовой оси, то есть если $\varphi(t) \in L_1(-\infty;\infty)$, то случайная величина $\xi$ будет случайной величиной непрерывного типа. Её плотность вероятности $p(x)$ - ограниченная функция и $p(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{-itx} \cdot \varphi(t)dt$.

Мы видим, что если характеристическая функция является *обратным преобразованием* Фурье плотности вероятности, то плотность вероятности будет *преобразованием* Фурье характеристической функции.

Если случайная величина $\xi$ дискретного типа и её значениями будут $x_k = a + kh$, где $a, h \in R$ и $k \in Z$, то набор $\{p_k\}$, $k = 1,2,..., n ..\infty$ называется *решётчатым* распределением вероятностей. В этом случае будет справедливо:

$$p_k = P(\xi = x_k) = \frac{h}{2\pi}\int\limits_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-itx_k} \cdot \varphi(t)dt.$$

Свойство 8. Теорема непрерывности.

Пусть $\{F_n(x)\}$ и $\{\varphi_n(t)\}$ последовательности соответствующих функций распределения и характеристических функций.

1. Если $F_n(x) \Rightarrow F(x)$ при $n \to \infty$, где $F(x)$ - функция распределения, то $\lim\limits_{n\to\infty} \varphi_n(t) = \varphi(t)$ для каждого $t \in R$, где $\varphi(t)$ - характеристическая функция, соответствующая функции распределения $F(x)$.

2. Если $\lim\limits_{n\to\infty} \varphi_n(t) = \varphi(t)$ при каждом $t \in R$ и $\varphi(t)$ - непрерывная функция в точке $t = 0$, то $F_n(x) \Rightarrow F(x)$ при $n \to \infty$, где $F(x)$ - функция распределения, соответствующая характеристической функции $\varphi(t)$.

***Комментарий***

Требование, чтобы предельная функция $F(x)$ была функцией распределения обязательно, что иллюстрирует следующий пример. Рассмотрим последовательность $\{F_n(x)\}$, $n = 1,2,...$, где $F_n(x) = \begin{cases} 0, & \textit{если} \quad x \le -n \\ \frac{1}{2}, & \textit{если} \quad -n < x \le n \\ 1, & \textit{если} \quad x > n \end{cases}$. Ясно, что для каждого $n$ функция $F_n(x)$ будет функцией распределения. Ясно, что $F_n(x) \Rightarrow F(x)$, где $F(x) \equiv \frac{1}{2}$ для $-\infty < x < \infty$, но функция $F(x) \equiv \frac{1}{2}$ функцией распределения не будет.

Свойство 8, называемое «Теорема непрерывности», позволяет при доказательстве теорем классической предельной проблемы «обойти» трудность, возникающую при определении функции распределения $F_{\eta_n}(x)$, являющейся свёрткой функций распределения $F_{\eta_n}(x) = F_1(x) * F_2(x) * ... * F_n(x)$.

Функциям распределения $\{F_k(x)\}$ случайных величин-слагаемых $\{\xi_k\}$,

$k=1,...,n$, однозначно соответствуют характеристические функции $\{\varphi_k(t)\}$. Каждой функции $F_{\eta_n}(x)$ соответствует характеристическая функция $\varphi_{\eta_n}(t)$, которая будет равна произведению характеристических функций $\{\varphi_k(t)\}$. При $n\to\infty$ последовательность $\{F_{\eta_n}(x)\}$ слабо сходится к некоторой функции $F(x)$, а последовательность $\{\varphi_{\eta_n}\}$ сходится к некоторой функции $\varphi(t)$, которая, согласно свойству 8, будет характеристической функцией соответствующей функции распределения $F(x)$ предельной случайной величины $\eta$. Схематически пути рассуждений при доказательстве предельных теорем можно представить так:

$$\begin{tikzcd}
\{F_k(x)\},\ k=1,...,n, \arrow[r, Rightarrow] \arrow[d] & F_1(x)*F_2(x)*...*F_n(x)=F_{\eta_n}(x) \arrow[r, Rightarrow] & \lim\limits_{n\to\infty} F_{\eta_n}=F_\eta(x). \\
\{\varphi_k(t)\},\ k=1,...,n, \arrow[r] & \varphi_1(t)\cdot\varphi_2(t)\cdot...\cdot\varphi_n(t)=\varphi_{\eta_n}(t) \arrow[r] & \lim\limits_{n\to\infty} \varphi_{\eta_n}=\varphi_\eta(t). \arrow[u]
\end{tikzcd}$$

Путь, отмеченный голубыми стрелками, будет более простым, а потому – более предпочтительным.

## §6.V Закон больших чисел

***Комментарий***

Теорема Я.Бернулли – первая теорема из Закона больших чисел, к которой мы пришли, рассматривая последовательность независимых одинаково распределённых бернуллиевских случайных величин, то есть $\xi_k\in B_n(p)$, $k=1,...,n$. Ослабевая требования, налагаемые члены последовательностей $\{\xi_k\}$, $k=1,...,n$, рассмотрим теоремы А.Я.Хинчина, П.Л.Чебышёва и Пуассона.

**Теория**

Теорема А.Я.Хинчина (1935г.)

*Пусть* $\{\xi_k\}$, $k=1,...,n$, *последовательность независимых, одинаково распределённых случайных величин, имеющих конечное математическое ожидание, то есть* $M\xi_k=m<\infty$. *Тогда:*

*1)* $\lim\limits_{n\to\infty} P\left(\left|\frac{1}{n}\sum\limits_{k=1}^{n}\xi_k-m\right|<\varepsilon\right)=1$;

*2)* $\tau'_n \xrightarrow{P} \tau_0$

*3) последовательность* $\{\xi_k\}$ *подчиняется Закону больших чисел.*

***Комментарий***

Все три формулировки теоремы – равнозначные утверждения, заявляющие, что при большом числе случайных величин-слагаемых среднее арифметическое их значений с вероятностью близкой к единице, то есть практически достоверно, будет очень мало отличаться от значения математического ожидания $m$.

Среднее арифметическое случайных величин $\frac{1}{n}\sum\limits_{k=1}^{n}\xi_k=\bar{\xi}$ - случайная величина, а значение $m$ математического ожидания каждой случайной величины $\xi_k$ и среднего арифметического этих случайных величин $\bar{\xi}$ равно $m$. Число $m$ –

постоянная величина. Теорема А.Я.Хинчина утверждает, что при больших значениях $n$ случайная величина $\bar{\xi}$ «ведёт себя» как постоянная, а случайная величина $\tau_n' = \bar{\xi} - m$ принимает значения, мало отличающиеся от нуля.

**Теория**

Доказательство теоремы.

Так как все случайные величины $\{\xi_k\}$, $k = 1,...,n$, одинаково распределены, то есть у них - одинаковые функции распределения $F(x)$ и одинаковые характеристические функции $\varphi(t)$. Так как случайные величины последовательности $\{\xi_k\}$ - независимые, то характеристическая функция их суммы равна произведению их характеристических функций (свойство 6): $\varphi_{\sum\limits_{k=1}^{n}\xi_k}(t) = [\varphi(t)]^n$. Используя свойство 5, запишем характеристическую функцию случайной величины $\tau_n' = \bar{\xi} - m$: $\varphi_{\tau_n'}(t) = \left[\varphi\left(\dfrac{t}{n}\right)\right]^n \cdot e^{-imt}$.

Выпишем первые два члена разложения функции $\varphi\left(\dfrac{t}{n}\right)$ в окрестности нуля в ряд Тейлора: $\varphi\left(\dfrac{t}{n}\right) = \varphi(0) + \dfrac{\varphi'(0)}{1!}\cdot\dfrac{t}{n} + o\left(\dfrac{t}{n}\right)$. Так как (свойства 1 и 4): $\varphi(0) = 1$ и $\varphi'(0) = i\cdot\alpha_1$, где $\alpha_1 = m$, то $\varphi_{\tau_n'}(t) = \left[1 + \dfrac{i\cdot m}{1!}\cdot\dfrac{t}{n} + o\left(\dfrac{t}{n}\right)\right]^n \cdot e^{-imt}$. Вычислим теперь, используя второй замечательный предел, предел последовательности характеристических функций $\{\varphi_{\tau_n'}(t)\}$ случайных величин $\tau_n'$, $n = 1,2,...$:

$$\lim_{n\to\infty}\varphi_{\tau_n'}(t) = \lim_{n\to\infty}\left[1 + \frac{i\cdot m\cdot t}{n} + o\left(\frac{t}{n}\right)\right]^n \cdot e^{-imt} = e^{imt}\cdot e^{-imt} = e^0 \equiv 1.$$

Получили: $\lim\limits_{n\to\infty}\varphi_{\tau_n'}(t) = \varphi(t) \equiv 1$. Функция $\varphi(t) \equiv 1$ является характеристической функцией вырожденного в нуле закона $\mathbf{E}(0)$. Следовательно, согласно свойству 8, $\tau_n' \xrightarrow{w} \tau_0$, где $P(\tau_0 = 0) = 1$. А так как случайная величина $\tau_0$ - вырожденная, то $\tau_n' \xrightarrow{P} \tau_0$. Что и требовалось доказать.

Следствие. Условия теоремы Бернулли являются частным случаем условий теоремы Хинчина, а потому справедливость теоремы Бернулли вытекает из теоремы Хинчина.

***Комментарий***

Теорема А.Я.Хинчина обосновывает «правило среднего арифметического», которое применяется при проведении измерений любого вида. Пусть интересующий нас «объект» имеет меру (длину, площадь, объём, вес, температуру плавления или замерзания, количество литров бензина, расходуемого на 100 км пути, вес батона, выпекаемого на хлебозаводе, объём минеральной воды в пластиковой бутылке и т.п.). Величина этой меры есть некоторое число $a$, которое мы желаем знать. Мы проводим $n$ измерений, наблюдений этого количественного признака, то есть мы получаем последовательность $\{\xi_k\}, k = 1,2,\ldots,n$ независимых одинаково распределённых случайных величин. Результат каждого измерения, или наблюдения, $x_k$ является значением случайной величины $\xi_k$. Интересующее нас значение $a$ является значением математического ожидания $M\xi_k = a$ каждой случайной величины $\xi_k$. Для оценки неизвестного нам значения $a$ мы вычисляем среднее арифметическое результатов $n$ измерений или наблюдений, то есть мы определяем число $\frac{1}{n}\sum_{k=1}^{n} x_k = \bar{x}$. Полученное число $\bar{x}$ мы принимаем в качестве приближённого значения $a$.

Согласно теореме Хинчина, полученное значение среднего арифметического $\bar{x}$ мало отличается от величины $a$, и чем больше будет сделано измерений или наблюдений, тем больше у нас будет уверенность в справедливости результата $\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - a\right| < \varepsilon$.

## Теория

<u>Теорема П.Л.Чебышёва (1867-1868г.)</u>

*Пусть* $\{\xi_k\}$, $k = 1,\ldots,n$, *последовательность независимых, разно распределённых случайных величин, дисперсии которых ограничены общей константой, то есть* $D\xi_k = \sigma_k^2 \le c$. *Тогда:*

*1)* $\lim\limits_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}M\xi_k\right| < \varepsilon\right) = 1$;

*2)* $\tau_n' \xrightarrow{P} \tau_0$

*3) последовательность* $\{\xi_k\}$ *подчиняется Закону больших чисел.*

### *Комментарий*

Как и в предыдущем случае все три утверждения – равнозначные утверждения, заявляющие, что при большом числе случайных величин-слагаемых среднее арифметическое их значений с вероятностью близкой к единице, то есть практически достоверно, будет очень мало отличаться от среднего арифметического значения их математических ожиданий.

Среднее арифметическое случайных величин $\frac{1}{n}\sum_{k=1}^{n}\xi_k = \bar{\xi}$ - случайная величина, а среднее арифметическое математических ожиданий $\frac{1}{n}\sum_{k=1}^{n}M\xi_k = M\bar{\xi}$ – постоянная величина. Теорема П.Л. Чебышёва утверждает, что при больших значениях $n$ случайная величина $\tau_n' = \bar{\xi} - M\bar{\xi}$ с вероятностью близкой к единице принимает значения, мало отличающиеся от нуля. Доказательство теоремы основано на известном из курса математического анализа неравенстве П.Л.Чебышёва, которое мы рассмотрим в виде леммы.

## Теория

Лемма. Неравенство П.Л.Чебышёва

*Пусть случайная величина $\xi$ принимает неотрицательные значения и $M\xi$ - её математическое ожидание. Тогда для любого положительного числа $a$ будет справедливо неравенство:* $P(\xi \geq a) \leq \dfrac{M\xi}{a}$.

Доказательство.

Так как значения $\xi \geq 0$, то $M\xi = \int\limits_0^\infty x dF(x)$. Учитывая вид области интегрирования, запишем $M\xi \geq \int\limits_a^\infty x dF(x) \geq a \cdot \int\limits_a^\infty dF(x)$, где $a > 0$. Значение интеграла $\int\limits_a^\infty dF(x)$ есть вероятность наступления случайного события $\{\xi \geq a\}$, следовательно $M\xi \geq a \cdot P(\xi \geq a)$, что и требовалось доказать.

Сделаем несколько трансформаций неравенства П.Л.Чебышёва:

Для любой случайной величины $\xi$, имеющей математическое ожидание $M\xi$, будет справедливо $P(|\xi| \geq a) \leq \dfrac{M|\xi|}{a}$. Случайные события $\{|\xi - M\xi| \geq a\}$ и $\{[\xi - M\xi]^2 \geq a^2\}$ - равновероятны, значит:

$$P(|\xi - M\xi| \geq a) = P([\xi - M\xi] \geq a^2) \leq \frac{M[\xi - M\xi]^2}{a^2} = \frac{D\xi}{a^2}.$$

Отсюда получаем неравенство $P(|\xi - M\xi| \geq a) \leq \dfrac{D\xi}{a^2}$, которое применим при доказательстве теоремы. Это неравенство иногда называют «вторым неравенством П.Л.Чебышёва».

Приступим теперь к доказательству сформулированной выше теоремы П.Л.Чебышёва.

Нам надо доказать, что будет справедливо $\lim\limits_{n\to\infty} P\left(\left|\dfrac{1}{n}\sum\limits_{k=1}^n \xi_k - \dfrac{1}{n}\sum\limits_{k=1}^n M\xi_k\right| < \varepsilon\right) = 1$, если последовательность случайных величин $\{\xi_k\}$, $k = 1,...,n$, удовлетворяет условиям теоремы.

Обозначим $\dfrac{1}{n}\sum\limits_{k=1}^n \xi_k = \eta_n$, тогда $\dfrac{1}{n}\sum\limits_{k=1}^n M\xi_k = M\eta_k$ и рассмотрим: $P(|\eta_n - M\eta_n| < \varepsilon) = 1 - P(|\eta_n - M\eta_n| \geq \varepsilon)$. Согласно неравенству П.Л.Чебышёва $P(|\eta_n - M\eta_n| \geq \varepsilon) \leq \dfrac{D\eta_n}{\varepsilon^2}$, значит $P(|\eta_n - M\eta_n| < \varepsilon) \geq 1 - \dfrac{D\eta_n}{\varepsilon^2}$.

Оценим, учитывая условия теоремы, величину $D\eta_n$. Так как случайные величины - члены последовательности $\{\xi_k\}$ независимые, а значения их дисперсий не превосходят постоянного числа $\boldsymbol{c}$, то $\boldsymbol{D\eta_n} = \frac{1}{\boldsymbol{n}^2}\sum_{\boldsymbol{k}=1}^{\boldsymbol{n}} \boldsymbol{D\xi_k} \le \frac{1}{\boldsymbol{n}^2}\sum_{\boldsymbol{k}=1}^{\boldsymbol{n}} \boldsymbol{c} = \frac{\boldsymbol{c}}{\boldsymbol{n}}$.

Следовательно, $\boldsymbol{P}\left(\left|\boldsymbol{\eta_n} - \boldsymbol{M\eta_n}\right| < \boldsymbol{\varepsilon}\right) \ge 1 - \frac{\boldsymbol{c}}{\boldsymbol{n\varepsilon}^2}$. Если $\boldsymbol{n} \to \infty$, то, так как вероятность не может быть больше единицы, получаем требуемое равенство:

$$\lim_{\boldsymbol{n}\to\infty} \boldsymbol{P}\left(\left|\frac{1}{\boldsymbol{n}}\sum_{\boldsymbol{k}=1}^{\boldsymbol{n}} \boldsymbol{\xi_k} - \frac{1}{\boldsymbol{n}}\sum_{\boldsymbol{k}=1}^{\boldsymbol{n}} \boldsymbol{M\xi_k}\right| < \boldsymbol{\varepsilon}\right) = 1.$$

Замечание. Теорема А.Я.Хинчина не является частным случаем теоремы П.Л.Чебышёва, так как существование конечного математического ожидания $\boldsymbol{M\xi_k} = \boldsymbol{m}$ не означает, что у случайных величин $\boldsymbol{\xi_k}, \boldsymbol{k} = 1,2,\dots,\boldsymbol{n}$, существуют конечные дисперсии, ограниченные общей постоянной $\boldsymbol{c}$.

У случайных величин $\boldsymbol{\xi_k}, \boldsymbol{k} = 1,2,\dots,\boldsymbol{n}$, вообще могут не существовать конечные дисперсии. И тогда эта последовательность случайных величин $\{\boldsymbol{\xi_k}\}$, $\boldsymbol{k} = 1,2,\dots,\boldsymbol{n}$, не входит во множество последовательностей случайных величин, рассматриваемых в классической предельной проблеме.

Теорема Пуассона (1837г.)

*Если $\{\boldsymbol{\xi_k}\}$, $\boldsymbol{k} = 1,2,\dots,\boldsymbol{n}$, - последовательность независимых разно распределённых бернуллиевских случайных величин: $\boldsymbol{\xi_k} \in \boldsymbol{B}_1(\boldsymbol{p_k})$, то*

*1)* $\lim_{\boldsymbol{n}\to\infty} \boldsymbol{P}\left(\left|\frac{1}{\boldsymbol{n}}\sum_{\boldsymbol{k}=1}^{\boldsymbol{n}} \boldsymbol{\xi_k} - \frac{1}{\boldsymbol{n}}\sum_{\boldsymbol{k}=1}^{\boldsymbol{n}} \boldsymbol{p_k}\right| < \boldsymbol{\varepsilon}\right) = 1$*;*

*2)* $\boldsymbol{\tau}'_{\boldsymbol{n}} \xrightarrow{\boldsymbol{P}} \boldsymbol{\tau}_0$

*3) последовательность* $\{\boldsymbol{\xi_k}\}$ *подчиняется Закону больших чисел.*

Ясно, что $\boldsymbol{M\xi_k} = \boldsymbol{p_k}$ и $\boldsymbol{D\xi_k} = \boldsymbol{p_k} \cdot \boldsymbol{q_k}$. Для любой вероятности $\boldsymbol{p_k}$ будет справедливо: $\boldsymbol{p_k} \cdot \boldsymbol{q_k} \le \frac{1}{4}$. Справедливость теоремы Пуассона следует из теоремы П.Л.Чебышёва, так все условия этой теоремы здесь выполнены.

## §7.V Центральная предельная теорема

### *Комментарий*

К понятию «Центральная предельная теорема» мы пришли, рассматривая случайные величины $\boldsymbol{\xi} \in \boldsymbol{B_n}(\boldsymbol{p})$, которые являются суммами независимых бернуллиевских случайных величин $\boldsymbol{\xi} = \sum_{\boldsymbol{k}=1}^{\boldsymbol{n}} \boldsymbol{\xi_k}$. Локальная и интегральная теоремы Муавра-Лапласа позволяют упростить вычисления вероятностей событий $\{\boldsymbol{\xi} = \boldsymbol{k}\}$ и $\{\boldsymbol{m}_1 \le \boldsymbol{\xi} \le \boldsymbol{m}_2\}$ при больших значениях $\boldsymbol{n}$. Суть этих теорем состоит в том, что при $\boldsymbol{n} \to \infty$ биномиальное распределение вероятностей аппроксимируется нормальным распределением. Ослабевая требования, налагаемые на члены последовательности $\{\boldsymbol{\xi_k}\}$, $\boldsymbol{k} = 1,\dots,\boldsymbol{n}$, рассмотрим теоремы П.Леви-А.Я.Хинчина, А.М.Ляпунова, Линдеберга-Феллера, в которых доказывается сходимость по распределению последовательности центрированных и нормированных сумм $\{\tau_n\}$, $n = 1,2,...$, где $\tau_n = \dfrac{\sum\limits_{k=1}^{n}\xi_k - \sum\limits_{k=1}^{n}M\xi_k}{\sqrt{\sum\limits_{k=1}^{n}D\xi_k}}$, к случайной величине $\tau_N$, где $\tau_N \in \mathcal{N}(0;1)$.

В частности, если все $\xi_k$ - независимые, имеют одинаковое распределение вероятностей и $M\xi_k = m$, $D\xi_k = \sigma^2$, то $\tau_n = \dfrac{\sum\limits_{k=1}^{n}\xi_k - n\cdot m}{\sigma\sqrt{n}}$.

**Теория**

Теорема П.Леви-А.Я.Хинчина (1935г.)

*Пусть* $\{\xi_k\}$, $k = 1,...,n$, *последовательность независимых, одинаково распределённых случайных величин, имеющих конечный начальный момент второго порядка, то есть* $M\xi_k^2 = \alpha_2 < \infty$. *Тогда:*

*1)* $\lim\limits_{n\to\infty} P\left(\dfrac{\sum\limits_{k=1}^{n}\xi_k - n\cdot m}{\sigma\sqrt{n}} < x\right) = F_0(x) = \dfrac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{z^2}{2}}dz$;

*2)* $\tau_n \xrightarrow{w} \tau_N$ *, где* $\tau_N \in \mathcal{N}(0;1)$;

*3) последовательность* $\{\xi_k\}$ *подчиняется Центральной предельной теореме.*

Доказательство. Если у случайных величин $\{\xi_k\}$, $k = 1,...,n$, существует конечный начальный момент второго порядка $M\xi_k^2 = \alpha_2$, то у них существует конечная дисперсия: $D\xi_k = \alpha_2 - m^2 = \sigma^2$.

Из последовательности $\{\xi_k\}$, $k = 1,...,n$, образуем последовательность центрированных случайных величин $\left\{\overset{o}{\xi}_k\right\}$, $k = 1,...,n$, где $\overset{o}{\xi}_k = \xi_k - m$. Так как все $\{\xi_k\}$, $k = 1,...,n$, имеют одинаковое распределение вероятностей, то $M\overset{o}{\xi}_k = \overset{o}{\alpha}_1 = 0$, $D\overset{o}{\xi}_k = D\xi_k = \sigma^2 = \overset{o}{\alpha}_2$ и $\varphi_{\overset{o}{\xi}}(t) = e^{-imt}\cdot\varphi_{\xi_k}(t)$.

Случайную величину $\tau_n$ запишем теперь так: $\tau_n = \dfrac{\sum\limits_{k=1}^{n}\overset{o}{\xi}_k}{\sigma\sqrt{n}}$.

Характеристическая функция случайной величины $\tau_n$, согласно свойствам 5 и 6 характеристических функций, будет иметь вид: $\varphi_{\tau_n}(t) = \left[\varphi_{\overset{o}{\xi}}\left(\dfrac{t}{\sigma\sqrt{n}}\right)\right]^n$.

Запишем первые три члена разложения функции $\varphi_{\overset{o}{\xi}}\left(\dfrac{t}{\sigma\sqrt{n}}\right)$ в окрестности нуля в ряд Тейлора: $\varphi_{\overset{o}{\xi}}\left(\dfrac{t}{\sigma\sqrt{n}}\right)=\varphi_{\overset{o}{\xi}}(0)+\dfrac{\varphi'_{\overset{o}{\xi}}(0)}{1!}\cdot\dfrac{t}{\sigma\sqrt{n}}+\dfrac{\varphi''_{\overset{o}{\xi}}(0)}{2!}\cdot\dfrac{t^2}{\sigma^2 n}+o\left(\dfrac{t^2}{\sigma^2 n}\right)$.

Так как (свойства 1 и 4): $\varphi_{\overset{o}{\xi}}(0)=1$, $\varphi'_{\overset{o}{\xi}}(0)=i\cdot\overset{o}{\alpha}_1=0$ и $\varphi''_{\overset{o}{\xi}}(0)=i^2\cdot\overset{o}{\alpha}_2=-\sigma^2$, то $\varphi_{\overset{o}{\xi}}\left(\dfrac{t}{\sigma\sqrt{n}}\right)=1+0+\dfrac{-\sigma^2}{2}\cdot\dfrac{t^2}{\sigma^2 n}+o\left(\dfrac{t^2}{\sigma^2 n}\right)$. Значит: $\varphi_{\tau_n}(t)=\left[1-\dfrac{t^2}{2n}+o\left(\dfrac{t^2}{2n}\right)\right]^n$.

Используя второй замечательный предел, вычислим предел последовательности характеристических функций $\{\varphi_{\tau_n}(t)\}$ при $n\to\infty$.

$$\lim_{n\to\infty}\varphi_{\tau_n}(t)=\lim_{n\to\infty}\left[1-\frac{t^2}{2n}+o\left(\frac{t^2}{2n}\right)\right]^n=e^{-\frac{t^2}{2}}.\quad \text{Получили:}\quad \lim_{n\to\infty}\varphi_{\tau_n}(t)=\varphi(t)=e^{-\frac{t^2}{2}}.$$

Функция $\varphi(t)=e^{-\frac{t^2}{2}}$ является характеристической функцией нормального закона $\mathcal{N}(0;1)$. Следовательно, согласно свойству 8, $\tau'_n\xrightarrow{w}\tau_N$, где $\tau_N\in\mathcal{N}(0;1)$. Что и требовалось доказать.

<u>Следствие 1.</u> *Справедливость интегральной теоремы Муавра-Лапласа следует из теоремы П.Леви-А.Я.Хинчина.*

Действительно, все случайные величины $\{\xi_k\}$, $k=1,\dots,n$, рассматриваемые в этой теореме, - независимые, одинаково распределены: $\xi_k\in B_1(p)$, начальный момент второго порядка равен: $\alpha_2=M\xi_k^2=p$, где $0<p<1$. Все условия теоремы П.Леви-А.Я.Хинчина выполнены.

<u>Следствие 2.</u> Случайную величину $\tau_n$ можно представить так:

$\tau_n=\dfrac{\sum\limits_{k=1}^{n}\xi_k-n\cdot m}{\sigma\sqrt{n}}:\dfrac{n}{n}=\dfrac{\bar{\xi}-m}{\dfrac{\sigma}{\sqrt{n}}}$. Среднее арифметическое случайных величин $\bar{\xi}=\dfrac{1}{n}\sum\limits_{k=1}^{n}\xi_k$ - случайная величина, её значениями будут средние арифметические значений одинаково распределённых случайных величин $\{\xi_k\}$.

Из теоремы П.Леви-А.Я.Хинчина следует, что каким бы ни был закон распределения вероятностей случайных величин $\{\xi_k\}$, их *среднее арифметическое* $\bar{\xi}$ *- случайная величина, подчиняющаяся нормальному распределению с параметрами* $M\bar{\xi}=m$ *и* $D\bar{\xi}=\dfrac{\sigma^2}{n}$, *то есть* $\bar{\xi}\in\mathcal{N}(m;\dfrac{\sigma}{\sqrt{n}})$.

***Комментарий***

Это свойство среднего арифметического $\bar{\xi}$ случайных величин используется, например, при экспериментальном определении значения среднего времени функционирования простых приборов, приспособлений, составных элементов сложных систем, при экспериментальном определении значения средней урожайности сельскохозяйственных культур, при оценке величины расхода бензина на 100 км пройденного пути рассматриваемой модели автомобиля и т.д.

*а)* Полученное в результате эксперимента, состоящего из большого числа наблюдений $\boldsymbol{n}$, значение $\bar{\xi}$, согласно ЗБЧ, мало отличается от измеряемой величины $\boldsymbol{m}$.

*б)* Так как $\bar{\xi}$ - случайная величина, то мы можем утверждать, что разброс её возможных значений около значения $\boldsymbol{m}$, то есть её дисперсия в $\boldsymbol{n}$ раз меньше дисперсии случайного результата каждого отдельного измерения.

*в)* И, наконец, при любом законе распределения вероятностей случайных величин $\xi_k$, мы можем оценить, применяя функцию Лапласа, вероятность отклонения на заданную величину $\varepsilon$ значения $\bar{\xi}$ от значения $\boldsymbol{m}$:

$$P\left(\left|\bar{\xi} - m\right| < \varepsilon\right) = 2 \cdot \Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right).$$

<u>*Замечание.*</u> Пуассон, а потом и Коши показали, что не всегда среднее арифметическое результатов наблюдений даёт лучший результат по сравнению с результатом отдельного наблюдения. В качестве примера для каждой случайной величины $\xi_k$ рассматривалась плотность вероятности $p(x) = \dfrac{1}{\pi\left(1 + x^2\right)}$, $(-\infty;\infty)$. Характеристическая функция любой случайной величины $\xi_k$ равна: $\varphi_{\xi_k}(t) = e^{-|t|}$. Характеристическая функция среднего арифметического $\bar{\xi}$ этих случайных величин будет равна: $\varphi_{\bar{\xi}}(t) = \left[e^{-\frac{|t|}{n}}\right]^n = e^{-|t|}$. То есть, среднее арифметическое $\bar{\xi}$ имеет то же распределение вероятностей, что и каждое слагаемое $\xi_k$. а потому разброс значений $\bar{\xi}$ около измеряемой величины $\boldsymbol{m}$ будет таким же как и разброс возможных значений любой $\xi_k$. Кстати, у случайной величины $\xi_k$, имеющей плотность вероятности $p(x) = \dfrac{1}{\pi\left(1 + x^2\right)}$, называемую распределением Коши, не существует ни математическое ожидание, ни дисперсия. Случайные величины, имеющие это распределение вероятностей, не входят во множество последовательностей, рассматриваемых в Классической предельной проблеме теории вероятностей.

**Теория**

Центрированную и нормированную сумму $\tau_n$ случайных величин $\{\xi_k\}$,

$\boldsymbol{k}=1,...,\boldsymbol{n}$, можно представить так: $\boldsymbol{\tau_n}=\dfrac{\sum_{k=1}^{n}\boldsymbol{\xi_k}-\sum_{k=1}^{n}\boldsymbol{M\xi_k}}{\sqrt{\sum_{k=1}^{n}\boldsymbol{D\xi_k}}}=\sum_{k=1}^{n}\left(\dfrac{\boldsymbol{\xi_k}-\boldsymbol{M\xi_k}}{\sqrt{\sum_{k=1}^{k}\boldsymbol{D\xi_k}}}\right)=\sum_{k=1}^{n}\boldsymbol{\xi_{nk}}$.

То есть, $\boldsymbol{\tau_n}$ является суммой элементов $\boldsymbol{n}$ - той строки схемы серий $\{\boldsymbol{\xi_{nk}}\}$, $\boldsymbol{k}=1,...,\boldsymbol{n}$, $\boldsymbol{n}=1,2,...$, порождаемой последовательностью независимых случайных величин $\{\boldsymbol{\xi_k}\}$, $\boldsymbol{k}=1,...,\boldsymbol{n}$.

<u>Теорема А.М.Ляпунова (1900-1901г.)</u>

*Пусть* $\{\boldsymbol{\xi_к}\}$, $\boldsymbol{k}=1,2,...,\boldsymbol{n}$, *последовательность независимых разно распределённых случайных величин, имеющих конечные дисперсии. Если выполняется «условие Ляпунова»:*

$$\sum_{k=1}^{n}\boldsymbol{M}\left|\boldsymbol{\xi_{nk}}\right|^{2+\delta}\to 0,\ \textit{при}\ \boldsymbol{n}\to\infty,\ \textit{то}\qquad \textit{(Л)}$$

*1)* $\boldsymbol{\tau_n}\xrightarrow{w}\boldsymbol{\tau_N}$ *, где* $\boldsymbol{\tau_N}\in\mathcal{N}(0;1)$*;*

*2) последовательность* $\{\boldsymbol{\xi_k}\}$ *подчиняется Центральной предельной теореме.*

Не анализируя суть условия Ляпунова, при выполнении которого последовательность $\{\boldsymbol{\xi_k}\}$ подчиняется Центральной предельной теореме, познакомимся с более наглядным с практической точки зрения «условием Линдеберга». Перед этим введём понятие: *асимптотически малые случайные величины.*

Все дисперсии $\boldsymbol{D\xi_k}>0$ и, если величины дисперсий слагаемых $\boldsymbol{\xi_k}$ конечны и примерно одного порядка, то дисперсия $\boldsymbol{D\xi_{nk}}=\dfrac{\boldsymbol{D\xi_k}}{\sum_{k=1}^{n}\boldsymbol{D\xi_k}}$ любой случайной величины $\boldsymbol{\xi_{nk}}$, при больших значениях $\boldsymbol{n}$, будет меньше любого наперёд заданного числа. Если значение дисперсии $\boldsymbol{D\xi_{nk}}$ мало, то с вероятностью близкой к единице практически все значения случайной величины $\boldsymbol{\xi_{nk}}$ сосредоточены около нуля.

<u>Определение.</u> Будем говорить, что случайные величины $\boldsymbol{\xi_{nk}}=\dfrac{\boldsymbol{\xi_k}-\boldsymbol{M\xi_k}}{\sqrt{\sum_{k=1}^{n}\boldsymbol{D\xi_k}}}$, где $\boldsymbol{k}=1,2,...,\boldsymbol{n}$; $\boldsymbol{n}=1,2,...$ *асимптотически малы*, если для любого $\boldsymbol{\varepsilon}>0$ будет выполняться: $\max\limits_{1\le k\le n}\boldsymbol{P}\left(\left|\boldsymbol{\xi_{nk}}\right|\ge\boldsymbol{\varepsilon}\right)\to 0$, при $\boldsymbol{n}\to\infty$.

<u>Теорема Линдеберга (1922г.)</u>

*Пусть* $\{\boldsymbol{\xi_к}\}$, $\boldsymbol{k}=1,2,...,\boldsymbol{n}$, *последовательность независимых случайных величин с конечными начальными моментами второго порядка.*

*Для того чтобы эта последовательность подчинялась Центральной предельной теореме, то есть, чтобы* $\boldsymbol{\tau_n}\xrightarrow{w}\boldsymbol{\tau_N}$*, где* $\boldsymbol{\tau_N}\in\mathcal{N}(0;1)$*,*

<u>*достаточно*</u> *выполнения «условия Линдеберга»:*

$$\text{для любого } \varepsilon > 0, \text{ при } n \to \infty, \quad \sum_{k=1}^{n} \int_{|x| \ge \varepsilon} x^2 dF_{nk}(x) \to 0. \qquad (L)$$

Добавление В.Феллера (1934г.):

*Если случайные величины* $\{\xi_{nk}\}$, *где* $k = 1,2,...,n$; $n = 1,2,...$ *асимптотически малы, то «условие Линдеберга» будет* <u>*необходимым*</u> *для того, чтобы эта последовательность подчинялась Центральной предельной теореме, то есть, чтобы* $\tau_n \xrightarrow{w} \tau_N$, *где* $\tau_N \in \mathcal{N}(0;1)$.

***Комментарий***

Для любого $k = 1,2,...,n$ дисперсия случайной величины $\xi_{nk}$ будет равна $D\xi_{nk} = \int_{-\infty}^{\infty} x^2 dF_{nk}(x)$. Ясно, что $\int_{-\infty}^{\infty} x^2 dF_{nk}(x) = \int_{|x|<\varepsilon} x^2 dF_{nk}(x) + \int_{|x|\ge\varepsilon} x^2 dF_{nk}(x)$ для любого $\varepsilon > 0$. Следовательно, условие *(L)* требует, чтобы при $n \to \infty$ все случайные величины $\xi_k$ были равномерно малы по сравнению с их суммой $\eta_n = \sum_{k=1}^{n} \xi_k$. То есть, все независимые случайные величины $\xi_k$ должны вносить примерно одинаковую, малую долю в формирование значения случайной величины $\eta_n$. В этом случае мы говорим, что последовательность $\{\xi_\kappa\}$, $k = 1,2,...,n$ подчиняется ЦПТ, то есть, их сумма имеет закон распределения вероятностей, мало отличающийся от нормального закона.

**Теория**

<u>Следствие 1.</u> Очевидно:

$$M|\xi_{nk}|^{2+\delta} = \int_{-\infty}^{\infty} |x|^{2+\delta} dF_{nk}(x) \ge \int_{|x|\ge\varepsilon} |x|^{2+\delta} dF_{nk}(x) \ge \varepsilon^{\delta} \int_{|x|\ge\varepsilon} x^2 dF_{nk}(x).$$

Просуммируем левую и правую части этого неравенства по $k$ от 1 до $n$:

*(Л)* $\sum_{k=1}^{n} M|\xi_{nk}|^{2+\delta} \ge \varepsilon^{\delta} \sum_{k=1}^{n} \int_{|x|\ge\varepsilon} x^2 dF_{nk}(x)$. *(L)* Ясно, что если значения $\varepsilon > 0$ и $\delta > 0$ - малы, то $\varepsilon^{\delta} \approx 1$. Следовательно, выполнение «условия Ляпунова» *(Л)* обеспечивает выполнение «условия Линдеберга» *(L)* .

<u>Следствие 2.</u> Пусть последовательность $\{\xi_\kappa\}$, $k = 1,2,...,n$, удовлетворяет всем условиям теоремы Леви-Хинчина. Так как $\xi_{nk} = \dfrac{\xi_k - m}{\sigma\sqrt{n}}$, то в интеграле $\int_{|x|\ge\varepsilon} x^2 dF_{nk}(x)$ сделаем замену переменной интегрирования $x = \dfrac{z - m}{\sigma\sqrt{n}}$ и просуммируем полученный интеграл по $k$ от 1 до $n$:

$$\sum_{k=1}^{n} \int_{\frac{|z-m|}{\sigma\sqrt{n}} \ge \varepsilon} \left( \frac{z-m}{\sigma\sqrt{n}} \right)^2 dF(z) = \frac{1}{\sigma^2} \int_{|z-m| \ge \varepsilon\sigma\sqrt{n}} (z-m)^2 dF(z).$$

Так как при $n \to \infty$ область интегрирования $\{z : |z - m| \ge \varepsilon\sigma\sqrt{n}\} \downarrow \varnothing$, то получаем, что $\int\limits_{|x|\geq\varepsilon} x^2 dF_{nk}(x) \to 0$. Значит $\sum\limits_{k=1}^{n} \int\limits_{|x|\geq\varepsilon} x^2 dF_{nk}(x) \to 0$ при $n \to \infty$.

«Условие Линдеберга» ***(L)*** – выполнено и, следовательно, теорема Леви-Хинчина вытекает из теоремы Линдеберга.

## §8.V Закон малых чисел. Теорема Пуассона

### *Комментарий*

Мы рассматривали в **§3.V** теорему Пуассона, применение которой позволяет аппроксимировать асимметричное биномиальное распределение $\xi \in B_n(p)$, где $n$ – количество независимых испытаний – «велико», а $p$ – вероятность наступления события в единичном испытании – «мала».

Понятия «велико» и «мало» - относительные понятия, зависящие от конкретного вида повторных независимых испытаний и нашей «платы» за наступление события в единичном испытании.

Мы построили «схему серий» бернуллиевских случайных величин $\{\xi_{nk}\}$, где $k = 1,2,\ldots,n$ и $n = 1,2,\ldots$. Здесь $\xi_{nk} \in B_1(p_n)$ для $k = 1,2,\ldots,n$ и $p_n = \dfrac{\lambda}{n}$.

Если $p_n$ вероятность наступления события в $n$-той серии повторных независимых испытаний, то при увеличении номера серии $n$ эта вероятность $p_n$ убывает со скоростью обратно пропорциональной увеличению номера серии $n$.

В этом случае предельным распределением вероятностей случайной величины $\xi \in B_n(p)$ будет распределение вероятностей Пуассона $\Pi(\lambda)$, которое мы назвали «законом редких событий». Докажем теорему Пуассона, применяя метод характеристических функций.

### Теория

Обозначим $\sum\limits_{k=1}^{n} \xi_{nk} = \tau''_n$. Ясно, что $\tau''_n \in B_n(p_n)$.

Теорема Пуассона (1837г.)

*Если $p_n \to 0$, но так, что $\lim\limits_{n\to\infty} p_n \cdot n = \lambda$, то для любого числа $m$ будет справедливо:* $\lim\limits_{n\to\infty} P(\tau''_n = m) = \dfrac{\lambda^m \cdot e^{-\lambda}}{m!}$ *или* $\tau''_n \xrightarrow{w} \tau_\lambda$.

Доказательство.

Характеристическая функция случайной величины $\xi_{nk} \in B_1(p_n)$ имеет вид: $\varphi_{nk}(t) = (q_n + p_n \cdot e^{it})$ для $k = 1,2,\ldots,n$. Так как случайные величины $\{\xi_{nk}\}$, $k = 1,2,\ldots,n$, - независимые, то $\varphi_{\tau''_n}(t) = \left[q_n + p_n \cdot e^{it}\right]^n$. Заменим $q_n = 1 - p_n$ и определим предел последовательности $\{\varphi_{\tau''_n}(t)\}$ при $n \to \infty$, учитывая, что $p_n \approx \dfrac{\lambda}{n}$ для больших значений $n$:

$$\lim_{n\to\infty} \varphi_{\tau''_n}(t) = \lim_{n\to\infty} \left[1 + (e^{it} - 1) \cdot p_n\right]^n = \lim_{n\to\infty} \left[1 + (e^{it} - 1) \cdot \frac{\lambda}{n}\right]^n = e^{\lambda(e^{it}-1)}.$$

Итак, $\lim\limits_{n\to\infty}\varphi_{\tau''_n}(t)=e^{\lambda(e^{it}-1)}=\varphi_\lambda(t)$ - характеристическая функция закона Пуассона. По свойству 8 характеристических функций заключаем, что $F_{\tau''_n}(x)\Rightarrow F_\lambda(x)$, где $F_\lambda(x)$ - функция распределения закона Пуассона с параметром $\lambda$. Значит $\tau''_n \xrightarrow{w} \tau_\lambda$, что и требовалось доказать.

## §9.V Усиленный закон больших чисел

### *Комментарий*

Рассматривая повторные независимые испытания, в которых $n\to\infty$, мы пришли к теореме Бернулли, утверждающей, что относительная частота $\frac{1}{n}\xi$ числа наступлений события $\xi$ сходится по вероятности к вероятности $p$ наступления события в единичном испытании: $\frac{1}{n}\xi \xrightarrow{P} p$. Это означает, что при достаточно большом $n$ вероятность наступления события $\left\{\left|\frac{1}{n}\xi - p\right| < \varepsilon\right\}$ может быть сделана больше, чем $1-\eta$, где $\eta > 0$ сколь угодно мало.

Но при этом можно утверждать, что с увеличением числа испытаний $n$ *будут повторяться моменты*, когда наступает событие $\left\{\left|\frac{1}{n}\xi - p\right| \ge \varepsilon\right\}$. Это означает, что из сходимости по вероятности следует, что траектория изменения относительной частоты $\frac{1}{n}\xi$ будет иногда «выходить» из полосы $(p-\varepsilon, p+\varepsilon)$.

В 1909г. Э.Борель показал, что при $p=\frac{1}{2}$ выполняется более «сильное» утверждение о поведении траектории изменения относительной частоты $\frac{1}{n}\xi$.

<u>Теорема Э.Бореля</u>

*Пусть* $\xi \in B_n\left(\frac{1}{2}\right)$. *Тогда для* $\forall\varepsilon>0$ *и* $\forall\delta>0$ *найдётся число* $n_0$*, что для любого натурального числа* **s** *будет выполняться:*

$$\sup_{n_0<n<n_0+s} P\left(\left|\frac{1}{n}\xi - \frac{1}{2}\right| < \varepsilon\right) > 1-\delta .$$

Утверждение теоремы Э.Бореля означает, что последовательность $\left\{\frac{1}{n}\xi\right\}$ сходится «почти наверное» к вероятности $p=\frac{1}{2}$.

### Теория

<u>Определение.</u> ***Усиленным законом больших чисел*** «**УЗБЧ**» называется совокупность теорем, в которых на последовательность случайных величин $\{\xi_k\}$

, $k=1,...,n$, налагаются условия, при которых их «сумма» $\tau'_n = \frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n}M\xi_k$ сходится почти наверное к вырожденной в нуле случайной величине: $P\left(\lim_{n\to\infty}\tau'_n = \tau_0\right)$, или $\tau'_n \xrightarrow{п.н.} \tau_0$.

В 1919г. Ф.П.Кантелли обобщил результат Э.Бореля на случай $p \neq \frac{1}{2}$. Так как $\xi = \sum_{k=1}^{n}\xi_k$, где $\xi_k \in B_1(p)$, то утверждение теоремы Ф.П.Кантелли можно записать так: $P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\xi_k = p\right) = 1$.

Теорему Ф.П.Кантелли можно считать аналогом теоремы Я.Бернулли, рассмотренной в ЗБЧ.

В ЗБЧ мы постепенно ослабляли требования к законам распределения элементов последовательностей $\{\xi_k\}$, $k=1,...,n$ и рассматривали теоремы А.Я.Хинчина и П.Л.Чебышёва. Аналогичные ослабления требований к законам распределения элементов последовательностей в УЗБЧ рассматриваются в теоремах А.Н.Колмогорова.

Теорема А.Н.Колмогорова (1933г.)

*Пусть* $\{\xi_k\}$, $k=1,2,...,n$; $n=1,2,...$ *- последовательность независимых, одинаково распределённых случайных величин. Тогда для того, чтобы эта последовательность подчинялась УЗБЧ, то есть чтобы* $\tau'_n \xrightarrow{п.н.} \tau_0$, *необходимо и достаточно, чтобы существовало* $M\xi_k = m < \infty$.

Теорема А.Н.Колмогорова (1930г.)

*Пусть* $\{\xi_k\}$, $k=1,...,n$ *- последовательность независимых, разно распределённых случайных величин, имеющих дисперсии* $D\xi_k = \sigma_k^2$ *и последовательность положительных чисел* $\{b_k\}$ *такова, что* $b_k \uparrow \infty$ *и* $\sum_{k=1}^{\infty}\frac{D\xi_k}{b_k^2} < \infty$. *Тогда* $\frac{\sum_{k=1}^{n}\xi_k - \sum_{k=1}^{n}M\xi_k}{b_n} \xrightarrow{п.н.} \tau_0$.

*В частности, если* $\sum_{k=1}^{\infty}\frac{D\xi_k}{k^2} < \infty$, *то* $\tau'_n \xrightarrow{п.н.} \tau_0$.

***Комментарий***

Пусть случайная величина $\xi$ равномерно распределена на отрезке $[0;1]$. Любое значение $\xi$ - число $x$ ($0 \le x \le 1$) можно представить в виде двоичной дроби: $x = \sum_{k=1}^{\infty}\frac{a_k}{2^k}$, где каждый коэффициент $a_k$ с равной возможностью может быть равным 0 или 1. Если $x$ – двоично-рациональное число, то есть - число вида $x = \frac{m}{2^n}$, где $m \in N$, то при $k = n$ будем считать, что $a_k = 1$, а при $k > n$ все

$a_k = 0$. Каждую бесконечную последовательность коэффициентов $\{a_1, a_2, ..., a_k, ....\}$ будем рассматривать как значение последовательности независимых случайных величин $\{\xi_1, \xi_2, ..., \xi_k, ....\}$, где все $\xi_k \in B_1\left(\frac{1}{2}\right)$ и $M\xi_k = \frac{1}{2}$.

Тогда согласно УЗБЧ мы можем утверждать, что $\frac{1}{n}\sum_{k=1}^{n}\xi_k \xrightarrow{п.н.} \frac{1}{2}$. То есть, почти наверное мы можем утверждать, что в произвольной бесконечной двоичной последовательности $\{a_1, a_2, ..., a_k, ...\}$ «половина» коэффициентов равна 0 и «половина» коэффициентов равна 1.

## §10.V Статистическое моделирование. Метод Монте-Карло

***Комментарий***

Рассмотрим примеры использования сходимости «почти наверное» последовательности сумм $\{\tau'_n\}$, $n = 1,2,...$ независимых случайных величин $\{\xi_k\}$, $k = 1,2,..., n$, в математической статистике.

Пусть $\xi$ некоторая случайная величина и $F(x)$ её функция распределения. Надо получить набор независимых случайных величин $\xi_1, \xi_2, ..., \xi_n$, имеющих то же распределение вероятностей, что и $\xi$. Значения этих случайных величин – набор чисел $(x_1, x_2, ..., x_n)$ будет возможными значениями случайной величины $\xi$, которые можно получить, наблюдая в эксперименте $n$ раз интересующий нас количественный признак – случайную величину $\xi$.

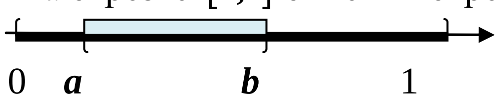
Значения функции распределения $F(x)$ мы можем рассматривать как значения случайной величины $\eta$, имеющей непрерывное равномерное распределение на отрезке $[0;1]$. Случайная величина $\xi = F^{-1}(\eta)$ будет *функцией* случайной величины $\eta$.

С помощью датчика случайных чисел мы получаем набор значений $(y_1, y_2, ..., y_n)$ случайной величины $\eta$. По этим числам, по формуле $x = F^{-1}(y)$, мы определяем набор чисел $(x_1, x_2, ..., x_n)$, являющиеся значениями независимых одинаково распределённых, как и наблюдаемая случайная величина $\xi$, случайных величин $\xi_1, \xi_2, ..., \xi_n$.

Метод вероятностно-статистического моделирования, называемым методом Монте-Карло, применяется при моделировании функционирования сложных систем, или при моделировании случайных процессов, в которых контролируются значения выходной характеристики, которая является случайной величиной.

1. Пусть $A$ – случайное событие и $P(A) = p$ - вероятность наступления этого события в единичном испытании.

На отрезке $[0;1]$ отметим отрезок $[a;b]$, длина которого равна $b - a = p$.



Датчик случайных чисел сформирует нам набор чисел $\left(x_1, x_2, \ldots, x_n\right)$, являющихся значениями случайной величины непрерывного типа, имеющей равномерное распределение вероятностей на сегменте $[0;1]$.

Определим последовательность независимых случайных величин $\xi_1, \xi_2, \ldots, \xi_n$. В этой последовательности каждая случайная $\xi_k$ принимает только два значения: $\xi_k = 1$, если $x_k \in [a;b]$ и $\xi_k = 0$, если $x_k \notin [a;b]$. Так как набор чисел $\left(x_1, x_2, \ldots, x_n\right)$ - это значения распределённой равномерно на $[0;1]$ некоторой случайной величины, то:

$P(\xi_k = 1) = \dfrac{mes[a;b]}{mes[0;1]} = p, \quad P(\xi_k = 0) = 1 - p = q$. Если $m$ – количество чисел набора $\left(x_1, x_2, \ldots, x_n\right)$, которые принадлежат отрезку $[a;b]$, то $\sum\limits_{k=1}^{n} \xi_k = m$, а $\dfrac{m}{n}$ - значение относительной частоты наступления события $A$. Согласно УЗБЧ $\dfrac{1}{n}\sum\limits_{k=1}^{n} \xi_k \xrightarrow{п.н.} \dfrac{1}{n}\sum\limits_{k=1}^{n} M\xi_k = p$.

2. Пусть $\{A_j\}, j = 1,2,\ldots, l$ – случайные события, образующие полную группу событий (ПГС), то есть: $A_j \cap A_k = \varnothing = V$ - невозможное событие и $\bigcup\limits_{j=1}^{l} A_j = \Omega = U$ - достоверное событие. Если $P(A_j) = p_j$, то $\sum\limits_{j=1}^{l} p_j = 1$. Для моделирования такой последовательности случайных событий и статистической оценки вероятностей их наступления полуотрезок $[0;1)$ разбиваем на $l$ непересекающихся полуинтервалов $\{[a_j; b_j)\}$: $\bigcup\limits_{j=1}^{l} [a_j; b_j) = [0;1)$. Длина полуинтервала $[a_j; b_j)$ равна вероятности наступления события $A_j$: $P(A_j) = \dfrac{mes[a_j; b_j)}{mes[0;1)}$.

Датчик случайных чисел сформирует нам набор чисел $\left(x_1, x_2, \ldots, x_n\right)$, являющиеся значениями равномерно распределённой на $[0;1)$ случайной величины. Если $m_j$ - количество чисел этого набора, которые принадлежат $[a_j; b_j)$, то $\dfrac{m_j}{n}$ относительная частота наступлений события $A_j$. Согласно УЗБЧ для каждого $j$ мы можем утверждать: $\dfrac{m_j}{n} \xrightarrow{п.н.} p_j$.

3. Статистический метод численного определения значения определённого интеграла.

Пусть надо вычислить значение определённого интеграла $I = \int\limits_{0}^{\frac{\pi}{2}} \cos x dx$. Статистическими методами можно получить приближённое значение этого интеграла, которое, благодаря УЗБЧ, с большой степенью уверенности будет мало отличаться от его значения.

Мера области интегрирования $mes\left[0;\frac{\pi}{2}\right]$ равна $\frac{\pi}{2}$.

Запишем интеграл так: $I=\int_0^{\frac{\pi}{2}}\frac{\pi}{2}\cos x\cdot\frac{1}{\frac{\pi}{2}-0}dx$.

Функция $p(x)=\begin{cases}0, & x\notin\left[0;\frac{\pi}{2}\right]\\ \frac{1}{\frac{\pi}{2}-0}, & x\in\left[0;\frac{\pi}{2}\right]\end{cases}$ является плотностью вероятности случайной величины $\xi$ равномерно распределённой на отрезке $\left[0;\frac{\pi}{2}\right]$. Мы можем рассматривать случайную величину $\eta=\frac{\pi}{2}\cos\xi$ как функцию $\eta=g(\xi)$ случайной величины $\xi$. Тогда, согласно общему определению математического ожидания $M[g(\xi)]=\int_R g(x)\cdot p(x)dx$, значение математического ожидания $M\left[\frac{\pi}{2}\cos\xi\right]$ будет являться значением определённого интеграла $I$.

Датчик случайных чисел даёт нам числа $(x_1,x_2,\dots,x_n)$, являющиеся значениями последовательности независимых случайных величин $\zeta_1,\zeta_2,\dots,\zeta_n$, равномерно распределённых на отрезке $[0;1]$. Если определить $\xi=\frac{\pi}{2}\zeta$, то по выборке $(x_1,x_2,\dots,x_n)$ мы получим числа $(y_1,y_2,\dots,y_n)$, где $y_i=\frac{\pi}{2}x_i$, являющиеся значениями равномерно распределённых на отрезке $\left[0;\frac{\pi}{2}\right]$ случайных величин $\xi_1,\xi_2,\dots,\xi_n$, плотность вероятности которых равна $p(x)$.

Так как так как $\eta_i=g(\xi_i)=\frac{\pi}{2}\cos\xi_i$, то по значениям случайных величин $\xi_1,\xi_2,\dots,\xi_n$ мы вычисляем значения случайных величин $\eta_1,\eta_2,\dots,\eta_n$, которые независимые и одинаково распределены. Определяем $\overline{\eta}=\frac{1}{n}\sum_{i=1}^{n}\eta_i=\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\pi}{2}\cos\xi_i\right)$. Согласно УЗБЧ $\overline{\eta}\xrightarrow{п.н.}I$, то есть $P(|\overline{\eta}-I|<\varepsilon)\geq\gamma$.

Возникает вопрос: «Каким должно быть число $n$, чтобы для выбранного $\varepsilon>0$, с уверенностью не меньшей, чем $\gamma\approx1$ выполнялось $P(|\overline{\eta}-I|<\varepsilon)\geq\gamma$»?

Пусть $\boldsymbol{\varepsilon}=0{,}05$ и $\boldsymbol{\gamma}=0{,}9$. Так как, согласно теореме Леви-Хинчина, среднее арифметическое $\overline{\boldsymbol{\eta}}$ случайных величин подчиняется нормальному распределению $\overline{\boldsymbol{\eta}}\in\mathcal{N}\left(\boldsymbol{I};\frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}}\right)$, то $\boldsymbol{P}\left(\left|\overline{\boldsymbol{\eta}}-\boldsymbol{I}\right|<\boldsymbol{\varepsilon}\right)=2\Phi\left(\frac{\boldsymbol{\varepsilon}\sqrt{\boldsymbol{n}}}{\boldsymbol{\sigma}}\right)$. Интересующее нас значение $\boldsymbol{n}$ определим из неравенства: $2\Phi\left(\frac{\boldsymbol{\varepsilon}\sqrt{\boldsymbol{n}}}{\boldsymbol{\sigma}}\right)\geq\boldsymbol{\gamma}$, то есть $\boldsymbol{n}\geq\frac{\boldsymbol{\sigma}^2}{\boldsymbol{\varepsilon}^2}\cdot\left[\boldsymbol{\Phi}^{-1}\left(\frac{\boldsymbol{\gamma}}{2}\right)\right]$.

Дисперсия $\boldsymbol{\sigma}^2=\boldsymbol{D}\boldsymbol{\eta}_{\boldsymbol{i}}=\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_1^2$, где $\boldsymbol{\alpha}_1=\boldsymbol{M}\boldsymbol{\eta}_{\boldsymbol{i}}=1$ и $\boldsymbol{\alpha}_2=\boldsymbol{M}\boldsymbol{\eta}_{\boldsymbol{i}}^2=\frac{\boldsymbol{\pi}^2}{8}$. По таблицам значений функции Лапласа определяем $\boldsymbol{\Phi}^{-1}\left(\frac{0.9}{2}\right)=1{,}645$. Следовательно,

$$\boldsymbol{n}\geq\frac{\frac{\boldsymbol{\pi}^2}{8}-1}{0{,}05^2}\cdot 1{,}645^2\approx 252{,}96.$$

Значит, для определения с уверенностью $\boldsymbol{\gamma}=0{,}9$ значения интеграла с точностью $\boldsymbol{\varepsilon}=0{,}05$ надо, используя датчик случайных чисел, получить выборку $(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_{\boldsymbol{n}})$, состоящую не менее чем из 253 чисел.

*Замечание.* Ясно, что, желая узнать значение подобного интеграла, никто не будет прибегать к методу Монте-Карло. Тем более, определяя минимально необходимый объём выборки $\boldsymbol{n}$, мы нашли значение этого интеграла, вычисляя значение $\boldsymbol{\alpha}_1=\boldsymbol{M}\boldsymbol{\eta}_{\boldsymbol{i}}=1$. Иногда советуют применять метод Монте-Карло при вычислении кратных интегралов от сложных функций. Но всё же представляется, что методы численного интегрирования и в таких случаях будут более предпочтительными.

**Таблица значений функции** $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

| | .,.0 | .,.1 | .,.2 | .,.3 | .,.4 | .,.5 | .,.6 | .,.7 | .,.8 | .,.9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0,0 | 0,3989 | 3989 | 3989 | 3988 | 3986 | 3984 | 3982 | 3980 | 3977 | 3973 |
| 0,1 | 3970 | 3965 | 3961 | 3956 | 3951 | 3945 | 3939 | 3932 | 3925 | 3918 |
| 0,2 | 3910 | 3902 | 3894 | 3885 | 3876 | 3867 | 3857 | 3847 | 3836 | 3825 |
| 0,3 | 3814 | 3802 | 3790 | 3778 | 3765 | 3752 | 3739 | 3726 | 3712 | 3697 |
| 0,4 | 3683 | 3668 | 3652 | 3637 | 3621 | 3605 | 3589 | 3572 | 3555 | 3538 |
| 0,5 | 0,3521 | 3503 | 3485 | 3467 | 3448 | 3429 | 3410 | 3391 | 3372 | 3352 |
| 0,6 | 3332 | 3312 | 3292 | 3271 | 3251 | 3230 | 3209 | 3187 | 3166 | 3144 |
| 0,7 | 3123 | 3101 | 3079 | 3056 | 3034 | 3011 | 2989 | 2966 | 2943 | 2920 |
| 0,8 | 2897 | 2874 | 2850 | 2827 | 2803 | 2780 | 2756 | 2732 | 2709 | 2685 |
| 0,9 | 2661 | 2637 | 2613 | 2589 | 2565 | 2541 | 2516 | 2492 | 2468 | 2444 |
| 1,0 | 0,2420 | 2396 | 2371 | 2347 | 2323 | 2299 | 2275 | 2251 | 2227 | 2203 |
| 1,1 | 2179 | 2155 | 2131 | 2107 | 2083 | 2059 | 2036 | 2012 | 1989 | 1965 |
| 1,2 | 1942 | 1919 | 1895 | 1872 | 1849 | 1826 | 1804 | 1781 | 1758 | 1736 |
| 1,3 | 1714 | 1691 | 1669 | 1647 | 1626 | 1604 | 1582 | 1561 | 1539 | 1518 |
| 1,4 | 1497 | 1476 | 1456 | 1435 | 1415 | 1394 | 1374 | 1354 | 1334 | 1315 |
| 1,5 | 0,1295 | 1276 | 1257 | 1238 | 1219 | 1200 | 1182 | 1163 | 1145 | 1127 |
| 1,6 | 1109 | 1092 | 1074 | 1057 | 1040 | 1023 | 1006 | 0989 | 0973 | 0957 |
| 1,7 | 0940 | 0925 | 0909 | 0893 | 0878 | 0863 | 0848 | 0833 | 0818 | 0804 |
| 1,8 | 0790 | 0775 | 0761 | 0748 | 0734 | 0721 | 0707 | 0694 | 0681 | 0669 |
| 1,9 | 0656 | 0644 | 0632 | 0620 | 0608 | 0596 | 0584 | 0573 | 0562 | 0551 |
| 2,0 | 0,0540 | 0529 | 0519 | 0508 | 0498 | 0488 | 0478 | 0468 | 0459 | 0449 |
| 2,1 | 0440 | 0431 | 0422 | 0413 | 0404 | 0396 | 0387 | 0379 | 0371 | 0363 |
| 2,2 | 0355 | 0347 | 0339 | 0332 | 0325 | 0317 | 0310 | 0303 | 0297 | 0290 |
| 2,3 | 0283 | 0277 | 0270 | 0264 | 0258 | 0252 | 0246 | 0241 | 0235 | 0229 |
| 2,4 | 0224 | 0219 | 0213 | 0208 | 0203 | 0198 | 0194 | 0189 | 0184 | 0180 |
| 2,5 | 0,0175 | 0171 | 0167 | 0163 | 0158 | 0154 | 0151 | 0147 | 0143 | 0139 |
| 2,6 | 0136 | 0132 | 0129 | 0126 | 0122 | 0119 | 0116 | 0113 | 0110 | 0107 |
| 2,7 | 0104 | 0101 | 0099 | 0096 | 0093 | 0091 | 0088 | 0086 | 0084 | 0081 |
| 2,8 | 0079 | 0077 | 0075 | 0073 | 0071 | 0069 | 0067 | 0065 | 0063 | 0061 |
| 2,9 | 0060 | 0058 | 0056 | 0055 | 0053 | 0051 | 0050 | 0048 | 0047 | 0046 |
| 3,0 | 0,0044 | 0043 | 0042 | 0040 | 0039 | 0038 | 0037 | 0036 | 0035 | 0034 |
| 3,1 | 0033 | 0032 | 0031 | 0030 | 0029 | 0028 | 0027 | 0026 | 0025 | 0025 |
| 3,2 | 0024 | 0023 | 0022 | 0022 | 0021 | 0020 | 0020 | 0019 | 0018 | 0018 |
| 3,3 | 0017 | 0017 | 0016 | 0016 | 0015 | 0015 | 0014 | 0014 | 0013 | 0013 |
| 3,4 | 0012 | 0012 | 0012 | 0011 | 0011 | 0010 | 0010 | 0010 | 0009 | 0009 |
| 3,5 | 0,0009 | 0008 | 0008 | 0008 | 0008 | 0007 | 0007 | 0007 | 0007 | 0006 |
| 3,6 | 0006 | 0006 | 0006 | 0005 | 0005 | 0005 | 0005 | 0005 | 0005 | 0004 |
| 3,7 | 0004 | 0004 | 0004 | 0004 | 0004 | 0004 | 0003 | 0003 | 0003 | 0003 |
| 3,8 | 0003 | 0003 | 0003 | 0003 | 0003 | 0002 | 0002 | 0002 | 0002 | 0002 |
| 3,9 | 0002 | 0002 | 0002 | 0002 | 0002 | 0002 | 0002 | 0002 | 0001 | 0001 |

**Таблица значений функции** $\Phi(x)=\dfrac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{z^2}{2}}dz$

| | ...,...0 | ...,...1 | ...,...2 | ...,...3 | ...,...4 | ...,...5 | ...,...6 | ...,...7 | ...,...8 | ...,...9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0,0 | 0,0000 | 0,0040 | 0,0080 | 0,0120 | 0,0160 | 0,0199 | 0,0239 | 0,0279 | 0,0319 | 0,0359 |
| 0,1 | 0398 | 0438 | 0478 | 0517 | 0557 | 0596 | 0636 | 0675 | 0714 | 0754 |
| 0,2 | 0793 | 0832 | 0871 | 0910 | 0948 | 0987 | 1026 | 1064 | 1103 | 1141 |
| 0,3 | 1179 | 1217 | 1255 | 1293 | 1331 | 1368 | 1406 | 1443 | 1480 | 1517 |
| 0,4 | 1554 | 1591 | 1628 | 1664 | 1700 | 1736 | 1772 | 1808 | 1844 | 1879 |
| 0,5 | 0,1915 | 0,1950 | 0,1985 | 0,2019 | 0,2054 | 0,2088 | 0,2123 | 0,2157 | 0,2190 | 0,2224 |
| 0,6 | 2258 | 2291 | 2324 | 2356 | 2389 | 2422 | 2454 | 2486 | 2518 | 2549 |
| 0,7 | 2580 | 2612 | 2642 | 2673 | 2704 | 2734 | 2764 | 2794 | 2823 | 2852 |
| 0,8 | 2881 | 2910 | 2939 | 2967 | 2996 | 3023 | 3051 | 3078 | 3106 | 3133 |
| 0,9 | 3159 | 3186 | 3212 | 3238 | 3264 | 3289 | 3315 | 3340 | 3365 | 3389 |
| 1,0 | 0,3413 | 0,3438 | 0,3461 | 0,3485 | 0,3508 | 0,3531 | 0,3554 | 0,3577 | 0,3599 | 0,3621 |
| 1,1 | 3643 | 3665 | 3686 | 3708 | 3729 | 3749 | 3770 | 3790 | 3810 | 3830 |
| 1,2 | 3849 | 3869 | 3888 | 3906 | 3925 | 3944 | 3962 | 3980 | 3997 | 4015 |
| 1,3 | 4032 | 4049 | 4066 | 4082 | 4099 | 4115 | 4131 | 4147 | 4162 | 4177 |
| 1,4 | 4192 | 4207 | 4222 | 4236 | 4251 | 4265 | 4279 | 4292 | 4306 | 4319 |
| 1,5 | 0,4332 | 0,4345 | 0,4357 | 0,4370 | 0,4382 | 0,4394 | 0,4406 | 0,4418 | 0,4430 | 0,4441 |
| 1,6 | 4452 | 4463 | 4474 | 4484 | 4495 | 4505 | 4515 | 4525 | 4535 | 4545 |
| 1,7 | 4554 | 4564 | 4573 | 4582 | 4591 | 4599 | 4608 | 4616 | 4625 | 4633 |
| 1,8 | 4641 | 4648 | 4656 | 4664 | 4671 | 4678 | 4686 | 4693 | 4700 | 4606 |
| 1,9 | 4713 | 4719 | 4726 | 4732 | 4738 | 4744 | 4750 | 4756 | 4762 | 4767 |
| 2,0 | 0,4772 | 0,4778 | 0,4783 | 0,4788 | 0,4793 | 0,4798 | 0,4803 | 0,4808 | 0,4812 | 0,4817 |
| 2,1 | 4821 | 4826 | 4830 | 4834 | 4838 | 4842 | 4846 | 4850 | 4854 | 4857 |
| 2,2 | 4861 | 4864 | 4868 | 4871 | 4874 | 4878 | 4881 | 4884 | 4887 | 4890 |
| 2,3 | 4893 | 4896 | 4898 | 4901 | 4904 | 4906 | 4909 | 4911 | 4913 | 4916 |
| 2,4 | 4918 | 4920 | 4922 | 4924 | 4927 | 4929 | 4930 | 4932 | 4934 | 4936 |
| 2,5 | 0,4938 | 0,4940 | 0,4941 | 0,4943 | 0,4945 | 0,4946 | 0,4948 | 0,4949 | 0,4951 | 0,4952 |
| 2,6 | 4953 | 4955 | 4956 | 4957 | 4958 | 4960 | 4961 | 4962 | 4963 | 4964 |
| 2,7 | 4965 | 4966 | 4967 | 4968 | 4969 | 4970 | 4971 | 4972 | 4973 | 4974 |
| 2,8 | 4974 | 4975 | 4976 | 4977 | 4977 | 4978 | 4979 | 4980 | 4980 | 4981 |
| 2,9 | 4981 | 4982 | 4982 | 4983 | 4984 | 4984 | 4985 | 4985 | 4986 | 4986 |
| 3,0 | 0,4986 | 0,4987 | 0,4987 | 0,4988 | 0,4988 | 0,4989 | 0,4989 | 0,4989 | 0,4990 | 0,4990 |
| 3,1 | 4990 | 4991 | 4991 | 4991 | 4992 | 4992 | 4992 | 4992 | 4993 | 4993 |
| 3,2 | 4993 | 4993 | 4994 | 4994 | 4994 | 4994 | 4994 | 4995 | 4995 | 4995 |
| 3,3 | 4995 | 4995 | 4996 | 4996 | 4996 | 4996 | 4996 | 4996 | 4996 | 4996 |
| 3,4 | 4997 | 4997 | 4997 | 4997 | 4997 | 4997 | 4997 | 4997 | 4998 | 4998 |
| 3,5 | 0,49977 | 0,49978 | 0,49978 | 0,49979 | 0,49980 | 0,49981 | 0,49981 | 0,49982 | 0,49983 | 0,49983 |
| 3,6 | 49984 | 49985 | 49985 | 49986 | 49986 | 49987 | 49987 | 49988 | 49988 | 49989 |
| 3,7 | 49989 | 49990 | 49990 | 49990 | 49991 | 49991 | 49992 | 49992 | 49992 | 49992 |
| 3,8 | 49993 | 49993 | 49993 | 49994 | 49994 | 49994 | 49994 | 49995 | 49995 | 49995 |
| 3,9 | 49995 | 49995 | 49996 | 49996 | 49996 | 49996 | 49996 | 49996 | 49997 | 49997 |

4,0 0,49997 4,5 0,499997 5,0 0,49999997

**Список литературы**


1. Боровков А.А. Теория вероятностей. -2-е издание. - М.: Наука, 1986.
2. Гихман И.И., Скороход А.В., Ядренко М.И. Теория вероятностей и Математическая статистика. – Киев: Вища школа, 1979.
3. Гнеденко Б.В. Курс теории вероятностей.- Издание шестое.- М.: Наука, 1988.
4. Гнеденко Б.В. , Колмогоров А.Н. Предельные распределения для сумм независимых случайных величин. – М.-Л.: Гостехиздат, 1949.
5. Климов Г.П. Теория вероятностей и математическая статистика. – М.: Изд. МГУ, 1983.
6. Колемаев В.А., Староверов О.В., Турундаевский В.Б. Теория вероятностей и математическая статистика. – М.: Высшая школа, 1991.
7. Колмогоров А.Н. Основные понятия теории вероятностей. 2-е издание. – М.: Наука, 1974.
8. Королюк В.С., Портенко Н.И., Скороход А.В., Турбин А.Ф, Справочник По теории вероятностей и математической статистике. – 2-е изд.-М.: Наука, 1985.
9. Лоэв М. Теория вероятностей. – М.: ИЛ, 1962.
10. Нейман Ю. Вводный курс теории вероятностей и математической статистики.- М.: 1968.
11. Прохоров Ю.В., Розанов Ю.А. Теория вероятностей. – 2-е изд.- М.: Наука, 1985.
12. Севастьянов Б.А. Курс теории вероятностей и математической статистики. - М.: Наука, 1982.
13. Тутубалин В.Н. Теория вероятностей. – М.: Изд. МГУ, 1972.
14. Феллер В. Введение в теорию вероятностей и её приложения. В 2-х томах, - М.: Мир, 1984.
15. Чистяков В.П. Курс теории вероятностей. – М.: Наука,1978.
16. Ширяев А.Н. Вероятность.- М.: Наука, 1980.