Министерство образования и науки Российской Федерации

**САНКТ-ПЕТЕРБУРГСКИЙ
ГОСУДАРСТВЕННЫЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ**

Д.А. Тархов, А.В.Андреев, Е.С. Единова,

А.М. Никулин, Т.А. Шемякина

МАТЕМАТИКА

ТЕОРИЯ ВЕРОЯТНОСТЕЙ
ЧАСТЬ 1

Учебное пособие

Санкт-Петербург

Издательство СПбГПУ

2012

УДК 512.8

Д.А. Тархов, А.В.Андреев, Е.С. Единова, А.М. Никулин, Т.А. Шемякина
Математика. Теория вероятностей. Часть 1: Учебное пособие / Д.А. Тархов, А.В.Андреев, Е.С. Единова, А.М. Никулин, Т.А. Шемякина, 2012. – 66с.

Учебное пособие соответствует содержанию Федерального Государственного образовательного стандарта высшего профессионального образования дисциплины "Высшая математика" направления подготовки и переподготовки бакалавров по специальности 280700.62 "Техносферная безопастность".

В пособии кратко изложены теоретические основы по курсу "Высшая математика" , который представлен начальными разделами теории вероятностей.

Предназначено для студентов высших учебных заведений технических и экономических направлений, изучающих дисциплину "Высшая математика". Пособие может быть использовано при подготовке бакалавров и в системе дополнительного профессионального образования, а также оно будет полезно для преподавателей дневных, вечерних и заочных отделений вузов и технических университетов.

Библиогр.: 6 назв.

Публикуется по решению методического совета факультета Комплексной Безопасности Санкт-Петербургского государственного политехнического университета.

Глава 1. Введение

1.1 Дискретная теория вероятностей

Одним из основных понятий теории вероятностей является понятие *случайного события* (поскольку в теории вероятностей рассматриваются лишь случайные события, то чаще слово "случайное"отбрасывается и говорят просто "событие"). Можно сказать, что под "событием"понимается всякий факт, который в результате *эксперимента* (ещё одно из основополагающих понятий теории вероятностей) может произойти или нет. Например, эксперимент состоит из однократного подбрасывания монеты - тогда под событием разумно понимать выпадение герба или решки (заведомо отбросим такие нереальные в реальной жизни факты, как постановка монеты на ребро, зависание в воздухе и т.д. и т.п.). Другие варианты событий, часто встречаемые в обычной жизни, - получение выигрыша в лотерее, получение счастливого билета на экзамене, рождение мальчика (или девочки) и т.д. и т.п.

Исходя из повседневного жизненного опыта ясно, что различные события можно сравнивать по степени их возможности. Так, если монета симметрична (то есть не фальшивая), то разумно считать, что оба возможных события - выпадение герба или решки - равновозможны, из многолетних наблюдений за количеством новорожденных известно, что частота рождения мальчика чуть выше, чем девочки и т.д.

Да и вообще, первоначальное понятие вероятности события (как меры возможности наступления события) связывалось именно с *частотой* появления данного события в серии (конечной) экспериментов, проводимых в *одинаковых условиях.*

Во введении речь будет идти о так называемой *дискретной* теории вероятностей. Обозначим $\Omega = \{\omega_1, \omega_2, ..., \omega_n, ...\}$ множество всех возможных исходов эксперимента. Множество Ω может быть конечным или счетным. Элементы ω_i называются *элементарными событиями* или *элементарными исходами*, а само Ω называется *пространством элементарных исходов*.

Предположим сначала, что в результате эксперимента может произойти *одно и только одно* из элементарных событий и предположим, что каждому ω_i отвечает неотрицательное число $\mathbf{P}(\omega_i)$, называемое *вероятностью* элементарного исхода ω_i, такое, что

1) $$0 \le \mathbf{P}(\omega_i) \le 1,$$

2) $$\sum_{\omega_i \in \Omega} \mathbf{P}(\omega_i) = 1.$$

Пусть в результате эксперимента произошло элементарное событие ω, $\omega \in \Omega$, которое может быть иногда ненаблюдаемо, но есть информация о его принадлежности некоторому множеству A из пространства элементарных исходов Ω, то есть

$$\omega \in A = \{\omega_{i_1}, \omega_{i_2}, ..., \omega_{i_k}, ...\} \subseteq \Omega.$$

В такой ситуации мы говорим, что произошло *случайное событие* или просто *событие* A (отметим , что никаких условий на множество A не накладывалось, т.о. событием будет являться *любое* подмножество пространства элементарных исходов Ω). Таким образом, любое событие, факт наступления которого *вполне определяется* тем, какое из элементарных событий наступило, как раз и называется случайным событием.

Случайные события мы будем обозначать большими буквами латинского алфавита (возможно, с индексом) $A, B, C, A_n, B_m, ...$

По условию, факт наступления события A зависит от ω: при некоторых ω_{i_j} из $\{\omega_1, \omega_2, ..., \omega_n, ...\}$ событие A наступает, при остальных - нет. Элементарное событие, при котором событие A наступает, назовём *благоприятствующим* событию A, остальные элементарные события назовём *неблагоприятствующими* событию A. Все вместе они образуют событие $\bar{A} = \Omega \setminus A$, называемое дополнением к событию A (отрицание события A). Таким образом любое событие вполне задаётся указанием тех ω_{i_j}, которые ему благоприятствуют.

В такой схеме вероятность $\mathbf{P}(A)$ события A определяется по формуле

$$\mathbf{P}(A) = \sum_{\omega_{i_j} \in A} \mathbf{P}(\omega_{i_j}).$$

Обозначим $\emptyset$ пустое множество, ему не благоприятствует ни одно элементарное событие, поэтому $\emptyset$ называют *невозможным событием*, при этом $\mathbf{P}(\emptyset) = 0$; пространству Ω благоприятствуют все элементарные события, поэтому Ω называют *достоверным событием*, и естественно, что $\mathbf{P}(\Omega) = 1$. Очевидно, что $\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A)$. Отметим следующие очевидные факты :

1. если $A \subseteq B \subseteq \Omega$, то $\mathbf{P}(A) \leq \mathbf{P}(B)$.

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

   Действительно:

$$\mathbf{P}(A \bigcup B) = \sum_{\omega \in A \bigcup B} \mathbf{P}(\omega) = \sum_{\omega \in A} \mathbf{P}(\omega) + \sum_{\omega \in B} \mathbf{P}(\omega) - \sum_{\omega \in A \bigcap B} \mathbf{P}(\omega) =$$

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

В частности, если $A \bigcap B = \emptyset$, то есть A и B *несовместные события*, то

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

3. Пусть $A_i \in \Omega$, $i \in \{1, ..., n\}$. Тогда

$$\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbf{P}(A_i) - \sum_{i<j} \mathbf{P}(A_i \bigcap A_j) +$$

$$+ \sum_{i<j<h} \mathbf{P}(A_i \bigcap A_j \bigcap A_h) - ... + (-1)^{n-1} \mathbf{P}(A_1 \bigcap ... \bigcap A_n).$$

Действительно, воспользуемся методом математической индукции (ММИ). Базу ММИ ($n = 2$) дает предыдущий пункт. Пусть утверждение верно для произвольных $n-1$ событий $A_1, ..., A_{n-1}$. Тогда, обозначив

$$B = \bigcup_{i=1}^{n-1} A_i,$$

имеем:

$$\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \mathbf{P}\left(B \bigcup A_n\right) = \mathbf{P}(B) + \mathbf{P}(A_n) - \mathbf{P}(B \bigcap A_n).$$

Учитывая, что

$$\mathbf{P}(B) = \mathbf{P}\left(\bigcup_{i=1}^{n-1} A_i\right), \ \mathbf{P}(B \bigcap A_n) = \mathbf{P}\left(\bigcup_{i=1}^{n-1} (A_i \bigcap A_n)\right),$$

получаем искомое утверждение:

$$\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \mathbf{P}(\bigcup_{i=1}^{n-1} A_i) + \mathbf{P}(A_n) - \mathbf{P}(\bigcup_{i=1}^{n-1} (A_i \bigcap A_n) =$$

$$= \sum_{i=1}^{n-1} \mathbf{P}(A_i) - \sum_{1 \le i<j \le n-1} \mathbf{P}(A_i \bigcap A_j) +$$

$$+ \sum_{1 \le i<j<h \le n-1} \mathbf{P}(A_i \bigcap A_j \bigcap A_h) - ... + (-1)^{n-2} \mathbf{P}(A_1 \bigcap ... \bigcap A_{n-1}) +$$

$$+\mathbf{P}(A_n)-\left(\sum_{i=1}^{n-1}\mathbf{P}(A_i\bigcap A_n)-\sum_{1\le i<j\le n-1}\mathbf{P}\left\{(A_i\bigcap A_n)\bigcap(A_j\bigcap A_n)\right\}+\right.$$

$$+\sum_{1\le i<j<h\le n-1}\mathbf{P}\left\{(A_i\bigcap A_n)\bigcap(A_j\bigcap A_n)\bigcap(A_h\bigcap A_n)\right\}-...$$

$$\left....+(-1)^{n-2}\mathbf{P}\left\{(A_1\bigcap A_n)\bigcap...\bigcap(A_{n-1}\bigcap A_n)\right\}\right)=$$

$$=\sum_{i=1}^{n}\mathbf{P}(A_i)-\sum_{1\le i<j\le n}\mathbf{P}(A_i\bigcap A_j)+$$

$$+\sum_{1\le i<j<h\le n}\mathbf{P}(A_i\bigcap A_j\bigcap A_h)-...+(-1)^{n-1}\mathbf{P}(A_1\bigcap...\bigcap A_n).$$

В частности, если события $A_i\in\Omega$, $i\in\{1,...,n\}$ *попарно несовместны*, то есть $A_i\bigcap A_j=\emptyset$, $i\neq j$, имеем (данное свойство называется *конечной аддитивностью* вероятности):

$$\mathbf{P}\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}\mathbf{P}(A_i).$$

Замечание.

1. Рассмотрим попарно несовместные события $A_i\in\Omega$, $i\in\{1,...,\infty\}$. Тогда (*счетная аддитивность* вероятности):

$$\mathbf{P}\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}\mathbf{P}(A_i).$$

   Это следует из

$$\mathbf{P}\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}\mathbf{P}(A_i)$$

   и

$$\mathbf{P}\left(\bigcup_{i=n+1}^{\infty}A_i\right)\to 0,\quad n\to\infty.$$

2.

$$\mathbf{P}(A\bigcup B)\le\mathbf{P}(A)+\mathbf{P}(B)\Rightarrow\mathbf{P}\left(\bigcup_{i=1}^{\infty}A_i\right)\le\sum_{i=1}^{\infty}\mathbf{P}(A_i).$$

Ясно, что всегда
$$0 \leq \mathbf{P}(A) \leq 1.$$
Наконец, приведем важное понятие независимости двух событий : говорят, что события A и B *независимы*, если
$$\mathbf{P}(A \bigcap B) = \mathbf{P}(A) \cdot \mathbf{P}(B).$$
Более подробно об основах теории вероятностей см., например, Тутубалин (1972), Браунли (1977).

Вероятность и уровень значимости.

Среди всех возможных событий для практики особый интерес представляют события, которые практически не могут появиться при однократном проведении эксперимента. Их называют *практически невозможными* событиями. И наоборот, если практически уверены, что событие должно появиться в эксперименте, то естественно его назвать *практически достоверным.* Все остальные события называются просто *возможными.* Формализацию такой классификации событий можно сделать в терминах так называемого уровня значимости.

Уровнем значимости называют число α, $0 < \alpha < 0.5$, такое что, если для события A верно соотношение $\mathbf{P}(A) < \alpha$, то экспериментатор считает, что событие A не может произойти при однократном проведении эксперимента. С другой стороны, очевидно, что должно произойти событие $\bar{A}$, при этом $\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A)$.

Если α мало, то $\mathbf{P}(\bar{A}) \approx 1$, то есть экспериментатор прав, считая, что событие $\bar{A}$ практически достоверно. Следуя этой логике любое событие такое, что $\alpha < \mathbf{P}(A) < 1 - \alpha$, называется возможным. На практике α полагают равным 0.1, 0.05, 0.01, 0.005, 0.001 и так далее. Выбор значения α определяется опытом экспериментатора, его знаниями и т.п.. а также его готовностью принимать рискованные решения.

1.2 Примеры

1. Рассматривается эксперимент с одним подбрасыванием симметричной кости. Тогда элементарные события - выпадение граней со значениями $i = 1, ..., 6$ и с вероятностями $\mathbf{P}(i) = \frac{1}{6}$. Рассмотрим следующие события:
$$A = \{\text{выпали чётные номера}\}, \quad B = \{\text{выпали нечётные номера}\},$$
$$C = \{\text{выпали 1 или 6 }\}, \quad D = \{\text{количество очков - от 2 до 5}\}.$$

Тогда

$$\mathbf{P}(A) = \mathbf{P}(B) = \frac{1}{2}, \quad \mathbf{P}(C) = \frac{1}{3}, \quad \mathbf{P}(D) = \frac{2}{3},$$

A и B - несовместные события, C и D также несовместные события,

$$A \bigcup B = \Omega, \quad \mathbf{P}(A \bigcup B) = \mathbf{P}(A) + \mathbf{P}(B) = \mathbf{P}(\Omega) = 1.$$

Далее:

$$A \bigcap B = \emptyset, \quad A \bigcap C = \{\text{выпал номер } 6\}.$$

Легко видеть, что

$$\mathbf{P}(A \bigcap C) = \frac{1}{6} = \mathbf{P}(A)\mathbf{P}(C),$$

$$\mathbf{P}(A \bigcap B) = 0 \neq \mathbf{P}(A)\mathbf{P}(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

Таким образом, события A и C являются независимыми событиями, а события A и B не являются независимыми.

2. В урне находятся 2 белых и 3 чёрных шара, из урны вынимается наугад 1 шар. Требуется найти вероятность того, что этот шар - белый.

   Итак, пусть А - событие, состоящее в появлении белого шара. Общее число случаев $n = 5$, число случаев, благоприятствующих A, $m = 2$. Следовательно

$$\mathbf{P}(A) = \frac{2}{5}.$$

3. В урне a белых и b чёрных шаров. Из урны вынимают 2 шара. Найти вероятность того, что оба шара белые.

   Пусть B данное событие. Общее число случаев:

   $n = C_{a+b}^2$ - число способов выбрать 2 предмета из общего количества $a + b$;

   число благоприятствующих событию B случаев:

   $m = C_a^2$ - число способов выбрать 2 шара из a белых.

   Таким образом

$$\mathbf{P}(B) = \frac{C_a^2}{C_{a+b}^2}, \quad \text{где } C_n^k = \frac{n!}{k!(n-k)!}.$$

1.3 Вероятность и частота

Пусть произведена *серия* из n опытов, в каждом из которых могло произойти или не произойти некое событие A. Тогда *частотой* (появления) *события* A в данной серии опытов будем называть отношение числа опытов μ_n, в которых появилось событие A, к общему числу n произведенных опытов. Обычно частоту $\frac{\mu_n}{n}$ события A называют *статистической вероятностью.* Обозначим её $\mathbf{P}^*(A)$:

$$\mathbf{P}^*(A) = \frac{\mu_n}{n}, \quad \text{где} \quad 0 \le \mu_n \le n.$$

Проблема в том, что частота $\frac{\mu_n}{n}$ события A в большой мере носит случайный характер: чаще всего она будет изменяться от одной серии опытов к другой. Например, пусть мы дважды по 10 раз подкинем монету, причём мы не знаем, фальшивая она или нет. В силу крайне ограниченного числа опытов вполне может получиться так, что в первой серии герб выпадет, скажем, 2 раза, а во второй - 7 раз. В таком случае частота в первый раз равна $\frac{1}{5}$, а во второй $\frac{7}{10}$.

К счастью, с увеличением числа опытов частота всё более будет лишаться случайного характера - массовость произведенных экспериментов нивелирует колебания частот, которые стабилизируются ("сходятся") к некоторому числу между 0 и 1. Например в случае с подлинной монетой частота появления герба (впрочем, как и решки) будет лишь незначительно уклоняться от $\frac{1}{2}$. Если, например, рассматривать большое количество новорожденных, то частота появления мальчика будет колебаться около 0.514.

Как выясняется, это свойство "устойчивости частоты"является одной из наиболее характерных закономерностей, наблюдаемых в случайных явлениях.

Поэтому число, к которому "выравнивается"частота при увеличении числа экспериментов, принимают за вероятность этого события.

Говоря о характере такой "сходимости"нужно отметить, что это так называемая "сходимость по вероятности". Это свойство выражается *теоремой Бернулли* (здесь мы забегаем вперед, поэтому пока опустим строгую формулировку и, тем более, доказательство):

$$\frac{\mu_n}{n} \to \mathbf{P}(A)\ ,\ n \to \infty \quad \textit{по вероятности}$$

при этом говорят, что случайная величина $\frac{\mu_n}{n}$ сходится *по вероятности* к числу (*неслучайному*) $\mathbf{P}(A)$. Это означает, что какое бы $\varepsilon > 0$ мы ни взяли, с ростом числа экспериментов n вероятность

$$\mathbf{P}(|\frac{\mu_n}{n} - \mathbf{P}(A)| < \varepsilon)$$

неограниченно приближается к 1.

В заключение отметим: в теории вероятностей принято пропускать знак $\bigcap$. В дальнейшем вместо $A \bigcap B$ будем писать AB, вместо $A_1 \bigcap A_2 \bigcap \cdots \bigcap A_n$ будем писать $A_1 A_2 \ldots A_n$ и т.д..

Глава 2. Аксиоматическое построение вероятности (аксиоматика Колмогорова)

2.1 Вероятностное пространство

Предложенное выше определение вероятности события через сумму вероятностей содержащихся в нем элементарных событий действенно в ситуации, когда пространство элементарных событий Ω состоит из не более чем счетного числа элементарных исходов. Легко представить себе задачу, когда когда это множество несчетно; например, случайное бросание точки в отрезок $[0, 1]$. В данном случае

$$\Omega = \{\omega\} = [0, 1] \quad , \quad \omega \in [0, 1]$$

и пространство элементарных событий Ω имеет континуум исходов. При этом если в экспериментах, имеющих конечное или счетное число исходов, любая их совокупность представляла событие, то мы встретимся с большими трудностями, если будем считать событием любое подмножество $[0, 1]$ (к тому же невозможно присвоить каждому элементарному исходу $\omega \in [0, 1]$ одну и ту же ненулевую вероятность так, чтобы их сумма равнялась 1). В качестве событий нужно выделить *специальный класс подмножеств* Ω.

Определение 1. *Пусть дано пространство элементарных событий* $\Omega = \{\omega\}$. *Назовём алгеброй на* Ω *семейство* $\mathcal{A} = \{A\}$ *подмножеств* Ω *таких, что*:

1. $\emptyset \in \mathcal{A}$, $\Omega \in \mathcal{A}$.
2. *Если* $A \in \mathcal{A}$, *тогда* $\bar{A} = \Omega \setminus A \in \mathcal{A}$

(то есть $\mathcal{A}$ инвариантна к операции перехода к дополнению).

3. *Если* $A \in \mathcal{A}$, $B \in \mathcal{A}$, *тогда*

$$A \cap B \in \mathcal{A}, \quad A \cup B \in \mathcal{A}.$$

Определение 2. *Пусть дана алгебра* $\mathcal{A}$ *на* Ω. *Назовём* $\mathcal{A}$ σ *-алгеброй* *на* Ω *если*

$$\forall n \in \mathbb{N} \quad A_n \in \mathcal{A}, \quad \text{тогда} \quad \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}, \quad \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}.$$

Замечание. Любая σ - алгебра представляет из себя семейство множеств, замкнутое относительно следующих теоретико- множественных операций:

конечное или счетное объединение,

конечное или счетное пересечение,

разность и симметрическая разность.

Теорема 1. *Любое (конечное или бесконечное) пересечение* σ *-алгебр на* Ω *также является* σ *-алгеброй на* Ω.

Доказательство: непосредственно следует из определения σ - алгебры. Действительно, рассмотрим (конечное или бесконечное) семейство σ - алгебр $\{\mathcal{A}_n\}$ на Ω и пусть $\mathcal{A} = \bigcap_n \mathcal{A}_n$.

1. $\forall n \quad \emptyset \in \mathcal{A}_n, \, \Omega \in \mathcal{A}_n \Rightarrow \emptyset \in \mathcal{A}, \, \Omega \in \mathcal{A}$.

2. $A \in \mathcal{A} \Rightarrow \forall n \, A \in \mathcal{A}_n \Rightarrow \forall n \, \bar{A} \in \mathcal{A}_n \Rightarrow \bar{A} \in \mathcal{A}$.

Аналогичным образом проверяется последний пункт (замкнутость относительно счетного объединения и счетного пересечения).

Следствие. Пусть $\mathcal{E}(\Omega)$ - множество всех подмножеств Ω, $(\Omega \neq \emptyset)$. Рассмотрим набор множеств $\mathcal{T} \subset \mathcal{E}(\Omega)$. Пересечение $\mathcal{F}_0$ всех σ-алгебр (они существуют, т.к. хотя бы $\mathcal{E}(\Omega)$, очевидно, является σ-алгеброй), содержащих $\mathcal{T}$, будет являться наименьшей или минимальной σ-алгеброй, содержащей $\mathcal{T}$. $\mathcal{F}_0$ называется σ-*алгеброй, порождённой* $\mathcal{T}$ и обозначается $\mathcal{F}_0 = \sigma(\mathcal{T})$.

Определение 3. *Рассмотрим пространство элементарных событий* Ω *и* σ*-алгебру* $\mathcal{A}$ *на* Ω. *Тогда пара* $(\Omega, \mathcal{A})$ *называется измеримым пространством.*

Определение 4. *Пусть* $(\Omega, \mathcal{A})$ *- измеримое пространство,* $\Omega = \{\omega\}$. *Назовём* $\mathcal{A} = \{A\}$ *-* σ*-алгеброй событий. Любое множество* $A \in \mathcal{A}$ *называется событием .*

Замечание . Специально отметим, что лишь множества из указанной в паре σ-алгебры являются событиями. Например, $\emptyset$ называется *невозможным событием*, напротив событие $\Omega = \bar{\emptyset}$ называется *достоверным событием.* Если события A и B таковы, что $AB = \emptyset$, то говорят о *несовместных событиях.*

Определение 5.*Пусть* $(\Omega, \mathcal{F})$ *- измеримое пространство. Рассмотрим числовую функцию* $\mathbf{P}$*, заданную на множествах из* $\mathcal{F}$*, обладающую следующими свойствами:*
1. $\forall\, A \in \mathcal{F} \quad \mathbf{P}(A) \geq 0.$
2. $\mathbf{P}(\Omega) = 1.$
3. $\forall\, \{A_n\} : \forall\, n\; A_n \in \mathcal{F}\,,\; A_i A_j = \emptyset\,,\; i \neq j :$

$$\mathbf{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbf{P}(A_n).$$

Таким образом заданную числовую функцию $\mathbf{P}$ будем называть *вероятностью на* $(\Omega, \mathcal{F})$.
Свойство вероятности, заданное пунктом 3, называется *счетной аддитивностью.*
Для дальнейшего нам понадобится следующее утверждение.
Теорема 2. *Эквивалентным пункту* 3 *определения будет требование аддитивности для конечного набора событий* A_i *и выполнение пункта* $3'$ *(свойство непрерывности сверху).*
$3'$. *Пусть последовательность*$\{B_n\}$ *событий такова, что* $B_{n+1} \subset B_n$ *и* $\bigcap_{n=1}^{\infty} B_n = B \in \mathcal{F}$*, тогда* $\mathbf{P}(B_n) \to \mathbf{P}(B)$ *при* $n \to \infty$.
Доказательство: Пусть выполнен пункт 3 и пусть $B_{n+1} \subset B_n, \bigcap_{n=1}^{\infty} B_n = B \in \mathcal{F}$. Тогда последовательность событий B, $C_k = B_k \bar{B}_{k+1}$, $k = 1, 2, ...$, состоит из попарно несовместных событий и $B_n = B \bigcup (\bigcup_{k=n}^{\infty} C_k)$ для любого n. Из свойства 3 получаем, что ряд в правой части равенства

$$\mathbf{P}(B_1) = \mathbf{P}(B) + \sum_{k=1}^{\infty} \mathbf{P}(C_k)$$

сходится, а это значит, что при $n \to \infty$

$$\mathbf{P}(B_n) = \mathbf{P}(B) + \sum_{k=n}^{\infty} \mathbf{P}(C_k) \to \mathbf{P}(B),$$

это и есть свойство $3'$.

Обратно, пусть $\{A_n\}$ - последовательность несовместных событий. Тогда

$$\mathbf{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \mathbf{P}\left(\bigcup_{k=1}^{n} A_k\right) + \mathbf{P}\left(\bigcup_{k=n+1}^{\infty} A_k\right)$$

и имеют место равенства

$$\sum_{k=1}^{\infty} \mathbf{P}(A_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbf{P}(A_k) = \lim_{n\to\infty} \mathbf{P}\left(\bigcup_{k=1}^{n} A_k\right) =$$

$$\lim_{n\to\infty}\left\{\mathbf{P}\left(\bigcup_{k=1}^{\infty}A_k\right)-\mathbf{P}\left(\bigcup_{k=n+1}^{\infty}A_k\right)\right\}=\mathbf{P}\left(\bigcup_{k=1}^{\infty}A_k\right).$$

Последнее равенство следует из $3'$.
Замечание. Если $B=\emptyset$, то говорят о *непрерывности сверху на пустом множестве.* Тогда $\mathbf{P}(B_n)\to 0$ при $n\to\infty$.
Определение 6. *Пусть* $(\Omega,\mathcal{F})$ *- измеримое пространство,* $\mathbf{P}$ *- вероятность на* $(\Omega,\mathcal{F})$. *Будем называть тройку* $(\Omega,\mathcal{F},\mathbf{P})$ *вероятностным пространством.*

Вероятность $\mathbf{P}$ на $(\Omega,\mathcal{F})$ иногда называют *распределением вероятностей* или просто *распределением* на $(\Omega,\mathcal{F})$ (напоминаем, что $\mathcal{F}$ - σ-алгебра).
Замечание. Часто в качестве пространства элементарных исходов Ω рассматривается подмножество числовой прямой $B\subset\mathbb{R}$ или сама прямая. В качестве σ-алгебры событий рассматривается σ-алгебра, порожденная всеми открытыми подмножествами B (или всей прямой). Такую σ-алгебру называют *борелевской* (минимальная σ-алгебра, содержащая B). Будем обозначать её $\mathcal{B}_B$. Элементы борелевской σ-алгебры называют *борелевскими множествами.*
Пример. $(\Omega,\mathcal{F})=([0,1],\mathcal{B}_{[0,1]})$ или $(\Omega,\mathcal{F})=(\mathbb{R},\mathcal{B}_{\mathbb{R}})$, т.е. $B=[0,1]$ или $B=\mathbb{R}$. Отметим, что наряду с интервалами $(a,b), a\in B, b\in B$ борелевскими множествами являются одноточечные множества $\{a\}$ и множества вида $(a,b]$, $[a,b]$, $[a,b)$, (a и b могут принимать, в случае $B=\mathbb{R}$, и бесконечные значения). Это утверждение следует, например, из представлений вида

$$\{a\}=\bigcap_1^{\infty}\left(a-\frac{1}{n},a+\frac{1}{n}\right),\quad (a,b]=\bigcap_1^{\infty}\left(a,b+\frac{1}{n}\right).$$

Таким образом, все счетные множества и счетные объединения и пересечения интервалов или отрезков также являются борелевскими. Из всего сказанного можно сделать вывод, что σ-алгебра, порожденная интервалами, σ-алгебра, порожденная отрезками, σ-алгебра, порожденная полуинтервалами и т.п., - все они представляют из себя одну и ту же борелевскую σ-алгебру.

2.2 Простейшие свойства вероятности

Пусть $(\Omega,\mathcal{F},\mathbf{P})$ - вероятностное пространство и условимся считать все рассматриваемые множества событиями: $A\in\mathcal{F}$, $B\in\mathcal{F}$, $A_k\in\mathcal{F}$, $B_l\in\mathcal{F}$. Тогда имеют место следующие свойства.
1. $\mathbf{P}(\emptyset)=0$.
Вытекает из пунктов 2 и 3 определения вероятности, и $\emptyset\bigcup\Omega=\Omega$.
2. $\mathbf{P}(\bar{A})=1-\mathbf{P}(A)$.

Это следует из
$$A\bigcup\bar{A}=\Omega\,,\ A\bar{A}=\emptyset.$$
3. $A\subseteq B\Rightarrow \mathbf{P}(A)\le\mathbf{P}(B)$.
Свойство *монотонности* вероятности, следует из её неотрицательности и очевидного равенства
$$\mathbf{P}(A)+\mathbf{P}(\bar{A}B)=\mathbf{P}(B).$$

4. $\mathbf{P}(A)\le 1$.
Очевидно, т.к. $A\subseteq\Omega$, $\mathbf{P}(\Omega)=1$.
5. $\mathbf{P}(A\bigcup B)=\mathbf{P}(A)+\mathbf{P}(B)-\mathbf{P}(AB)$.
Следует из
$$A\bigcup B=A\bigcup(B\backslash(AB))\,,\ A(B\backslash(AB))=\emptyset\,,\ \mathbf{P}(B\backslash(AB))=\mathbf{P}(B)-\mathbf{P}(AB).$$

6. $\mathbf{P}(A\bigcup B)\le\mathbf{P}(A)+\mathbf{P}(B)$.
Очевидное следствие предыдущего пункта.

7.
$$\mathbf{P}(\bigcup_{k=1}^{n}A_k)=\sum_{k=1}^{n}\mathbf{P}(A_k)-\sum_{k<l}\mathbf{P}(A_kA_l)+$$
$$+\sum_{k<l<m}\mathbf{P}(A_kA_lA_m)-\ldots(-1)^{n-1}\mathbf{P}(A_1\ldots A_n).$$
Свойство доказывалось для не более чем счетного (*дискретного*) Ω. В общем случае доказательство не изменится.
8. Счетная полуаддитивность .
$$\mathbf{P}(\bigcup_{n=1}^{\infty}A_n)\le\sum_{n=1}^{\infty}\mathbf{P}(A_n).$$

Действительно,
$$\bigcup_{n=1}^{\infty}A_n=\bigcup_{n=1}^{\infty}(B_nA_n)\,,\ \text{где}\ B_n=\Omega\backslash\bigcup_{k=1}^{n-1}A_k.$$

Отметим *дизъюнктность* событий $(A_kB_k)\bigcap(A_lB_l)=\emptyset$, $k\ne l$. Из всего этого следует
$$\mathbf{P}(\bigcup_{n=1}^{\infty}A_n)=\sum_{n=1}^{\infty}\mathbf{P}(B_nA_n)\le\sum_{n=1}^{\infty}\mathbf{P}(A_n).$$

9. *Непрерывность снизу.* Рассмотрим возрастающую последовательность событий $\{A_n\}$, $A_n \subseteq A_{n+1}$ и пусть $A = \bigcup_{n=1}^{\infty} A_n$. Тогда

$$\lim_{n\to\infty} \mathbf{P}(A_n) = \mathbf{P}(A).$$

На самом деле это свойство является лишь другой формой свойства *непрерывности сверху.* Введем в рассмотрение множества $B_n = A\backslash A_n$ и заметим, что $B_{n+1} \subseteq B_n$, $\bigcap_{n=1}^{\infty} B_n = \emptyset$. По непрерывности сверху на пустом множестве $\mathbf{P}(B_n) \to 0$, $n \to \infty$. Осталось лишь отметить, что

$$\mathbf{P}(B_n) = \mathbf{P}(A\backslash A_n) = \mathbf{P}(A) - \mathbf{P}(A_n) \Rightarrow \mathbf{P}(A) - \mathbf{P}(A_n) \to 0\ ,\ n \to \infty.$$

2.3 Независимость. Условная вероятность. Формула полной вероятности. Формула Байеса

Пусть $(\Omega, \mathcal{F}, \mathbf{P})$ - вероятностное пространство.
Определение 7. *Рассмотрим события A и B, и пусть $\mathbf{P}(B) > 0$. Условной вероятностью события A при условии, что произошло событие B, будем называть*

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(AB)}{\mathbf{P}(B)}.$$

Определение 8. *События A и B называются независимыми, если*

$$\mathbf{P}(AB) = \mathbf{P}(A)\mathbf{P}(B).$$

Отметим простейшие свойства условной вероятности и независимых событий.

1. Если $\mathbf{P}(B) > 0$, то независимость событий A и B эквивалентна $\mathbf{P}(A|B) = \mathbf{P}(A)$.

   Утверждение следует из определения.

2. Из независимости событий A и B следует независимость событий $\bar{A}$ и B.

$$\mathbf{P}(\bar{A}B) = \mathbf{P}((\Omega\backslash A)B) = \mathbf{P}(B) - \mathbf{P}(AB) = \mathbf{P}(B) - \mathbf{P}(A)\mathbf{P}(B) =$$
$$= \mathbf{P}(B)(1 - \mathbf{P}(A)) = \mathbf{P}(B)\mathbf{P}(\bar{A}).$$

3. Пусть $BC = \emptyset$, а также независимы события A и B, A и C. Тогда будут также независимы события A и $B\bigcup C$.

$$\mathbf{P}(A(B\bigcup C)) = \mathbf{P}((AB)\bigcup(AC)) = \mathbf{P}(AB) + \mathbf{P}(AC) =$$

$$= \mathbf{P}(A)\mathbf{P}(B) + \mathbf{P}(A)\mathbf{P}(C) = \mathbf{P}(A)(\mathbf{P}(B) + \mathbf{P}(C)) = \mathbf{P}(A)\mathbf{P}(B\bigcup C).$$

Определение 9.*Рассмотрим набор событий* $A_1, \dots, A_n$.

1. *События называются попарно независимыми, если*

$$\forall i \neq j \quad \mathbf{P}(A_i A_j) = \mathbf{P}(A_i)\mathbf{P}(A_j).$$

2. *События называются независимыми в совокупности, если*

$$\forall k \in \{1, \dots, n\} \, \forall i_1 \neq i_2 \neq \dots \neq i_k :$$

$$\mathbf{P}(A_{i_1} A_{i_2} \dots A_{i_k}) = \mathbf{P}(A_{i_1})\mathbf{P}(A_{i_2}) \dots \mathbf{P}(A_{i_k}).$$

Замечание. Очевидно, что из независимости в совокупности следует попарная независимость. Обратное утверждение, как будет показано ниже, неверно.

Пример Бернштейна. Возьмем тетраэдр, три грани которого покрашены в красный, синий и зелёный цвета соответственно, а на четвёртой грани присутствуют все три цвета. Эксперимент состоит в однократном подбрасывании тетраэдра, после чего наблюдается цвет (цвета) на грани, на которую упал тетраэдр. Рассмотрим события:

К= { на грани есть красный цвет },
С= { на грани есть синий цвет },
З= { на грани есть зеленый цвет }.

Каждый цвет присутствует на двух гранях (из четырех), любые два цвета одновременно и все три цвета одновременно - на одной. Поэтому

$\mathbf{P}$(К)= $\mathbf{P}$(С)= $\mathbf{P}$(З)= $\frac{1}{2}$.
$\mathbf{P}$(КС)= $\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} =$ $\mathbf{P}$(К)$\mathbf{P}$(С),
$\mathbf{P}$(КЗ)= $\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} =$ $\mathbf{P}$(К)$\mathbf{P}$(З),
$\mathbf{P}$(СЗ)= $\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} =$ $\mathbf{P}$(С)$\mathbf{P}$(З),
$\mathbf{P}$(КСЗ)= $\frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} =$ $\mathbf{P}$(К)$\mathbf{P}$(С)$\mathbf{P}$(З).

Таким образом, события К, С, З являются попарно независимыми, но не являются независимыми в совокупности.

Определение 10.*Рассмотрим набор событий* $B_1, \dots, B_n$, *удовлетворяющий следующим свойствам:*

1. $\forall k \quad \mathbf{P}(B_k) > 0$.

2. $\forall k \neq l \quad B_k B_l = \emptyset$.

3. $B_1 \bigcup B_2 \bigcup \dots \bigcup B_n = \Omega$.

Такой набор мы будем называть полной системой событий.

Теорема 3. Формула полной вероятности. *Пусть A - событие, $\{B_i\}_{i=1}^n$ - полная система событий. Тогда*

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

Доказательство. Заметим, что $\forall i \neq j\ (AB_i)(AB_j) = \emptyset$. Поэтому

$$\mathbf{P}(A) = \mathbf{P}\left(\bigcup_{i=1}^{n}(AB_i)\right) = \sum_{i=1}^{n} \mathbf{P}(AB_i) = \sum_{i=1}^{n} \mathbf{P}(A|B_i)\mathbf{P}(B_i).$$

Замечание.

$$\mathbf{P}(B_k|A) = \frac{\mathbf{P}(B_k A)}{\mathbf{P}(A)} = \frac{\mathbf{P}(AB_k)}{\mathbf{P}(A)} = \frac{\mathbf{P}(A|B_k)\mathbf{P}(B_k)}{\mathbf{P}(A)}.$$

Нами получена *Формула Байеса.* В окончательном виде она выглядит следующим образом:

$$\mathbf{P}(B_k|A) = \frac{\mathbf{P}(A|B_k)\mathbf{P}(B_k)}{\sum_{i=1}^{n} \mathbf{P}(A|B_i)\mathbf{P}(B_i)}.$$

Глава 3. Случайные величины. Законы распределения

3.1 Случайная величина

Рассмотрим пару измеримых пространств $(\Omega, \mathcal{F})$ и $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ и функцию

$$\xi = \xi(\omega) : \Omega \to \mathbb{R},$$

обладающую свойством *измеримости*, т.е.

$$\forall B \in \mathcal{B}_{\mathbb{R}} \quad \xi^{-1}(B) \in \mathcal{F}.$$

Говорят также, что ξ осуществляет *измеримое отображение*

$$\xi : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}}).$$

Определение 1.*Будем называть такую функцию $\xi = \xi(\omega)$ случайной величиной, отображающей пространство элементарных исходов Ω в множество действительных чисел $\mathbb{R}$.*

В силу измеримости каждая случайная величина порождает на измеримом пространстве $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ свою функцию множеств:

$$\forall B \in \mathcal{B}_{\mathbb{R}} : \quad \mathcal{P}_\xi(B) = \mathbf{P}(\omega : \xi(\omega) \in B) = \mathbf{P}(\xi^{-1}(B)).$$

Отметим корректность последнего равенства: вероятность $\mathbf{P}$ определена на $\{\xi^{-1}(B)\}$ для любого борелевского множества B на прямой $\mathbb{R}$ в силу измеримости случайной величины ξ.

Замечание. Будем называть числовую функцию $g : \mathbb{R} \to \mathbb{R}$ *борелевской*, если

$$\forall B \in \mathcal{B}_{\mathbb{R}} \;\Rightarrow\; g^{-1}(B) \in \mathcal{B}_{\mathbb{R}}.$$

Пусть ξ - случайная величина. Тогда

$$\forall B \in \mathcal{B}_{\mathbb{R}} \;\Rightarrow\; \{g(\xi) \in B\} = \{\xi \in g^{-1}(B)\} \in \mathcal{F},$$

т.е. $g(\xi)$ также будет случайной величиной. Рассмотрим непрерывную функцию $g : \mathbb{R} \to \mathbb{R}$, $g \in C(\mathbb{R})$. Из курса математического анализа известно, что для любого открытого множества $B \subseteq \mathbb{R}$ (т.е. B - борелевское множество) его прообраз $g^{-1}(B)$ также открыт (т.е. $g^{-1}(B)-$также борелевское множество). Из этого нетрудно получить (строгое доказательство мы опустим), что любая непрерывная функция является борелевской, т.е. для непрерывной g суперпозиция $g(\xi)$ также будет случайной величиной.

С помощью непосредственной проверки (здесь мы её опустим в силу её очевидности) легко убедиться, что индуцированная случайной величиной ξ функция множеств $\mathcal{P}_\xi : \mathcal{B}_{\mathbb{R}} \to \mathbb{R}$ удовлетворяет всем трем пунктам из определения вероятности, т.о. $\mathcal{P}_\xi$ является вероятностью на измеримом пространстве $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, а тройка $(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \mathcal{P}_\xi)$ является вероятностным пространством. Фактически каждая случайная величина порождает собственное вероятностное пространство, причем если исходное вероятностное пространство $(\Omega, \mathcal{F}, \mathbf{P})$ могло иметь любую природу, то вновь полученное связано с числовой прямой, поэтому проще для изучения.

Определение 2. *Вероятность $\mathcal{P}_\xi$ на $\mathcal{B}_{\mathbb{R}}$ называется распределением случайной величины ξ .*

Распределение случайной величины определено для любых борелевских множеств на прямой, в частности на лучах $(-\infty, x)$, $x \in \mathbb{R}$, которые также являются борелевскими множествами в силу своей открытости.

3.2 Функция распределения

Определение 3. *Функцией распределения случайной величины ξ называется функция числового аргумента*

$$F(x) = F_\xi(x) = \mathcal{P}_\xi((-\infty, x)) = \mathbf{P}(\xi < x) \, , \; x \in \mathbb{R}.$$

Замечание.В силу определения распределение случайной величины полностью определяет её функцию распределения. На самом деле (доказательство см., например, А.А. Боровков "Теория вероятностей"(1986)) верно и обратное: функция распределения случайной величины полностью определяет её распределение.

Необходимо отметить, что $F(x)$ *неслучайная величина*, она лишь характеризует то, как случайная величина X "вела себя"до "момента $x \in \mathbb{R}$"(точнее, вероятность "вести себя"до этого момента). В любом случае следует отметить, что функция распределения, вообще говоря, *не определяет полностью поведение своей случайной величины* (в частности, существуют *различные*, то есть принимающие разные значения, случайные величины, имеющие тем не менее *одинаковые* функции распределения; в принципе, это можно попробовать объяснить тем, что если сами случайные величины "ведут себя" именно "случайным образом", то такие их характеристики, как функции распределения, имеют детерминированный характер).

Теорема 1. Свойства функции распределения.*Пусть* $F(x)-$*функция распределения произвольной случайной величины. Тогда имеют место следующие свойства* $F(x)$.

1. *Монотонность:*

$$\forall x_1 \leq x_2 \ \Rightarrow \ F(x_1) \leq F(x_2).$$

2. $$\lim_{x\to-\infty} F(x) = 0 \quad , \quad \lim_{x\to+\infty} F(x) = 1.$$

3. *Непрерывность слева на* $\mathbb{R}$:

$$\forall x_0 \in \mathbb{R} : \lim_{x\to x_0-0} F(x) = F(x_0).$$

Доказательство: пусть $F(x)-$функция распределения некоей случайной величины ξ:

$$\xi : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \mathcal{P}_\xi) \ , \ F(x) = \mathbf{P}(\xi < x).$$

1. Используя свойство 3 вероятности, получим:

$$x_1 \leq x_2 \ \Rightarrow \ \{\xi < x_1\} \subseteq \{\xi < x_2\} \ \Rightarrow$$

$$\Rightarrow \ \mathcal{P}_\xi((-\infty, x_1)) \leq \mathcal{P}_\xi((-\infty, x_2)) \ \Leftrightarrow \ F(x_1) \leq F(x_2).$$

2. Рассмотрим числовые последовательности

$$\{x_n\}_n \ , \ x_{n+1} \leq x_n \ , \ \lim_{n\to\infty} x_n = -\infty$$

и

$$\{y_n\}_n \ , \ y_n \leq y_{n+1} \ , \ \lim_{n\to\infty} y_n = +\infty,$$

а также семейства множеств

$$A_n = \{\xi < x_n\} \quad , \quad B_n = \{\xi < y_n\}.$$

Заметим, что

$$A_{n+1} \subseteq A_n \ , \ \bigcap_{n=1}^{\infty} A_n = \emptyset \ , \ B_n \subseteq B_{n+1} \ , \ \bigcup_{n=1}^{\infty} B_n = \Omega.$$

В силу измеримости случайной величины $A_n \in \mathcal{F}$, $B_n \in \mathcal{F}$, поэтому на них определена вероятность $\mathbf{P}$. Тогда по свойствам вероятности (непрерывность сверху на пустом множестве и непрерывность снизу соотвественно) имеем

$$\lim_{n\to\infty} \mathbf{P}(A_n) = 0 \quad , \quad \lim_{n\to\infty} \mathbf{P}(B_n) = 1,$$

откуда, используя уже доказанную монотонность $F(x)$, окончательно получаем

$$\lim_{x\to-\infty} F(x) = 0 \quad , \quad \lim_{x\to+\infty} F(x) = 1.$$

3. Рассмотрим числовую последовательность

$$\{z_n\}_n \ , \ z_n \leq z_{n+1} \ , \ z_n < x_0 \ , \ \lim_{n\to\infty} z_n = x_0$$

и семейство множеств, удовлетворяющее свойствам

$$C_n = \{\xi < z_n\} \ , \ C_n \subseteq C_{n+1} \ , \ \bigcup_{n=1}^{\infty} C_n = \{\xi < x_0\} = C.$$

Аналогично предыдущему пукту получаем (непрерывность снизу)

$$\lim_{n\to\infty} \mathbf{P}(C_n) = \mathbf{P}(C) = F(x_0).$$

Окончательно, используя уже доказанную монотонность $F(x)$, имеем

$$\lim_{x\to x_0-0} F(x) = F(x_0).$$

Замечание.

1. Функция распределения не обязана быть непрерывной справа.

$$F(x+0)-F(x) = \lim_{n\to\infty}\left(F(x+\frac{1}{n}) - F(x)\right) = \lim_{n\to\infty}\mathbf{P}\left(\xi\in[x, x+\frac{1}{n})\right) =$$

$$= \mathbf{P}\left(\bigcap_{n=1}^{\infty}\left\{\xi\in[x, x+\frac{1}{n})\right\}\right) = \mathbf{P}(\xi = x).$$

   Т.е. функция распределения F непрерывна в точке x тогда и только тогда, когда $\mathbf{P}(\xi = x) = 0$.

2. Часто $F(x)$ определяют чуть иначе: $F_\xi(x) = \mathbf{P}(\xi \le x)$. Легко убедиться, что тогда изменится лишь третье свойство: функция распределения будет непрерывной справа на $\mathbb{R}$.

3. На самом деле имеет место следующий результат, который здесь мы приводим без доказательства: *если функция $F(x)$ обладает свойствами $1-3$, то существуют вероятностное пространство $(\Omega, \mathcal{F}, \mathbf{P})$ и случайная величина ξ на нём такие, что*

$$\forall x\in\mathbb{R}: \quad \mathbf{P}(\xi < x) = F_\xi(x) = F(x)$$

   (доказательство можно найти , например: А.А. Боровков "Теория вероятностей"(1986)). Отметим, что такие пространство и случайная величина не единственны.

3.3 Типы распределений

Дискретные распределения.
Говорят, что случайная величина ξ имеет *дискретное распределение*, если множество её значений не более чем счетно. Наглядно это выражается *рядом распределения* случайной величины.

Значения	$\cdots$	x_{-k}	x_{-k+1}	$\cdots$	x_m	x_{m+1}	$\cdots$
Вероятности	$\cdots$	p_{-k}	p_{-k+1}	$\cdots$	p_m	p_{m+1}	$\cdots$

$$-\infty < \cdots < x_{-k} < x_{-k+1} < \cdots < x_m < x_{m+1} < \cdots < \infty,$$

$$\forall i \quad p_i > 0\ ,\ \sum_{i\in\mathbb{Z}} p_i = 1\ ,\ p_i = \mathbf{P}(\xi = x_i).$$

Т.е. в верхнем ряду указаны различные конечные значения, принимаемые случайной величиной, а в нижнем - вероятности их принять (здесь случайная величина принимает счетное число значений).

Особой характеристикой функций распределений этого типа является то, что они *ступенчатые*, то есть принимают лишь конечное или счётное число значений от 0 до 1 (во вполне определённых точках; между ними значения функции *постоянны*): до первого значения, принимаемого ξ (включая его), $F(x) = 0$; от первого (исключая) до второго (включая) - значение вероятности принять первое значение; от второго (исключая) до третьего (включая) - сумма вероятностей принять первое и второе значения и т.д. В общем виде:

$$x \in (x_{k-1}, x_k] \;\Rightarrow\; F(x) = \sum_{i \le k-1} p_i,$$

$$x \in (x_k, x_{k+1}] \;\Rightarrow\; F(x) = \sum_{i \le k-1} p_i + p_k = \sum_{i \le k} p_i,$$

т.е. в точке $x = x_k$ функкция $F(x)$ (непрерывная слева) совершает "скачок" величиной

$$F(x_k + 0) - F(x_k) = p_k = \mathbf{P}(\xi = x_k).$$

Если можно вычленить наибольшее значение (обозначим его x_{max}), принимаемое случайной величиной, то получим:

$$x > x_{max} \;\Rightarrow\; F(x) = 1.$$

Приведем несколько примеров (в некоторых из них укажем ряд распределения и функцию распределения).

1. **Атомическое** (вырожденное) распределение.

   $$\xi \in I_a \in \mathbb{R}, \quad \text{то есть } \mathbf{P}(\xi = a) = 1.$$

   Таким образом, постоянные являются тоже случайными величинами. Функция распределения испытывает скачок в точке с абсциссой $x = a$.

   $$F_\xi(x) = \begin{cases} 0, & x \le a, \\ 1, & x > a. \end{cases}$$

2. **Испытание Бернулли** .

   Будем говорить, что случайная величина ξ является испытанием Бернулли, если в эксперименте, в котором есть лишь 2 исхода (0, называемый *неудачей*, и 1, называемый *успехом*), она принимает значение 1 с вероятностью $p \in (0;1)$ и 0 с вероятностью $q = 1 - p$. Будем писать $\xi \in B_p$. Ряд распределения выглядит следующим образом:

Значения	0	1
Вероятности	q	p

Функция распределения имеет два скачка: в точках с абсциссами $x = 0$ и $x = 1$.

$$F_\xi(x) = \begin{cases} 0, & x \le 0, \\ q, & x \in (0;1], \\ 1, & x > 1. \end{cases}$$

3. **Биномиальное распределение**.

$$\xi \in B(n,p), \quad n \in \mathbb{N}, \quad p \in (0,1).$$

Это распределение описывает вероятность того, что случайная величина ξ в n *одинаковых* экспериментах, в каждом из которых есть лишь 2 исхода: *успех* (обозначаемый цифрой 1) с вероятностью $p \in (0;1)$ и *неудача* (обозначаемая цифрой 0) с вероятностью $q = 1 - p$, принимает *определенное* число *успехов* $k \in \{0, 1, \ldots, n\}$. В этой модели:

$$\mathbf{P}(\xi = k) = C_n^k p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}.$$

В дальнейшем мы покажем связь между биномиально распределеннной случайной величиной и испытаниями Бернулли.

4. **Дискретное равномерное распределение**.

Ряд распределения случайной величины с таким распределением имеет вид :

Значения	0	1	$\cdots$	n
Вероятности	$\frac{1}{n+1}$	$\frac{1}{n+1}$	$\cdots$	$\frac{1}{n+1}$

5. **Геометрическое распределение**. Случайная величина принимает счетное число значений (как и в следующем пункте):

$$\mathbf{P}(\xi = k) = p^k(1-p)\ ,\ k \in \{0, 1, \ldots, n, \ldots\}.$$

Распределение описывает вероятность того, что в последовательных одинаковых испытаниях Бернулли первая неудача произошла в $k+$ 1-ом эксперименте.

6. **Пуассоновское распределение**.

$$\xi \in \Pi_\lambda, \quad \lambda > 0.$$

Это распределение характеризуется тем, что случайная величина ξ принимает целочисленные значения $n \in \{0, 1, 2, ..., n, ...\}$, при этом

$$\mathbf{P}(\xi = n) = \frac{\lambda^n}{n!} e^{-\lambda}.$$

Здесь функция распределения имеет вид суммы:

$$F_\xi(x) = \begin{cases} 0, & x \leq 0, \\ \sum_{i=0}^{k} \frac{\lambda^i}{i!} e^{-\lambda}, & x \in (k, k+1],\ k \in \mathbb{N} \bigcup \{0\}. \end{cases}$$

На самом деле выясняется, что существует достаточно много явлений, хорошо описываемых распределением Пуассона.

Более содержательные примеры дискретных распределений, встречающихся часто в теории надежности и контроле качества, можно найти в книгах Воинова и Никулина (1989, 1993, 1996), Браунли (1977), Кокса и Оукса (1988).

Абсолютно непрерывные распределения.

Рассмотрим случайную величину

$$\xi : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathbb{R}, \mathcal{B}_\mathbb{R}, \mathcal{P}_\xi)$$

такую, что

$$\exists f : \mathbb{R} \to \mathbb{R} \ : \ \forall x \in \mathbb{R}\ f(x) \geq 0 \ , \ \int_{-\infty}^{\infty} f(x) dx = 1$$

и для любого отрезка $[a, b] \subset \mathbb{R}$ вероятность $\mathbf{P}(a \leq \xi \leq b)$ события $\{a \leq \xi \leq b\}$ вычисляется по формуле

$$\mathbf{P}(a \leq \xi \leq b) = \int_a^b f(x) dx.$$

Определение 4. *Функция $f(x) = f_\xi(x)$ называется плотностью распределения вероятностей случайной величины ξ.*

Замечание. Сделаем несколько важных дополнений, которые могут быть нам полезными в дальнейшем.

1. Из свойств интеграла Римана следует, что
$$\forall x \in \mathbb{R} \quad \mathbf{P}(\xi = x) = 0,$$
откуда получаем
$$\forall a, b \in \mathbb{R} \quad \mathbf{P}(a \le \xi \le b) = \mathbf{P}(a \le \xi < b) = \mathbf{P}(a < \xi \le b) =$$
$$= \mathbf{P}(a < \xi < b) = \int_a^b f_\xi(x)dx.$$
В дальнейшем будем использовать следующее обозначение: мы будем писать $< a, b >$, если не имеет значения, включает ли промежуток от a до b в себя граничные точки a и b, или нет.

2. Естественным оказывается требование на интеграл от плотности по всей прямой:
$$\mathcal{P}_\xi(\mathbb{R}) = \mathbf{P}(\xi \in \mathbb{R}) = \int_{-\infty}^{\infty} f_\xi(x)dx = 1.$$

В данной ситуации все функции распределения *непрерывны*, более того их называют *абсолютно непрерывными* функциями, то есть каждая из них восстанавливается своей производной $F'(x)$, которая, на самом деле, и является плотностью распределения $f(x) = F'(x)$:
$$F(x) = \int_{\infty}^{x} f(t)dt.$$

Замечание. Для интегрируемой по Риману функции её интеграл *по множеству меры* 0 равен нулю (определение и простейшие примеры множеств меры 0 - см. **приложение 1**, там же - свойство *почти всюду* (*п.в.*)). Поэтому свойства плотностей и функций абсолютно непрерывных распределений имеют место на прямой *почти всюду*, т.е. с точностью до множества меры 0. В частности, определять плотность также можно *почти всюду* на прямой.

Приведём несколько примеров абсолютно непрерывных распределений ($f(x)$ - плотность, $F(x\ )$- функция распределения).

1. **Равномерное распределение** с параметрами $a, b \in \mathbb{R}$. $a < b$,
$$\xi \in U[a,b] \Leftrightarrow f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0, & x \in (-\infty, a) \cup (b, \infty) \end{cases}$$
$$F(x) = \begin{cases} 0, & x \le a \\ \frac{x-a}{b-a}, & x \in (a,b] \\ 1, & x > b \end{cases}$$

2. **Экспоненциальное распределение** с параметром $\alpha > 0$ (иногда таже говорят **показательное распределение**).

$$\xi \in \Gamma_\alpha \Leftrightarrow f(x) = \begin{cases} 0, & x \leq 0 \\ \alpha \exp(-\alpha x), & x > 0 \end{cases}$$

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - \exp(-\alpha x), & x > 0 \end{cases}$$

3. **Нормальное распределение** с параметрами $\mu \in \mathbb{R}$, $\sigma^2 > 0$ (иногда таже говорят **гауссовское распределение**).

$$\xi \in N(\mu, \sigma^2) \Leftrightarrow f_{\mu,\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ x \in \mathbb{R},$$

таким образом,

$$F_{\mu,\sigma^2}(x) = \mathbf{P}(\xi < x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

Отдельный случай, часто встречаемый на практике, так называемое *стандартное* нормальное распределение (с параметрами $\mu = 0$, $\sigma^2 = 1$)

$$\xi \in N(0,1) \Leftrightarrow f_{0,1}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

$$F_{0,1}(x) = \mathbf{P}(\xi < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

Заметим, что

$$\forall x \in \mathbb{R}: \quad F_{\mu,\sigma^2}(x) = F_{0,1}\left(\frac{x-\mu}{\sigma}\right).$$

4. **Распределение Коши** с параметрами $\mu \in \mathbb{R}$, $\sigma^2 > 0$.

$$\xi \in K(\mu, \sigma^2) \Leftrightarrow f_{\mu,\sigma^2}(x) = \frac{1}{\pi\sigma} \frac{1}{1 + \left(\frac{x-\mu}{\sigma}\right)^2}, \ x \in \mathbb{R},$$

таким образом (функцию распределения Коши обозначают $K(x)$),

$$K_{\mu,\sigma^2}(x) = \mathbf{P}(\xi < x) = \frac{1}{\pi\sigma} \int_{-\infty}^{x} \frac{1}{1 + \left(\frac{t-\mu}{\sigma}\right)^2} dt.$$

Отдельный случай, часто встречаемый на практике, так называемое *стандартное* распределение Коши (с параметрами $\mu = 0$, $\sigma^2 = 1$)

$$\xi \in K(0,1) \Leftrightarrow f_{0,1}(x) = \frac{1}{\pi}\frac{1}{1+x^2},$$

$$K_{0,1}(x) = \mathbf{P}(\xi < x) = \frac{1}{\pi}\int_{-\infty}^{x} \frac{1}{1+t^2}dt.$$

Заметим, что

$$\forall x \in \mathbb{R} \quad K_{\mu,\sigma^2}(x) = K_{0,1}\left(\frac{x-\mu}{\sigma}\right).$$

В распределениях пунктов 3 и 4 параметры μ и σ имеют смысл параметров *сдвига* и *масштаба* распределения соответственно.
Сингулярные распределения.
Данный тип распределения характеризуется тем, что функция распределения $F(x)$ непрерывна, но точки роста образуют множество меры 0.
Определение 5. *Точка x называется точкой роста функции $F(x)$, если*

$$\forall \varepsilon > 0 \quad F(x+\varepsilon) - F(x-\varepsilon) > 0.$$

Т.о., здесь $F(x)$ непрерывна на $\mathbb{R}$, $\frac{dF(x)}{dx} = 0$ почти всюду, $F(+\infty) - F(-\infty) = 1$.

Примером такой функции распределения может служить *кривая Кантора (Канторова лестница)*, все изменения которой сосредоточены на отрезке $[0,1]$(см., например, Боровков "Теория вероятностей"(1986)).
Замечание. Известна теорема Лебега: любое распределение $\mathcal{P}$ может быть *единственным образом* представлено в виде *смеси*:

$$\exists! \alpha > 0\,,\ \beta > 0\,,\ \gamma > 0\,,\ : \alpha + \beta + \gamma = 1 \quad \exists! \mathcal{P}_\alpha\,,\ \mathcal{P}_\beta\,,\ \mathcal{P}_\gamma : \forall B \in \mathcal{B}_{\mathbb{R}} \Rightarrow$$

$$\Rightarrow \mathcal{P}(B) = \alpha\mathcal{P}_\alpha(B) + \beta\mathcal{P}_\beta(B) + \gamma\mathcal{P}_\gamma(B).$$

Здесь $\mathcal{P}_\alpha, \mathcal{P}_\beta, \mathcal{P}\gamma$ - распределения дискретного, абсолютно непрерывного и сингулярного типов соответственно. В терминах функций распределения, используя аналогичные обозначения, запишем (говорят о представлении функции распределения в виде суммы трех *компонент* - дискретной, абсолютно непрерывной и сингулярной):

$$\forall x \in \mathbb{R} \Rightarrow F(x) = \alpha F_\alpha(x) + \beta F_\beta(x) + \gamma F_\gamma(x).$$

Из трёх компонент разрывна лишь одна - дискретная, причем число точек разрыва не более чем счётно. Отсюда следует, что любая функция распределения имеет не более чем счётное число точек разрыва.

Глава 4. Многомерные случайные величины. Независимость случайных величин

4.1 Основные определения

По аналогии с борелевской σ-алгеброй на прямой $\mathcal{B}_{\mathbb{R}}$ введём борелевскую σ-алгебру в пространстве $\mathbb{R}^n$: $\mathcal{B}^n_{\mathbb{R}} = \mathcal{B}^n$ – σ-алгебра, порожденная открытыми множествами в $\mathbb{R}^n$ (на самом деле она совпадает с σ-алгеброй, порожденной множествами $B = B_1 \times ... \times B_n$, $B_i \in \mathcal{B}_{\mathbb{R}}$).
Определение 1. *Рассмотрим случайные величины $\xi_1, \xi_2, ..., \xi_n$, заданные на одном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbf{P})$. Каждому $\omega \in \Omega$ эти случайные величины ставят в соответствие n-мерный вектор*

$$\bar{\xi}(\omega) = (\xi_1(\omega), \xi_2(\omega), ..., \xi_n(\omega)).$$

Отображение $\bar{\xi} : \Omega \to \mathbb{R}^n$, задаваемое $\xi_1, \xi_2, ..., \xi_n$, называется случайным вектором или многомерной случайной величиной.
Замечание. Отображение $\bar{\xi} : \Omega \to \mathbb{R}^n$ следует рассматривать как измеримое отображение пространства $(\Omega, \mathcal{F})$ на пространство $(\mathbb{R}^n, \mathcal{B}^n)$. Поэтому для любого борелевского множества $B \in \mathcal{B}^n$ определена функция

$$\mathcal{P}_{\bar{\xi}}(B) = \mathbf{P}(\bar{\xi} \in B) : \mathcal{B}^n \to [0, 1].$$

Очевидно что эта функция множеств $\mathcal{P}_{\bar{\xi}}$ является вероятностью на $(\mathbb{R}^n, \mathcal{B}^n)$, поэтому можно сказать (по аналогии с одномерным случаем), что случайный вектор $\bar{\xi}$ порождает вероятностное пространство $(\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}_{\bar{\xi}})$.
Определение 2 .

1. *Функция $\mathcal{P}_{\bar{\xi}}$ называется распределением случайного вектора $\bar{\xi}$.*

2. *Функция*

$$F_{\bar{\xi}}(\bar{x}) = F_{\xi_1,...,\xi_n}(x_1, ..., x_n) = \mathbf{P}(\xi_1 < x_1, ..., \xi_n < x_n),$$

$$\bar{x} = (x_1, ..., x_n) \in \mathbb{R}^n$$

называется функцией распределения случайного вектора $\bar{\xi} = (\xi_1, \xi_2, ..., \xi_n)$ (или функцией совместного распределения случайных величин $\xi_1, \xi_2, ..., \xi_n$).

Приведем свойства функции распределения случайного вектора $\bar{\xi} = (\xi_1, \xi_2, ..., \xi_n)$, называющиеся *свойствами согласованности*:

1. $$\lim_{x_n \to \infty} F_{\xi_1,...,\xi_n}(x_1, ..., x_n) = F_{\xi_1,...,\xi_{n-1}}(x_1, ..., x_{n-1}).$$

2. $$\lim_{x_k \to -\infty} F_{\xi_1,...,\xi_n}(x_1, ..., x_n) = 0\ ,\ k \in \{1, ..., n\}.$$

В первом пункте несущественно, что предел рассматривается по последнему аргументу: если устремить к $+\infty$ произвольную k-ую координату, то в пределе получим функцию распределения *подвектора* (без этой k-ой координаты)

$$F_{\xi_1,...,\xi_{k-1},\xi_{k+1},...,\xi_n}(x_1, ..., x_{k-1}, x_{k+1}, ..., x_n).$$

Во втором пункте предел оказывается равным нулю вне зависимости от того, какая из координат устремляется к $-\infty$.

Если посмотреть на свойства функции распределения (зависящей от одной переменной) случайной величины (пункты 1—3), то можно заметить, что *свойства согласованности* являются аналогом пункта 2 и доказываются так же. Монотонность и непрерывность слева по любому аргументу (аналоги пунктов 1 и 3 соответственно) тоже проверяются аналогичнымспособом, поэтому эти доказательства мы опустим. Как и в одномерном случае сделаем

Замечание.

1. Верен следующий факт: каждая функция $F(x_1, ..., x_n)$ с набором указанных выше свойств будет функцией распределения некоторого случайного вектора.

2. По определению функция распределения случайного вектора полностью определяется его распределением. Верно и обратное : функция распределения $F_{\bar{\xi}} = F_{\xi_1,...,\xi_n}$ однозначно определяет распределение $\mathcal{P}_{\bar{\xi}}$.

4.2 Основные типы распределений

Аналогично одномерному случаю рассмотрим различные типы распределений многомерных случайных величин (здесь мы ограничимся лишь дискретным и абсолютно непрерывным распределениями).

Дискретные распределения.

Будем говорить, что случайный вектор имеет дискретное распределение, если он принимает не более чем счетное число значений.

Пример. Полиномиальное распределение.
Рассмотрим вероятностное пространство $(\Omega, \mathcal{F}, \mathbf{P})$, полную систему событий в нём $\{A_i\}_{i=1}^m$, $m \geq 2$, фиксированное $n \in \mathbb{N}$ и вектор вероятностей

$$\bar{p} = (p_1, ..., p_m) \ , \ p_i > 0 \ , \ \sum_{i=1}^m p_i = 1 \ : \ \mathbf{P}(A_i) = p_i.$$

Говорят, что целочисленный вектор $\bar{\nu} = (\nu_1, ..., \nu_m)$ имеет полиномиальное распределение с параметрами n и $\bar{p}$, если:

$$\bar{\nu} \in B(n, \bar{p}) \ \Leftrightarrow \ \forall \bar{k} = (k_1, ..., k_m) \ : \ k_i \in \mathbb{N} \bigcup \{0\} \ , \ \sum_{i=1}^m k_i = n \ \Rightarrow$$

$$\mathbf{P}(\bar{\nu} = \bar{k}) = \frac{n!}{k_1! \cdots k_m!} p_1^{k_1} \cdots p_m^{k_m}.$$

Представим себе описанное распределение следующим образом: независимо друг от друга проводятся n одинаковых экспериментов. Результатом каждого из них является появление *одного из m событий* $A_1, ..., A_m$ (вероятность появления A_i в каждом испытании равна p_i), а координаты случайного вектора $\bar{\nu}$ можно предсталять себе как количество появлений соответствующих событий в n проведенных экспериментах. Если $m = 2$, то мы получаем биномиальное распределение.

Абсолютно непрерывные распределения.

Рассмотрим случайный вектор

$$\bar{\xi} : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}_{\bar{\xi}})$$

такой, что

$$\exists f : \mathbb{R}^n \to \mathbb{R} \ : \ \forall \bar{x} \in \mathbb{R}^n \ f(\bar{x}) \geq 0 \ , \ \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(x_1, ..., x_n) dx_1 ... dx_n = 1$$

и для любого замкнутого параллелепипеда $\Pi_n = [a_1, b_1] \times ... \times [a_n, b_n] \subset \mathbb{R}^n$ вероятность $\mathbf{P}(\bar{\xi} \in \Pi_n)$ вычисляется по формуле

$$\mathbf{P}(\bar{\xi} \in \Pi_n) = \mathcal{P}_{\bar{\xi}}(\Pi_n) = \int_{a_1}^{b_1} ... \int_{a_n}^{b_n} f(x_1, ..., x_n) dx_1 ... dx_n.$$

Определение 3. *Функция* $f(\bar{x}) = f_{\bar{\xi}}(\bar{x}) = f_{\xi_1,...,\xi_n}(x_1,...,x_n)$ *называется плотностью распределения вероятностей случайного вектора* $\bar{\xi}$.

Имеют место аналоги замечаний, данных для одномерного случая. Например, если $\Pi_n^* =< a_1, b_1 > \times ... \times < a_n, b_n >$, то

$$\mathcal{P}_{\bar{\xi}}(\Pi_n) = \mathcal{P}_{\bar{\xi}}(\Pi_n^*) = \int_{a_1}^{b_1} ... \int_{a_n}^{b_n} f(x_1,...,x_n)dx_1...dx_n,$$

$$\mathcal{P}_{\bar{\xi}}(\mathbb{R}^n) = \mathbf{P}(\bar{\xi} \in \mathbb{R}^n) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f_{\bar{\xi}}(x_1,...,x_n)dx_1...dx_n = 1,$$

$$F_{\xi_1,...,\xi_n}(x_1,...,x_n) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} f_{\xi_1,...,\xi_n}(t_1,...,t_n)dt_1...dt_n,$$

$$\frac{\partial^n F_{\xi_1,...,\xi_n}(x_1,...,x_n)}{\partial x_1...\partial x_n} = f_{\xi_1,...,\xi_n}(x_1,...,x_n).$$

Отметим очевидный факт: если случайный вектор $\bar{\xi} = (\xi_1,...,\xi_n)$ имеет плотность $f(x_1,...,x_n)$, то любой *подвектор* $\bar{\xi}^{(m)} = (\xi_{k_1},...,\xi_{k_m}), k_m < n$ также имеет плотность, равную (пусть для простоты $k_i = i$, $i \in \{1,...,m\}$)

$$f(x_1,...,x_m) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(x_1,...,x_n)dx_{m+1}...dx_n.$$

В частности, абсолютно непрерывными будут все координаты случайного вектора ($m = 1$).

Также нужно отметить, что свойства плотностей и функций абсолютно непрерывных распределений имеют место в $\mathbb{R}^n$ *почти всюду*, т.е. с точностью до множества меры 0. В частности, определять плотность в $\mathbb{R}^n$ также можно *почти всюду*.

Пример. Многомерное нормальное распределение.
Пусть

$$\bar{\alpha} = (\alpha_1,...,\alpha_n) \, , \; \alpha_i \in \mathbb{R} \, , \; \Sigma^2 = \|\sigma_{ij}\|_{i=1,j=1}^n,$$

где Σ^2- симметричная положительно определенная матрица. Пусть также

$$A = \|a_{ij}\|_{i=1,j=1}^n \, , \; A^{-1} = \Sigma^2.$$

Случайный вектор $\bar{\xi} = (\xi_1,...,\xi_n)$ имеет нормальное распределение с параметрами $\bar{\alpha}$ и Σ^2

$$\bar{\xi} \in N(\bar{\alpha}, \Sigma^2),$$

если это распределение имеет плотность

$$\varphi_{\bar{\alpha},\Sigma^2}(\bar{x}) = \frac{\sqrt{\det A}}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}(\bar{x}-\bar{\alpha})A(\bar{x}-\bar{\alpha})^T\right\}.$$

4.3 Независимость случайных величин

Определение 4.

1. *Случайные величины* $\xi_1, ..., \xi_n$*, заданные на одном вероятностном пространстве* $(\Omega, \mathcal{F}, \mathbf{P})$*, называются независимыми, если*

$$\forall B_1, ..., B_n \in \mathcal{B}_{\mathbb{R}} \Rightarrow \mathbf{P}(\xi_1 \in B_1, ..., \xi_n \in B_n) = \mathbf{P}(\xi_1 \in B_1) \dots \mathbf{P}(\xi_n \in B_n).$$

2. *Случайные величины последовательности* $\{\xi_n\}_{n=1}^{\infty}$*, заданные на одном вероятностном пространстве* $(\Omega, \mathcal{F}, \mathbf{P})$*, называются независимыми, если любой конечный набор из этой последовательности состоит из независимых случайных величин.*

Нашей целью является получение каких либо критериев независимости случайных величин, желательно в терминах функций распределения. Для этого нам понадобятся дополнительные утверждения.
Определение 5. *Рассмотрим вероятностное пространство* $(\Omega, \mathcal{F}, \mathbf{P})$.

1. *Рассмотрим алгебры событий* $\mathcal{A}_1$ *и* $\mathcal{A}_2$*, содержащиеся в* σ*-алгебре* $\mathcal{F}$ *.Говорят, что алгебры событий* $\mathcal{A}_1$ *и* $\mathcal{A}_2$ *независимы, если*

$$\forall A_1 \in \mathcal{A}_1 \, \forall A_2 \in \mathcal{A}_2 \ \Rightarrow \ \mathbf{P}(A_1 A_2) = \mathbf{P}(A_1)\mathbf{P}(A_2).$$

2. *Говорят, что алгебры событий* $\{\mathcal{A}_n\}_{n\in\mathbb{N}}$ *независимы, если*

$$\forall k \in \mathbb{N} \forall n_1 \neq \dots \neq n_k \forall A_{n_i} \in \mathcal{A}_{n_i} \Rightarrow \mathbf{P}\left(\bigcap_{i=1}^{k} A_{n_i}\right) = \mathbf{P}(A_{n_1}) \dots \mathbf{P}(A_{n_k}).$$

Замечание. На самом деле верен важный, интуитивно очевидный результат, доказательство которого мы здесь опустим (см., например, Боровков "Теория вероятностей" (1986)): σ-алгеры, порожденные независимыми классами событий, также будут независимы.
Определение 6. *Рассмотрим вероятностное пространство* $(\Omega, \mathcal{F}, \mathbf{P})$ *и случайную величину* ξ *на нём .Назовем* σ*-алгеброй, порождённой случайной величиной* ξ*,* σ*-алгебру* $\mathcal{F}_\xi = \sigma(\xi) \subseteq \mathcal{F}$*, порождённую классом событий* $\mathcal{A}$*, состоящим из множеств :*

$$A = \xi^{-1}(B) = \{\omega : \ \xi(\omega) \in B\} \in \mathcal{F} \quad \forall B \in \mathcal{B}_{\mathbb{R}}.$$

Таким образом $\mathcal{F}_\xi = \sigma(\xi) = \sigma(\mathcal{A})$.
(Точно так же определяется σ-алгебра, порожденная случайным вектором).

Используя данные выше определения и замечания, сформулируем очевидный вывод, связывающий независимость случайных величин (заданных на одном вероятностном пространстве) и σ-алгебр: *для независимости случайных величин* $\xi_1, ..., \xi_n$ *необходимо и достаточно, чтобы были независимы порожденные ими* σ*-алгебры* $\sigma(\xi_1), ..., \sigma(\xi_n)$ (доказательство см., например: Боровков "Теория вероятностей" (1986)).

Теорема 1. Критерий независимости случайных величин в терминах функций распределения . *Случайные величины* $\xi_1, ..., \xi_n$*, заданные на одном вероятностном пространстве* $(\Omega, \mathcal{F}, \mathbf{P})$*, независимы тогда и только тогда, когда*

$$F_{\xi_1,...,\xi_n}(x_1, ..., x_n) = F_{\xi_1}(x_1) \dots F_{\xi_n}(x_n),$$

где $F_{\xi_1,...,\xi_n}(x_1, ..., x_n)$ *- совместная функция распределения случайных величин* $\xi_1, ..., \xi_n$*,* $F_{\xi_i}(x_i)$ *- функция распределения случайной величины* ξ_i.

Доказательство. В одну сторону утверждение теоремы очевидно, так что достаточно убедиться, что из равенства для совместной функции распределения

$$F_{\xi_1,...,\xi_n}(x_1, ..., x_n) = F_{\xi_1}(x_1) \dots F_{\xi_n}(x_n)$$

следует независимость случайных величин $\xi_1, ..., \xi_n$. Пусть для простоты (для краткости записей) $n = 2$. Введем обозначения

$$\Delta_1 = [x_1, x_2) \ , \ \Delta_2 = [y_1, y_2) \ , \ x_i, y_j \in \mathbb{R}.$$

В данных обозначениях имеем:

$$\mathbf{P}(\xi_1 \in \Delta_1, \xi_2 \in \Delta_2) = \mathbf{P}(x_1 \le \xi_1 < x_2, y_1 \le \xi_2 < y_2) =$$

$$= \mathbf{P}(\xi_1 < x_2, y_1 \le \xi_2 < y_2) - \mathbf{P}(\xi_1 < x_1, y_1 \le \xi_2 < y_2) =$$

$$= \{F_{\xi_1,\xi_2}(x_2, y_2) - F_{\xi_1,\xi_2}(x_2, y_1)\} - \{F_{\xi_1,\xi_2}(x_1, y_2) - F_{\xi_1,\xi_2}(x_1, y_1)\} =$$

$$= F_{\xi_1,\xi_2}(x_2, y_2) - F_{\xi_1,\xi_2}(x_1, y_2) - F_{\xi_1,\xi_2}(x_2, y_1) + F_{\xi_1,\xi_2}(x_1, y_1) =$$

$$= (F_{\xi_1}(x_2) - F_{\xi_1}(x_1)) \cdot (F_{\xi_2}(y_2) - F_{\xi_2}(y_1)) = \mathbf{P}(\xi_1 \in \Delta_1)\mathbf{P}(\xi_2 \in \Delta_2).$$

Тогда для двух наборов полуинтервалов

$$\{\Delta_1^{(i)}\}_{i=1}^n \ , \ \Delta_1^{(i)} = [x_1^{(i)}, x_2^{(i)}) \subset \mathbb{R} \ , \ \Delta_1^{(k)}\Delta_1^{(l)} = \emptyset, \ k \ne l,$$

$$\{\Delta_2^{(j)}\}_{j=1}^m \ , \ \Delta_2^{(j)} = [y_1^{(j)}, y_2^{(j)}) \subset \mathbb{R} \ , \ \Delta_2^{(k)}\Delta_2^{(l)} = \emptyset, \ k \ne l$$

имеем

$$\mathbf{P}\left(\xi_1 \in \bigcup_{i=1}^{n} \Delta_1^{(i)}, \xi_2 \in \bigcup_{j=1}^{m} \Delta_2^{(j)}\right) = \sum_{i,j} \mathbf{P}(\xi_1 \in \Delta_1^{(i)}, \xi_2 \in \Delta_2^{(j)}) =$$

$$= \sum_{i,j} \mathbf{P}(\xi_1 \in \Delta_1^{(i)})\mathbf{P}(\xi_2 \in \Delta_2^{(j)}) = \mathbf{P}\left(\xi_1 \in \bigcup_{i=1}^{n} \Delta_1^{(i)}\right) \mathbf{P}\left(\xi_2 \in \bigcup_{j=1}^{m} \Delta_2^{(j)}\right).$$

Теперь введем классы событий $\mathcal{A}_1$ и $\mathcal{A}_2$, состоящие из всевозможных событий вида $\{\xi_1 \in \Delta_1\} \subseteq \mathcal{F}$ и $\{\xi_2 \in \Delta_2\} \subseteq \mathcal{F}$ соответственно. Мы только что показали независимость этих классов, но тогда независимы порожденные ими σ-алгебры $\sigma(\mathcal{A}_1)$ и $\sigma(\mathcal{A}_1)$. Осталось заметить, что $\sigma(\xi_1) = \sigma(\mathcal{A}_1)$, $\sigma(\xi_2) = \sigma(\mathcal{A}_2)$. Т.о., независимы σ-алгебры, порожденные случайными величинами ξ_1 и ξ_2 , что эквивалентно независимости случайных величин ξ_1 и ξ_2 .

Следствие. *Пусть случайные величины $\xi_1, ..., \xi_n$, заданные на одном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbf{P})$, абсолютно непрерывно распределены с плотностями $f_1(x), ..., f_n(x)$ соответственно. Тогда для независимости $\xi_1, ..., \xi_n$ необходимо и достаточно, чтобы случайный вектор $\bar{\xi} = (\xi_1, ..., \xi_n)$ был абсолютно непрерывно распределен с плотностью*

$$f(x_1, \ldots, x_n) = f_1(x_1) \ldots f_n(x_n).$$

Доказательство.

1. **Необходимость.** Пусть функция распределения случайной величины ξ_i равна

$$F_{\xi_i}(x_i) = \int_{-\infty}^{x_i} f_i(t_i) dt_i,$$

и $\xi_1, ..., \xi_n$ - независимы. Тогда их совместная функция распределения будет следующей:

$$F_{\xi_1,\ldots\xi_n}(x_1, ..., x_n) = F_{\xi_1}(x_1) \ldots F_{\xi_n}(x_n) =$$

$$= \int_{-\infty}^{x_1} f_1(t_1)dt_1 \ldots \int_{-\infty}^{x_n} f_n(t_n)dt_n = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} f_1(t_1)...f_n(t_n)dt_1...dt_n.$$

Необходимость доказана.

Достаточность. Пусть функция распределения случайного вектора $\bar{\xi}$ имеет вид

$$F_{\xi_1,\ldots\xi_n}(x_1, ..., x_n) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} f_1(t_1)...f_n(t_n)dt_1...dt_n.$$

Тогда, в силу равенства

$$\int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} f_1(t_1)...f_n(t_n)dt_1...dt_n = \int_{-\infty}^{x_1} f_1(t_1)dt_1 \ldots \int_{-\infty}^{x_n} f_n(t_n)dt_n,$$

получим

$$F_{\xi_1,...\xi_n}(x_1, ..., x_n) = F_{\xi_1}(x_1) \ldots F_{\xi_n}(x_n).$$

Достаточность доказана.

Теорема 2. *Пусть случайные величины $\xi_1, ..., \xi_n$, заданные на одном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbf{P})$, независимы. Рассмотрим борелевские функции $g_1(x), ..., g_n(x)$. Тогда случайные величины $g_1(\xi_1), ..., g_n(\xi_n)$ также будут независимы.*

Доказательство. Пусть для простоты (для краткости) $n = 2$.Заметим, что по свойствам функций $g_i(x)$ имеем

$$\forall i \, \forall B_i \in \mathcal{B}_{\mathbb{R}} \Rightarrow \{x \in \mathbb{R} : \, g_i(x) \in B_i\} = g_i^{-1}(B_i) = B_i^* \in \mathcal{B}_{\mathbb{R}}.$$

Тогда

$$\forall i \, \forall B_i \in \mathcal{B}_{\mathbb{R}} \Rightarrow \{\omega \in \Omega : \, g_i(\xi_i(\omega)) \in B_i\} = \{\omega \in \Omega : \, \xi_i(\omega) \in B_i^*\}.$$

Отсюда получаем требуемый результат

$$\mathbf{P}(g_1(\xi_1) \in B_1, g_2(\xi_2) \in B_2) = \mathbf{P}(\xi_1 \in B_1^*, \xi_2 \in B_2^*) =$$

$$= \mathbf{P}(\xi_1 \in B_1^*)\mathbf{P}(\xi_2 \in B_2^*) = \mathbf{P}(g_1(\xi_1) \in B_1)\mathbf{P}(g_2(\xi_2) \in B_2).$$

Пусть случайные величины $\xi_1, ..., \xi_n$ заданы на одном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbf{P})$. В дальнейшем (для краткости) не станем это специально оговаривать и будем использовать следующие аббревиатуры: вместо словосочетания *независимые случайные величины* будем писать НСВ; если же вдобавок эти случайные величины имеют одинаковые функции распределения (*одинаково распределены*) , будем писать НОРСВ. То же будет относиться и к бесконечным последовательностям случайных величин.

4.4 Примеры

1. **Биномиальное распределение**. Пусть $\xi_1, ..., \xi_n$ - НОРСВ, где ξ_i - испытание Бернулли с вероятностью успеха $p \in (0, 1)$ (вероятность неудачи $q = 1 - p$). Введем новую случайную величину

$$S_n = \xi_1 + ... + \xi_n.$$

Заметим, что S_n покажет нам количество значений ξ_i, оказавшихся равными 1, т.е. *количество успехов в n испытаниях Бернулли.* Т.о.

$$S_n \in \{0, 1, ..., n\}.$$

Пусть ровно $k \in \{0, 1, ..., n\}$ испытаний Бернулли оказались успешными, причем нам известны их номера. Для простоты пусть это будут первые k испытаний. Используя независимость, получим

$$\mathbf{P}(\xi_1 = 1, ..., \xi_k = 1, \xi_{k+1} = 0, ..., \xi_n = 0) =$$

$$= \mathbf{P}(\xi_1 = 1)...\mathbf{P}(\xi_k = 1)\mathbf{P}(\xi_{k+1} = 0)...\mathbf{P}(\xi_n = 0) = p^k q^{n-k}.$$

Пусть $A_{i_1,...,i_k}$, $1 \le i_1 < ... < i_k \le n$ - событие, заключающееся в том, что испытания с номерами $i_1, ..., i_k$ окончились успехом, остальные - неудачей. Очевидно, что

$$\mathbf{P}(A_{i_1,...,i_k}) = \mathbf{P}(A_{1,...,k}) = p^k q^{n-k},$$

а также то, что эти события попарно несовместны:

$$A_{i_1,...,i_k} A_{j_1,...,j_k} = \emptyset\ ,\ \{i_1, ..., i_k\} \ne \{j_1, ..., j_k\}.$$

Всего существует C_n^k различных событий $A_{i_1,...,i_k}$. Отсюда получаем (событие $\{S_n = k\}$ произойдет, если произойдет одно из событий $A_{i_1,...,i_k}$):

$$\mathbf{P}(S_n = k) = \mathbf{P}\left(\bigcup_{i_1,...,i_k} A_{i_1,...,i_k}\right) = \sum_{i_1,...,i_k} \mathbf{P}(A_{i_1,...,i_k}) = C_n^k p^k q^{n-k}.$$

Т.о. мы получили, что

$$S_n = \xi_1 + ... + \xi_n \in B(n, p),$$

иначе говоря, *биномиально распределенная случайная величина может быть представлена суммой независимых и одинаково распределенных испытаний Бернулли с той же вероятностью успеха в каждом испытании.*

В заключение установим полезный факт, касающийся поведения вероятностей $\mathbf{P}(S_n = k)$, $k \in \{0, 1, ..., n-1\}$, при фиксированном количестве испытаний n.

$$\mathbf{P}(S_n = k) = C_n^k p^k q^{n-k} \ \Rightarrow\ \frac{\mathbf{P}(S_n = k+1)}{\mathbf{P}(S_n = k)} = \frac{(n-k)p}{(k+1)q}.$$

Из этого равенства получаем:

$$\mathbf{P}(S_n = k+1) > \mathbf{P}(S_n = k) \;\Leftrightarrow\; k < np - q,$$

$$\mathbf{P}(S_n = k+1) = \mathbf{P}(S_n = k) \;\Leftrightarrow\; k = np - q,$$

$$\mathbf{P}(S_n = k+1) < \mathbf{P}(S_n = k) \;\Leftrightarrow\; k > np - q,$$

т.е. с ростом k вероятность $\mathbf{P}(S_n = k)$ возрастает, достигает максимума, а потом убывает. Рассмотрим два случая (пока полагаем $np{-}q > 0$):

(a) Пусть $np{-}q$ является натуральным числом. Тогда максимальное значение вероятности $\mathbf{P}(S_n = k)$ достигается при двух значениях k:

$$k_0^{(1)} = np - q\ ,\ k_0^{(2)} = np - q + 1 = np + p.$$

(b) Пусть $np{-}q$ не является натуральным числом. Тогда максимальное значение вероятности $\mathbf{P}(S_n = k)$ достигается при значении $k = k_0$, равному наименьшему целому числу, большему $k_0^{(1)}$:

$$k_0 = [np - q] + 1 = [np + p].$$

Осталось лишь рассмотреть случай $np{-}q \le 0$. Совершенно очевидно, что

$$np{-}q = 0 \Rightarrow \mathbf{P}(S_n = 0) = \mathbf{P}(S_n = 1) > \mathbf{P}(S_n = 2) > ... > \mathbf{P}(S_n = n),$$

$$np{-}q < 0 \Rightarrow \mathbf{P}(S_n = 0) > \mathbf{P}(S_n = 1) > \mathbf{P}(S_n = 2) > ... > \mathbf{P}(S_n = n).$$

2. **Порядковые статистики**. Пусть $\xi_1, ..., \xi_n$ - НОРСВ. Построим по ним новые случайные величины. Для каждого элементарного события ω переставим значения $\xi_1(\omega), ..., \xi_n(\omega)$ в порядке возрастания, заодно переименовав их:

$$\omega \to \xi_1(\omega), \xi_2(\omega), ..., \xi_n(\omega) \to \xi_{(1)}(\omega) \le \xi_{(2)}(\omega) \le ... \le \xi_{(n)}(\omega).$$

Мы построили новый набор случайных величин

$$\xi_{(1)}, \xi_{(2)}, ..., \xi_{(n)} : \Omega \to \mathbb{R}\ ,\ \xi_{(1)} \le \xi_{(2)} \le ... \le \xi_{(n)},$$

называемых *порядковыми статистиками*. Вектор, составленный из них в порядке возрастания, называется *вектором порядковых статистик*. Случайную величину $\xi_{(1)}$ будем называть *наименьшей порядковой статистикой*, а $\xi_{(n)}$ будем называть *наибольшей порядковой статистикой*. Оказывается, что построенные по набору НОРСВ порядковые статистики не будут ни независимыми, ни одинаково распределенными. Для доказательства этого факта нам достаточно будет рассмотрения *наименьшей и наибольшей порядковых статистик.*

Введем обозначения:

$$F(x) = \mathbf{P}(\xi_i < x) \forall i \in \{1, ..., n\}, F_1(x) = \mathbf{P}(\xi_{(1)} < x), F_n(x) = \mathbf{P}(\xi_{(n)} < x).$$

Найдем функции распределения *наибольшей порядковой статистики* $(F_n(x))$ и *наименьшей порядковой статистики* $(F_1(x))$.

$$F_n(x) = \mathbf{P}(\xi_{(n)} < x) = \mathbf{P}(\xi_{(1)} < x, \xi_{(2)} < x, ..., \xi_{(n)} < x) =$$

$$= \mathbf{P}(\xi_1 < x, \xi_2 < x, ..., \xi_n < x) = \mathbf{P}(\xi_1 < x)\mathbf{P}(\xi_2 < x) \dots \mathbf{P}(\xi_n < x) = F^n(x).$$

Второе равенство следует из того, что если *наибольшая порядковая статистика* принимает значения, меньшие x, то остальные порядковые статистики, её не превосходящие, тоже принимают значения, меньшие x. Третье равенство следует из того, что если все переставленные в порядке возрастания значения $\xi_1(\omega), \xi_2(\omega), ..., \xi_n(\omega)$ оказались меньше x, то и до перестановки они были меньше x. Далее используются независимость и одинаковая распределенность $\xi_1, ..., \xi_n$.

Теперь найдем функцию распределения *наименьшей порядковой статистики* $(F_1(x))$:

$$F_1(x) = \mathbf{P}(\xi_{(1)} < x) = 1 - \mathbf{P}(\xi_{(1)} \geq x) =$$

$$= 1 - \mathbf{P}(\xi_{(1)} \geq x, \xi_{(2)} \geq x, \dots, \xi_{(n)} \geq x) = 1 - \mathbf{P}(\xi_1 \geq x, \xi_2 \geq x, \dots, \xi_n \geq x) =$$
$$= 1 - \mathbf{P}(\xi_1 \geq x)\mathbf{P}(\xi_2 \geq x) \dots \mathbf{P}(\xi_n \geq x) = 1 -$$
$$- (1 - \mathbf{P}(\xi_1 < x))(1 - \mathbf{P}(\xi_2 < x)) \dots (1 - \mathbf{P}(\xi_n < x)) = 1 - (1 - F(x))^n.$$

Во втором равенстве мы перешли к дополнительному событию. Третье равенство следует из того, что если *наименьшая порядковая статистика* принимает значения, большие или равные x, то остальные порядковые статистики, её не меньшие, также будут принимать значения, лежащие правее x. Все дальнейшие рассуждения полностью аналогичны приведённым выше.

Итак, мы показали, что *наименьшая и наибольшая порядковые статистики* имеют разные распределения. Покажем, что они не являются независимыми. Пусть

$$F_{1,n}(x, y) = \mathbf{P}(\xi_{(1)} < x, \xi_{(n)} < y)$$

их совместная функция распределения. Очевидно, что если $x > y$, то

$$F_{1,n}(x,y) = F_{1,n}(y,y) = \mathbf{P}(\xi_{(1)} < y, \xi_{(2)} < y, ..., \xi_{(n)} < y) = F^n(y).$$

Первое равенство следует из того, что *наименьшая порядковая статистика* не может быть больше *наибольшей порядковой статистики*, далее рассуждения аналогичны приведённым выше. Пусть теперь $x \le y$:

$$F_{1,n}(x,y) = \mathbf{P}(\xi_{(1)} < x, \xi_{(n)} < y) =$$

$$\mathbf{P}(\xi_{(1)} < y, \xi_{(n)} < y) - \mathbf{P}(\xi_{(1)} \in [x,y), \xi_{(n)} < y) =$$

$$= F^n(y) - \mathbf{P}(\xi_{(1)} \in [x,y), \xi_{(2)} \in [x,y), ..., \xi_{(n)} \in [x,y)) =$$

$$= F^n(y) - \mathbf{P}(\xi_1 \in [x,y), \xi_2 \in [x,y), ..., \xi_n \in [x,y)) = F^n(y) - (F(y) - F(x))^n.$$

В итоге

$$F_{1,n}(x,y) = \begin{cases} F^n(y) - (F(y) - F(x))^n, & x \le y \\ F^n(y), & x > y. \end{cases}$$

Для независимости необходимо и достаточно выполнение равенства

$$F_{1,n}(x,y) = F_1(x)F_n(y)\ ,\ x \in \mathbb{R}\ ,\ y \in \mathbb{R},$$

где

$$F_1(x)F_n(y) = (1 - (1 - F(x))^n)\, F^n(y).$$

Т.о. *наименьшая и наибольшая порядковые статистики* не являются независимыми.

Замечание .

1. Используя похожие рассудения, нетрудно получить формулу для функции распределения произвольной порядковой статистики:

$$F_k(x) = \mathbf{P}(\xi_{(r)} < x) = \sum_{k=r}^{n} C_n^k F^k(x)(1 - F(x))^{n-k}.$$

2. Пусть исходные случайные величины $\xi_1, ..., \xi_n$ абсолютно непрерывны, и плотность произвольной из них равна $f(x) = F'(x)$. Тогда все порядковые статистики также абсолютно непрерывны, и плотность, например, *наибольшей порядковой статистики* равна

$$f_n(x) = F_n'(x) = nF^{n-1}(x)f(x).$$

Очевидно также, что в это случае

$$\mathbf{P}(\xi_{(1)} < \xi_{(2)} < ... < \xi_{(n)}) = 1.$$

3. **Функция распределения суммы**. Рассмотрим случайные величины $\xi_1, ..., \xi_n$, заданные на одном вероятностном пространстве и пусть случайный вектор $\bar{\xi} = (\xi_1, ..., \xi_n)$ абсолютно непрерывен с плотностью $f(x_1, ..., x_n)$. Требуется найти функцию распределения $F(x)$ суммы

$$\eta = \xi_1 + ... + \xi_n.$$

Искомая функция равна вероятности попадания точки $(\xi_1, ..., \xi_n)$ в полупространство $\xi_1 + ... + \xi_n < x$ и, следовательно,

$$\mathbf{P}(\eta < x) = F(x) = \int ... \int_{\sum_{k=1}^{n} x_k < x} f(x_1, ..., x_n) dx_1 ... dx_n.$$

Для простоты рассмотрим случай $n = 2$. Имеем:

$$F(x) = \int\int_{x_1 + x_2 < x} f(x_1, x_2) dx_1 dx_2 = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{x - x_1} f(x_1, x_2) dx_2.$$

Пусть пока ξ_1 и ξ_2 - НСВ. Тогда $f(x_1, x_2) = f_1(x_1) f_2(x_2)$, где $f_1(x_1), f_2(x_2)$ - плотности распределений ξ_1 и ξ_2 соответственно.

$$F(x) = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{x - x_1} f(x_1, x_2) dx_2 = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{x - x_1} f_1(x_1) f_2(x_2) dx_2 =$$

$$= \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{x} f_1(x_1) f_2(z - x_1) dz = \int_{-\infty}^{x} dz \left\{ \int_{-\infty}^{\infty} f_1(x_1) f_2(z - x_1) dx_1 \right\}.$$

Отметим, что в общем случае (без независимости ξ_1 и ξ_2) получим

$$F(x) = \int_{-\infty}^{x} dz \int_{-\infty}^{\infty} f(x_1, z - x_1) dx_1.$$

Последние равенства доказывают, что если многомерное распределение слагаемых имеет плотность распределения, то их сумма также имеет плотность распределения. В случае двух независимых слагаемых эта плотность имеет вид:

$$f(x) = \int_{-\infty}^{\infty} f_1(z) f_2(x - z) dz = \int_{-\infty}^{\infty} f_1(x - z) f_2(z) dz.$$

Здесь первое равенство получено дифференцированием формулы для функции распределения суммы (в случае независимых слагаемых). Второе равенство очевидно из соображений симметрии.

В общем случае

$$f(x) = (F(x))' = \left\{ \int_{-\infty}^{x} dz \int_{-\infty}^{\infty} f(x_1, z - x_1) dx_1 \right\}'_x =$$

$$\left\{ \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} f(x_1, z - x_1) dx_1 \right) dz \right\}'_x = \int_{-\infty}^{\infty} f(x_1, x - x_1) dx_1 =$$

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

Последнее равенство получено переименованием x_1 в z.

Глава 5. Числовые характеристики случайных величин. Неравенства для случайных величин

5.1 Основные определения

Рассмотрим случайную величину

$$\xi : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \mathcal{P}_\xi)$$

с функцией распределения $F_\xi(x) = \mathbf{P}(\xi < x)$ и борелевскую (измеримую) функцию $g : \mathbb{R} \to \mathbb{R}$ (практически всегда мы будем рассматривать лишь непрерывные функции). Как отмечалось ранее, $\eta = g(\xi)$ также будет случайной величиной на исходном вероятностном пространстве.

В дальнейшем мы будем использовать *интеграл Стилтьеса* (см. **приложение 2.)**

Определение 1.

1. *Математическим ожиданием (средним значением) случайной величины ξ назовём число*

$$\mathbf{E}(\xi) = \mathbf{E}\xi = \int_{-\infty}^{\infty} x dF_\xi(x),$$

если интеграл в правой части существует.

2. *Математическим ожиданием случайной величины $\eta = g(\xi)$ назовём число*

$$\mathbf{E}\eta = \mathbf{E}(g(\xi)) = \int_{-\infty}^{\infty} g(x) dF_\xi(x) = \int_{-\infty}^{\infty} x dF_{g(\xi)}(x).$$

Данное определение сформулировано в общем виде скорее для "привыкания"к математическому (вероятностному) языку. Нам же чаще всего придется сталкиваться с дискретными или абсолютно непрерывными распределениями, где интеграл Стилтьеса превращается в числовой ряд или интеграл Римана соответственно. Тем не менее следует иметь в виду, что свойства математических ожиданий функций от случайных величин, даже если доказательства будут даваться только для дискретного и абсолютно непрерывного случаев, будут справедливы для произвольных распределений (и будут формулироваться в общем виде для произвольного распределения).

Замечание.

1. Пусть $F_\xi(x)$ - функция распределения дискретной случайной величины ξ, заданной рядом распределения

Значения	$\cdots$	x_{-k}	x_{-k+1}	$\cdots$	x_m	x_{m+1}	$\cdots$
Вероятности	$\cdots$	p_{-k}	p_{-k+1}	$\cdots$	p_m	p_{m+1}	$\cdots$

$$-\infty < \cdots < x_{-k} < x_{-k+1} < \cdots < x_m < x_{m+1} < \cdots < \infty,$$

$$\forall i \quad p_i > 0 \, , \, \sum_{i \in \mathbb{Z}} p_i = 1 \, , \, p_i = \mathbf{P}(\xi = x_i).$$

Тогда математическое ожидание случайной величины (интеграл Стилтьеса) превратится в сумму (см. **приложение 2**):

$$\mathbf{E}\xi = \int_{-\infty}^{\infty} x dF_\xi(x) = \sum_{k \in \mathbb{Z}} x_k \mathbf{P}(\xi = x_k) = \sum_{k \in \mathbb{Z}} x_k p_k,$$

$$\mathbf{E}(g(\xi)) = \sum_{k \in \mathbb{Z}} g(x_k) p_k.$$

2. Пусть $F_\xi(x)$ - абсолютно непрерывная функция распределения случайной величины ξ с плотностью $f_\xi(x)$. Тогда математическое ожидание случайной величины (интеграл Стилтьеса) превратится в интеграл Римана (см. **приложение 2**):

$$\mathbf{E}\xi = \int_{-\infty}^{\infty} x dF_\xi(x) = \int_{-\infty}^{\infty} x f_\xi(x) dx \, , \, \mathbf{E}(g(\xi)) = \int_{-\infty}^{\infty} g(x) f_\xi(x) dx.$$

3. Часто в литературе математическое ожидание случайной величины обозначают $\mathbf{M}\xi$.

Определение 2.*Введем следующие понятия :*

1. $\mathbf{E}(\xi^k)$ - *момент порядка k случайной величины ξ,*
2. $\mathbf{E}(|\xi|^k)$ - *абсолютный момент порядка k случайной величины ξ,*
3. $\mathbf{E}\left\{(\xi - \mathbf{E}\xi)^k\right\}$ - *центральный момент порядка k случайной величины ξ ,*
4. $\mathbf{E}\left\{|\xi - \mathbf{E}\xi|^k\right\}$ - *абсолютный центральный случайной величины ξ момент порядка k.*
5. $\mathbf{E}(\xi_1^{k_1}...\xi_n^{k_n})$ - *смешанный момент порядка $k = k_1 + ... + k_n$ случайного вектора $\bar{\xi} = (\xi_1, ..., \xi_n)$,*
6. $\mathbf{E}\left\{(\xi_1 - \mathbf{E}\xi_1)^{k_1}...(\xi_n - \mathbf{E}\xi_n)^{k_n}\right\}$ - *центральный смешанный момент порядка $k = k_1 + ... + k_n$ случайного вектора $\bar{\xi} = (\xi_1, ..., \xi_n)$.*

Замечание. В пунктах 5 и 6 определения 2 соотвествующие моменты нужно понимать следующим образом (поясним на примере п.5):

$$\mathbf{E}(\xi_1^{k_1}...\xi_n^{k_n}) = \mathbf{E}\zeta = \int_{-\infty}^{\infty} x dF_\zeta(x) = \int_{-\infty}^{\infty} x dF_{\xi_1^{k_1}...\xi_n^{k_n}}(x),$$

$$\zeta = \xi_1^{k_1}...\xi_n^{k_n} \ , \ F_\zeta(x) = F_{\xi_1^{k_1}...\xi_n^{k_n}}(x) = \mathbf{P}(\zeta < x).$$

На практике это часто приводит к трудностям, связанным с получением $F_\zeta(x)$ в явном виде. Поэтому удобнее пользоваться иным способом задания математического ожидания для функций координат случайных векторов, а именно:

$$\bar{\xi} = (\xi_1, ..., \xi_n) \ , \ F_{\bar{\xi}}(\bar{x}) = F_{\xi_1,...,\xi_n}(x_1, ..., x_n) \ , \ g = g(x_1, ..., x_n) : \mathbb{R}^n \to \mathbb{R},$$

$$\mathbf{E}(g(\xi_1, ..., \xi_n)) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} g(x_1, ..., x_n) dF_{\xi_1,...,\xi_n}(x_1, ..., x_n).$$

Мы не будем давать формального определения n-мерного интеграла Стилтьеса, потому что фактически нам понадобятся только дискретные и абсолютно непрерывные многомерные случайные величины. Для них мы конкретизируем эту формулу. Рассмотрим дискретный случай:

$$\bar{\xi} = (\xi_1, ..., \xi_n) \in \left\{(a_{i_1}^{(1)}, ..., a_{i_n}^{(n)})\right\}_{i_1,...,i_n \in \mathbb{N}} \subset \mathbb{R},$$

$$(a_{i_1}^{(1)}, ..., a_{i_n}^{(n)}) \neq (a_{j_1}^{(1)}, ..., a_{j_n}^{(n)}) \ , \ (i_1, ..., i_n) \neq (j_1, ..., j_n),$$

$$\mathbf{P}\left((\xi_1, ..., \xi_n) = (a_{i_1}^{(1)}, ..., a_{i_n}^{(n)})\right) = p_{i_1,...,i_n} \ , \ \sum_{i_1,...,i_n} p_{i_1,...,i_n} = 1,$$

$$\mathbf{E}(g(\xi_1, ..., \xi_n)) = \sum_{i_1,...,i_n} g(a_{i_1}^{(1)}, ..., a_{i_n}^{(n)}) p_{i_1,...,i_n}.$$

В абсолютно непрерывном случае: обозначим $f_{\xi_1,...,\xi_n}(x_1, ..., x_n)$ плотность распределения случайного ветора $\bar{\xi} = (\xi_1, ..., \xi_n)$. Тогда

$$\mathbf{E}(g(\xi_1, ..., \xi_n)) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} g(x_1, ..., x_n) f_{\xi_1,...,\xi_n}(x_1, ..., x_n) dx_1 ... dx_n,$$

(подразумевается *измеримость* $g(x_1, ..., x_n)$:

$$\forall B \in \mathcal{B}_{\mathbb{R}} \Rightarrow g^{-1}(B) \in \mathcal{B}^n.)$$

В частности:

$$\mathbf{E}\xi_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} x_1 f_{\xi_1,\xi_2,...,\xi_n}(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n =$$

$$= \int_{-\infty}^{\infty} x_1 dx_1 \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f_{\xi_1,\xi_2,...,\xi_n}(x_1, x_2, ..., x_n) dx_2 ... dx_n = \int_{-\infty}^{\infty} x_1 f_{\xi_1}(x_1) dx_1,$$

где

$$f_{\xi_1}(x_1) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f_{\xi_1,\xi_2,...,\xi_n}(x_1, x_2, ..., x_n) dx_2 ... dx_n$$

- плотность распределения ξ_1.

5.2 Простейшие свойства математического ожидания и дисперсии

Рассмотрим произвольные случайные величины ξ и η, заданные на одном вероятностном пространстве. Тогда

1. $$\forall a, b \in \mathbb{R} \Rightarrow \mathbf{E}a = a \ , \ \mathbf{E}(b\xi) = b\mathbf{E}\xi.$$

2. $$\mathbf{E}(\xi + \eta) = \mathbf{E}\xi + \mathbf{E}\eta,$$

   если существуют любые 2 из присутствующих математических ожиданий. Свойство очевидным образом обобщается на сумму конечного числа случайных величин:

   $$\mathbf{E}(\xi_1 + ... + \xi_n) = \mathbf{E}\xi_1 + ... + \mathbf{E}\xi_n.$$

3. $$\mathbf{P}(a \le \xi \le b) = 1 \Rightarrow a \le \mathbf{E}\xi \le b. \quad \text{Всегда } |\mathbf{E}\xi| \le \mathbf{E}|\xi|.$$

4.
$$\mathbf{P}(\xi \geq 0) = 1\ ,\ \mathbf{E}\xi = 0 \ \Rightarrow\ \mathbf{P}(\xi = 0) = 1.$$

5. Пусть случайные величины ξ и η независимы. Тогда
$$\mathbf{E}(\xi\eta) = \mathbf{E}\xi\mathbf{E}\eta.$$

   Обратное утверждение неверно. Свойство очевидным образом обобщается на набор конечного числа независимых случайных величин: если $\xi_1, ..., \xi_n$ - НСВ, то
$$\mathbf{E}(\xi_1...\xi_n) = \mathbf{E}\xi_1...\mathbf{E}\xi_n.$$

6.
$$\mathbf{P}(\xi \leq \eta) = 1 \ \Rightarrow\ \mathbf{E}\xi \leq \mathbf{E}\eta.$$

Фактически основные свойства математических ожиданий (1, 3 и 4) совпадают со свойствами интеграла Стилтьеса, но доказательства свойств 2 и 5 в общем случае достаточно трудоёмки. Поэтому, учитывая, что большинство практически важных распределений относится к дискретному или абсолютно непрерывному типам распределений, приведём доказательства пунктов 2 и 5 лишь для них; тем не менее свойства верны для произвольных распределений. В качестве иллюстрации использования интеграла Стилтьеса докажем сначала свойство 3.

Доказательство п.3. Пусть $F_\xi(x)$ - функция распределения ξ. Тогда
$$a \leq \xi \leq b \ \Rightarrow\ F_\xi(x) = 0, \text{ если } x \leq a\ ,\ F_\xi(x) = 1, \text{ если } x > b.$$

Используя свойства интеграла Стилтьеса (см. приложение 2),
$$a\int_{a-}^{b-} dF_\xi(x) = \int_{a-}^{b-} adF_\xi(x) \leq \int_{a-}^{b-} xdF_\xi(x) \leq \int_{a-}^{b-} bdF_\xi(x) = b\int_{a-}^{b-} dF_\xi(x).$$

По свойствам данной интегрирующей функции $F_\xi(x)$ (постоянство на $(-\infty, a]$ и (b, ∞)) получим, что для любой константы $C \in \mathbb{R}$
$$\int_{-\infty}^{a+} CdF_\xi(x) = \int_{-\infty}^{a+} xdF_\xi(x) = \int_{b+}^{\infty} xdF_\xi(x) = \int_{b+}^{\infty} CdF_\xi(x) = 0$$

Используя аддитивность интеграла Стилтьеса как функции множеств, окончательно получим:
$$a = \int_{-\infty}^{\infty} adF_\xi(x) \leq \int_{-\infty}^{\infty} xdF_\xi(x) = \mathbf{E}\xi \leq b = \int_{-\infty}^{\infty} bdF_\xi(x)$$

Вторая часть п.3 очевидным образом следует из определения интеграла Стилтьеса (см. приложение 2).

Доказательство п.2.

Дискретный случай. Пусть ξ - случайная величина, заданная рядом распределения

Значения	a_1	a_2	$\cdots$	a_n	$\cdots$
Вероятности	p_1	p_2	$\cdots$	p_n	$\cdots$,

и пусть η - случайная величина, заданная рядом распределения

Значения	b_1	b_2	$\cdots$	b_k	$\cdots$
Вероятности	q_1	q_2	$\cdots$	q_k	$\cdots$.

Возможные значения случайной величины $\xi+\eta$ (тоже дискретной) имеют вид $a_n + b_k$, $n, k \in \mathbb{N}$. Введем вероятности $\{g_{nk}\}_{n,k}$ и события $\{G_{nk}\}_{n,k}$:

$$G_{nk} = \{\xi = a_n, \eta = b_k\}, \bigcup_{n,k} G_{nk} = \Omega, G_{ni}G_{nj} = \emptyset, i \neq j, G_{lk}G_{mk} = \emptyset, l \neq m,$$

$$g_{nk} = \mathbf{P}(\xi = a_n, \eta = b_k) = \mathbf{P}(G_{nk}) , \sum_{n,k} g_{nk} = 1.$$

В данном случае математическое ожидание (интеграл Стилтьеса) представляет собой ряд, поэтому

$$\mathbf{E}(\xi + \eta) = \sum_{n,k} (a_n + b_k) g_{nk} = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} (a_n + b_k) g_{nk} =$$

$$= \sum_{n=1}^{\infty} a_n \left(\sum_{k=1}^{\infty} g_{nk} \right) + \sum_{k=1}^{\infty} b_k \left(\sum_{n=1}^{\infty} g_{nk} \right) .$$

Заметим:

$$\bigcup_{k=1}^{\infty} G_{nk} = \{\xi = a_n\} , \bigcup_{n=1}^{\infty} G_{nk} = \{\eta = b_k\}.$$

Учитывая указанные выше свойства событий G_{nk}, получим:

$$\sum_{n=1}^{\infty} a_n \left(\sum_{k=1}^{\infty} g_{nk} \right) = \sum_{n=1}^{\infty} a_n p_n = \mathbf{E}\xi,$$

$$\sum_{k=1}^{\infty} b_k \left(\sum_{n=1}^{\infty} g_{nk} \right) = \sum_{k=1}^{\infty} b_k q_k = \mathbf{E}\eta,$$

что завершает доказательство в дискретном случае.
Абсолютно непрерывный случай. Рассмотрим абсолютно непрерывно распределенный случайный вектор (ξ, η) с плотностью распределения $f(x, y)$ и случайную величину $\zeta = \xi + \eta$. Обозначим

$$F_\zeta(x) = \mathbf{P}(\zeta < x).$$

В данном случае математическое ожидание (интеграл Стилтьеса) представляет собой интеграл Римана, формула для плотности суммы нами выведена (см. **4.4**), поэтому

$$\mathbf{E}\zeta = \mathbf{E}(\xi + \eta) = \int_{-\infty}^{\infty} x dF_\zeta(x) = \int_{-\infty}^{\infty} x \left(\int_{-\infty}^{\infty} f(z, x - z) dz \right) dx =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(z, x - z) dz dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (z + y) f(z, y) dz dy =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z f(z, y) dz dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(z, y) dz dy =$$

$$= \int_{-\infty}^{\infty} z f_\xi(z) dz + \int_{-\infty}^{\infty} y f_\eta(y) dy = \mathbf{E}\xi + \mathbf{E}\eta,$$

где $f_\xi(z)$ и $f_\eta(y)$ - плотности, соотвественно, ξ и η:

$$f_\xi(z) = \int_{-\infty}^{\infty} f(z, y) dy \; , \; f_\eta(y) = \int_{-\infty}^{\infty} f(z, y) dz,$$

что завершает доказательство в абсолютно непрерывном случае.
Доказательство п.5.
Дискретный случай. Возьмём случайные величины ξ и η из соответствующей части доказательства п.2.

$$\mathbf{E}(\xi\eta) = \sum_{n,k} (a_n b_k p_n q_k) = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} (a_n b_k p_n q_k) =$$

$$= \left(\sum_{n=1}^{\infty} a_n p_n \right) \left(\sum_{k=1}^{\infty} b_k q_k \right) = \mathbf{E}\xi \mathbf{E}\eta,$$

что завершает доказательство в дискретном случае.
Абсолютно непрерывный случай. Рассмотрим абсолютно непрерывно распределенный случайный вектор (ξ, η) с плотностью распределения

$f(x, y)$. Можно было бы ввести случайную величину $\zeta = \xi\eta$, найти, как в примере с суммой $\xi + \eta$, её функцию распределения $F_\zeta(x)$ и плотность $f_\zeta(x)$, после чего посчитать

$$\mathbf{E}(\zeta) = \int_{-\infty}^{\infty} x f_\zeta(x) dx.$$

Этот способ оказывается довольно трудоёмким, поэтому мы поступим иначе. Рассмотрим функцию $g(x, y) = xy$, $(x, y) \in \mathbb{R}^2$ и заметим, что, в силу независимости ξ и η, $f(x, y) = f_\xi(x) f_\eta(y)$, где плотности $f_\xi(x)$ и $f_\eta(y)$ - плотности ξ и η соответственно. Тогда

$$\mathbf{E}(\xi\eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_\xi(x) f_\eta(y) dx dy = \int_{-\infty}^{\infty} x f_\xi(x) dx \int_{-\infty}^{\infty} y f_\eta(y) dy = \mathbf{E}\xi\mathbf{E}\eta,$$

что завершает доказательство в абсолютно непрерывном случае.
Следствие 1. Если случайные величины ξ и η независимы, а функции $g(x)$ и $h(x)$ - борелевские, то

$$\mathbf{E}(g(\xi)h(\eta)) = \mathbf{E}(g(\xi))\mathbf{E}(h(\eta)).$$

Доказательство. Следует из того, что $g(\xi)$ и $h(\eta)$, как уже показывалось, также будут независимы.

Осталось показать, что обратное утверждение неверно.
Пример. Пусть ξ и ζ независимы, $\mathbf{E}\xi = \mathbf{E}\zeta = 0$ и пусть $\eta = \xi\zeta$. Тогда ξ и η, очевидно, зависимы (за исключением некоторых тривиальных случаев, например $\xi = 0$ и т.п.). Однако

$$\mathbf{E}(\xi\eta) = \mathbf{E}(\xi^2\zeta) = \mathbf{E}(\xi^2)\mathbf{E}\zeta = 0 = \mathbf{E}\xi\mathbf{E}\eta.$$

Доказательство п.6. Рассмотрим случайную величину $\zeta = \eta - \xi$. Очевидно, что $\mathbf{P}(\zeta \geq 0) = 1$. Теперь утверждение следует из п.3 (a=0).
Определение 3. *Дисперсией случайной величины называется её центральный момент второго порядка.*

$$\mathbf{D}(\xi) = \mathbf{D}\xi = \mathbf{E}\left\{(\xi - \mathbf{E}\xi)^2\right\}.$$

Величина $\sigma = \sqrt{\mathbf{D}\xi}$ называется стандартным уклонением.

Отметим ещё одну формулу для нахождения дисперсии:

$$\mathbf{D}\xi = \mathbf{E}\left\{(\xi - \mathbf{E}\xi)^2\right\} = \mathbf{E}\left\{\xi^2 - 2\xi\mathbf{E}\xi + (\mathbf{E}\xi)^2\right\} =$$

$$= \mathbf{E}(\xi^2) - 2\mathbf{E}\xi\mathbf{E}\xi + (\mathbf{E}\xi)^2 = \mathbf{E}(\xi^2) - (\mathbf{E}\xi)^2.$$

Простейшие свойства дисперсии. Рассмотрим произвольные случайные величины ξ и η, заданные на одном вероятностном пространстве. Тогда

1.
$$\mathbf{D}\xi \geq 0\,;\ \mathbf{D}\xi = 0 \ \Leftrightarrow\ \mathbf{P}(\xi = c) = 1\,,\ c = const.$$

Неотрицательность дисперсии следует из свойств интеграла Стилтьеса. Далее:

$$\mathbf{P}(\xi = c) = 1 \ \Rightarrow\ \mathbf{E}(\xi^2) = (\mathbf{E}\xi)^2 = c^2 \ \Rightarrow\ \mathbf{D}\xi = 0.$$

Обратно

$$\mathbf{D}\xi = \mathbf{E}\left\{(\xi - \mathbf{E}\xi)^2\right\} = 0 \ \Rightarrow\ \mathbf{P}(\xi - \mathbf{E}\xi = 0) = 1 \ \Leftrightarrow\ \mathbf{P}(\xi = \mathbf{E}\xi) = 1,$$

(следует из неотрицательности $(\xi - \mathbf{E}\xi)^2$ и свойства 4 математического ожидания).

2.
$$\forall a, b \in \mathbb{R} \ \Rightarrow\ \mathbf{D}(a + b\xi) = b^2\mathbf{D}\xi.$$

Это свойство следует из определения.

3. Если случайные величины ξ и η независимы, то дисперсия аддитивна:

$$\mathbf{D}(\xi + \eta) = \mathbf{D}\xi + \mathbf{D}\eta.$$

Действительно, т.к. $\mathbf{E}(\xi + \eta) = \mathbf{E}\xi + \mathbf{E}\eta$ и $\mathbf{E}(\xi\eta) = \mathbf{E}\xi\mathbf{E}\eta$, имеем

$$\mathbf{D}(\xi + \eta) = \mathbf{E}\{(\xi + \eta)^2\} - (\mathbf{E}\xi + \mathbf{E}\eta)^2 = \mathbf{E}(\xi^2) + 2\mathbf{E}\xi\mathbf{E}\eta + \mathbf{E}(\eta^2) -$$

$$-(\mathbf{E}\xi)^2 - (\mathbf{E}\eta)^2 - 2\mathbf{E}\xi\mathbf{E}\eta = \mathbf{E}(\xi^2) - (\mathbf{E}\xi)^2 + \mathbf{E}(\eta^2) - (\mathbf{E}\eta)^2 = \mathbf{D}\xi + \mathbf{D}\eta.$$

Заметим, что здесь достаточно потребовать не независимость ξ и η, а лишь $\mathbf{E}(\xi\eta) = \mathbf{E}\xi\mathbf{E}\eta$. Свойство очевидным образом обобщается на сумму конечного числа независимых случайных величин:

$$\mathbf{D}(\xi_1 + ... + \xi_n) = \mathbf{D}\xi_1 + ... + \mathbf{D}\xi_n.$$

Замечание. Если представить себе распределение случайной величины ξ как единичную массу, "размазанную" по прямой, то математическое ожидание $\mathbf{E}\xi$ есть координата центра тяжести, дисперсия $\mathbf{D}\xi$ – момент инерции распределения этой массы на прямой.

5.3 Примеры

1. **Испытание Бернулли.** Случайная величина $\xi \in B_p$, поэтому

$$\mathbf{E}\xi = 0 \cdot q + 1 \cdot p = p.$$

Очевидно, что $\xi = \xi^2$, поэтому

$$\mathbf{E}(\xi^2) = p,\, (\mathbf{E}\xi)^2 = p^2 \Rightarrow \mathbf{D}\xi = \mathbf{E}(\xi^2) - (\mathbf{E}\xi)^2 = p - p^2 = p(1-p) = pq.$$

2. **Биномиальное распределение.** Воспользуемся представлением биномиально распределенной случайной величины $S_n \in B(n,p)$:

$$S_n = \xi_1 + ... + \xi_n\ ,\ \xi_i \in B_p\ ,\ \xi_1, ..., \xi_n \text{ - НОРСВ.}$$

Тогда по свойствам математического ожидания и дисперсии имеем:

$$\mathbf{E}S_n = \mathbf{E}(\xi_1 + ... + \xi_n) = \mathbf{E}\xi_1 + ... + \mathbf{E}\xi_n = np,$$

$$\mathbf{D}S_n = \mathbf{D}(\xi_1 + ... + \xi_n) = \mathbf{D}\xi_1 + ... + \mathbf{D}\xi_n = npq.$$

3. **Пуассоновское распределение.** Напомним, что

$$\xi \in \Pi_\lambda,\ \lambda > 0 \ \Leftrightarrow\ \xi \in \{0, 1, 2, ..., n, ...\} \text{ и } \mathbf{P}(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda}.$$

$$\mathbf{E}\xi = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=1}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} =$$

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

$$\mathbf{E}(\xi^2) = \sum_{k=0}^{\infty} k^2\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=1}^{\infty} k^2\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{k=1}^{\infty} k\frac{\lambda^{k-1}}{(k-1)!} =$$

$$= \lambda e^{-\lambda}\left(\sum_{k=1}^{\infty}(k-1)\frac{\lambda^{k-1}}{(k-1)!} + \sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\right) =$$

$$= \lambda e^{-\lambda}\left(\sum_{k=2}^{\infty}(k-1)\frac{\lambda^{k-1}}{(k-1)!} + e^{\lambda}\right) = \lambda e^{-\lambda}\left(\lambda\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!} + e^{\lambda}\right) =$$

$$= \lambda e^{-\lambda}(\lambda e^{\lambda} + e^{\lambda}) = \lambda^2 + \lambda.$$

$$\mathbf{D}\xi = \mathbf{E}(\xi^2) - (\mathbf{E}\xi)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

4. **Равномерное распределение.** Случайная величина $\xi \in U[a,b]$. Напомним, что это распределение является абсолютно непрерывным.

$$\mathbf{E}\xi = \int_{-\infty}^{\infty} x f_\xi(x)dx = \frac{1}{b-a}\int_a^b xdx = \left.\frac{x^2}{2(b-a)}\right|_a^b = \frac{a+b}{2}.$$

$$\mathbf{E}(\xi^2) = \int_{-\infty}^{\infty} x^2 f_\xi(x)dx = \frac{1}{b-a}\int_a^b x^2dx = \left.\frac{x^3}{3(b-a)}\right|_a^b = \frac{a^2+ab+b^2}{3}.$$

$$\mathbf{D}\xi = \mathbf{E}(\xi^2) - (\mathbf{E}\xi)^2 = \frac{(b-a)^2}{12}.$$

5. **Экспоненциальное распределение .** Наша случайная величина $\xi \in \Gamma_\alpha$, $\alpha > 0$. Это распределение также является абсолютно непрерывным.

$$\mathbf{E}\xi = \int_{-\infty}^{\infty} x f_\xi(x)dx = \alpha\int_0^{\infty} xe^{-\alpha x}dx =$$

$$= \alpha\left(-\frac{x}{\alpha}e^{-\alpha x}\Big|_0^{\infty} + \frac{1}{\alpha}\int_0^{\infty} e^{-\alpha x}dx\right) = \int_0^{\infty} e^{-\alpha x}dx = -\frac{1}{\alpha}e^{-\alpha x}\Big|_0^{\infty} = \frac{1}{\alpha}.$$

$$\mathbf{E}(\xi^2) = \int_{-\infty}^{\infty} x^2 f_\xi(x)dx = \alpha\int_0^{\infty} x^2e^{-\alpha x}dx =$$

$$= \alpha\left(-\frac{x^2}{\alpha}e^{-\alpha x}\Big|_0^{\infty} + \frac{2}{\alpha}\int_0^{\infty} xe^{-\alpha x}dx\right) = \frac{2}{\alpha}\alpha\int_0^{\infty} e^{-\alpha x}dx = \frac{2}{\alpha^2}.$$

$$\mathbf{D}\xi = \mathbf{E}(\xi^2) - (\mathbf{E}\xi)^2 = \frac{1}{\alpha^2}.$$

6. **Нормальное распределение .** Рассмотрим абсолютно непрерывную случайную величину $\xi \in N(\mu, \sigma^2)$, $\mu \in \mathbb{R}, \sigma^2 > 0$.

$$\mathbf{E}\xi = \int_{-\infty}^{\infty} x f_\xi(x)dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} xe^{-\frac{(x-\mu)^2}{2\sigma^2}}dx =$$

$$= \frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{\infty} \frac{x-\mu}{\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx + \int_{-\infty}^{\infty} \frac{\mu}{\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx\right).$$

Сделав в каждом интеграле замену $\frac{x-\mu}{\sigma} = t$, получим

$$\mathbf{E}\xi = \frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty} te^{-\frac{t^2}{2}}dt + \mu\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt = \mu.$$

Здесь первый интеграл равен нулю как сходящийся интеграл от нечетной функции по всей прямой. Второй интеграл равен единице по свойству плотности (под интегралом стоит плотность $N(0,1)$).

Для вычисления дисперсии здесь удобнее использовать формулу, данную в определении:

$$\mathbf{D}\xi = \mathbf{E}\left\{(\xi - \mathbf{E}\xi)^2\right\} = \int_{-\infty}^{\infty} (x - \mathbf{E}\xi)^2 f_\xi(x)dx =$$

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

$$= \sigma^2 \left(-\frac{te^{-\frac{t^2}{2}}}{\sqrt{2\pi}} \Bigg|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \right) = \sigma^2.$$

Здесь после замены $\frac{x-\mu}{\sigma} = t$ мы использовали формулу интегрирования по частям:

$$u = t\ ,\ dv = te^{-\frac{t^2}{2}}dt\ \Rightarrow\ du = dt\ ,\ v = -e^{-\frac{t^2}{2}};$$

в последнем равенстве двойная подстановка обращается в ноль, а интеграл равен единице.

Мы получили, что параметры нормального распределения совпадают с его математическим ожиданием и дисперсией.

7. **Распределение Коши.** Рассмотрим $\xi \in K(0,1)$.

$$\mathbf{E}\xi = \int_{-\infty}^{\infty} x f_\xi(x)dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x}{1+x^2} dx =$$

$$= \frac{1}{\pi} \left\{ \int_{-\infty}^{0} \frac{x}{1+x^2} dx + \int_{0}^{\infty} \frac{x}{1+x^2} dx \right\}.$$

Оба интеграла в скобках расходятся, поэтому расходится интеграл по всей прямой (на самом деле для расходимости суммы достаточно расходимости одного из интегралов). Действительно, сделав замену $1 + x^2 = t$, получим

$$\int_0^{\infty} \frac{x}{1+x^2} dx = \frac{1}{2} \int_1^{\infty} \frac{dt}{t}$$

расходящийся интеграл. Т.о. у $K(0,1)$ (очевидно, что и у любого $K(\mu, \sigma^2)$) математическое ожидание не существует. Позже будет доказано, что не существуют и моменты любого высшего порядка, в частности дисперсия.

5.4 Коэффициент корреляции

Коэффициент корреляции представляет собой количественную меру зависимости двух случайных величин друг от друга. В его определении и свойствах будем предполагать существование у случайных величин конечного математического ожидания и конечной ненулевой дисперсии.

Определение 4. *Случайную величину ζ называют нормированной случайной величиной, если* $\mathbf{E}\zeta = 0$, $\mathbf{D}\zeta = \mathbf{E}(\zeta^2) = 1$.

Пусть ξ - случайная величина (с конечным математическим ожиданием и конечной ненулевой дисперсией). Построим по ней новую случайную величину:

$$\widetilde{\xi} = \frac{\xi - \mathbf{E}\xi}{\sqrt{\mathbf{D}\xi}} \text{ . Легко проверить, что: } \quad \mathbf{E}\widetilde{\xi} = 0 \text{ , } \mathbf{D}\widetilde{\xi} = \mathbf{E}(\widetilde{\xi}^2) = 1 :$$

$$\mathbf{E}\widetilde{\xi} = \frac{1}{\sqrt{\mathbf{D}\xi}}\mathbf{E}(\xi - \mathbf{E}\xi) = \frac{1}{\sqrt{\mathbf{D}\xi}}(\mathbf{E}\xi - \mathbf{E}\,(\mathbf{E}\xi)) = \frac{1}{\sqrt{\mathbf{D}\xi}}(\mathbf{E}\xi - \mathbf{E}\xi) = 0,$$

$$\mathbf{D}\widetilde{\xi} = \frac{1}{\mathbf{D}\xi}\mathbf{D}(\xi - \mathbf{E}\xi) = \frac{1}{\mathbf{D}\xi}\mathbf{D}\xi = 1.$$

Т.о. по любой случайной величине ξ (при наличии определённых условий на $\mathbf{E}\xi$ и $\mathbf{D}\xi$) всегда можно построить нормированную $\widetilde{\xi}$.

Определение 5.

1. *Коэффициентом ковариации между случайными величинами ξ и η называют число*

$$cov(\xi, \eta) = \mathbf{E}\,\{(\xi - \mathbf{E}\xi)(\eta - \mathbf{E}\eta)\}$$

2. *Коэффициентом корреляции между случайными величинами ξ и η называют число*

$$\rho(\xi, \eta) = \frac{cov(\xi, \eta)}{\sqrt{\mathbf{D}\xi\mathbf{D}\eta}} = \mathbf{E}\left\{\frac{(\xi - \mathbf{E}\xi)(\eta - \mathbf{E}\eta)}{\sqrt{\mathbf{D}\xi\mathbf{D}\eta}}\right\} = \mathbf{E}(\widetilde{\xi}\widetilde{\eta}).$$

Замечание. Иногда удобнее вычислять коэффициент корреляции по другой формуле:

$$\rho(\xi, \eta) = \mathbf{E}\left\{\frac{(\xi - \mathbf{E}\xi)(\eta - \mathbf{E}\eta)}{\sqrt{\mathbf{D}\xi\mathbf{D}\eta}}\right\} = \frac{\mathbf{E}\,\{\xi\eta - \xi\mathbf{E}\eta - \eta\mathbf{E}\xi + \mathbf{E}\xi\mathbf{E}\eta\}}{\sqrt{\mathbf{D}\xi\mathbf{D}\eta}} =$$

$$= \frac{\mathbf{E}(\xi\eta) - \mathbf{E}\xi\mathbf{E}\eta}{\sqrt{\mathbf{D}\xi\mathbf{D}\eta}}.$$

Простейшие свойства коэффициента корреляции.

1. $$|\rho(\xi,\eta)| \le 1.$$

   Действительно, из неотрицательности дисперсии получим:

   $$0 \le \mathbf{D}(\widetilde{\xi} \pm \widetilde{\eta}) = \mathbf{E}\left\{(\widetilde{\xi} \pm \widetilde{\eta})^2\right\} = 2 \pm 2\rho(\xi,\eta) \ \Rightarrow\ |\rho(\xi,\eta)| \le 1.$$

2. Если ξ и η независимы, то $\rho(\xi,\eta) = 0$.
   В самом деле, из независимости ξ и η следует независимость их нормировок $\widetilde{\xi}$ и $\widetilde{\eta}$. По свойствам математического ожидания (п.5) имеем:
   $$\rho(\xi,\eta) = \mathbf{E}(\widetilde{\xi}\widetilde{\eta}) = \mathbf{E}\widetilde{\xi}\mathbf{E}\widetilde{\eta} = 0.$$

3. $$|\rho(\xi,\eta)| = 1 \ \Leftrightarrow\ \exists a \in \mathbb{R}\, \exists b \ne 0\ :\ \mathbf{P}(\eta = a + b\xi) = 1.$$

   В самом деле, пусть

   $$\mathbf{P}(\eta = a + b\xi) = 1 \text{ и обозначим } \mathbf{E}\xi = \alpha\ ,\ \sqrt{\mathbf{D}\xi} = \beta.$$

   Получаем

   $$\rho(\xi,\eta) = \mathbf{E}\left\{\frac{\xi - \alpha}{\beta} \cdot \frac{a + b\xi - a - b\alpha}{|b|\beta}\right\} = sign b = \pm 1.$$

   Обратно, пусть, например, $\rho(\xi,\eta) = 1$. Используя свойства дисперсии, получим:

   $$\mathbf{D}(\widetilde{\xi} - \widetilde{\eta}) = 2 - 2\rho(\xi,\eta) = 0 \ \Leftrightarrow\ \mathbf{P}(\widetilde{\xi} - \widetilde{\eta} = c) = 1 \ \Leftrightarrow$$

   $$\Leftrightarrow\ \mathbf{P}\left\{\frac{\xi - \mathbf{E}\xi}{\sqrt{\mathbf{D}\xi}} = \frac{\eta - \mathbf{E}\eta}{\sqrt{\mathbf{D}\eta}}\right\} = 1 \ \Leftrightarrow\ \exists a, b :\ \mathbf{P}(\eta = a + b\xi) = 1.$$

   Если же $\rho(\xi,\eta) = -1$, то, взяв $\mathbf{D}(\widetilde{\xi} - \widetilde{\eta})$, с помощью аналогичных выкладок получим
   $$\mathbf{P}(\widetilde{\xi} - \widetilde{\eta} = c) = 1.$$

Определение 6.*Случайные величины ξ и η называются:*

1. *отрицательно коррелированными, если $\rho(\xi,\eta) < 0$,*

2. *некоррелированными, если* $\rho(\xi, \eta) = 0$,

3. *положительно коррелированными, если* $\rho(\xi, \eta) > 0$.

Заметим, что некоррелированные случайные величины не обязаны быть независимыми. Чтобы убедиться в этом можно взять случайные величины из примера к доказательству п.5 свойств математического ожидания.

5.5 Неравенства для случайных величин

Рассмотрим произвольные случайные величины ξ и η и докажем ряд важных неравенств.

1. **Неравенство Гёльдера**. Пусть $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$. Тогда

$$\mathbf{E}\{|\xi\eta|\} \leq \{\mathbf{E}(|\xi|^p)\}^{\frac{1}{p}} \{\mathbf{E}(|\eta|^q)\}^{\frac{1}{q}} .$$

Доказательство. Рассмотрим функцию $y = x^p, x > 0$. Она выпукла вниз, $x(1) = 1$, $x'(1) = p$. Тогда, очевидно,

$$\forall x > 0 : \ p(x-1) \leq x^p - 1.$$

Пусть

$$a > 0 \, , \, b > 0 \, , \, x = \left(\frac{a}{b}\right)^{\frac{1}{p}} > 0.$$

Тогда:

$$p\left\{\left(\frac{a}{b}\right)^{\frac{1}{p}} - 1\right\} \leq \frac{a}{b} - 1 \ \Rightarrow \ \frac{a^{\frac{1}{p}} - b^{\frac{1}{p}}}{b^{\frac{1}{p}}} \leq \frac{a-b}{pb} \ \Rightarrow$$

$$\Rightarrow \ a^{\frac{1}{p}} b^{1-\frac{1}{p}} - b^{\frac{1}{p}} b^{1-\frac{1}{p}} \leq \frac{a}{p} - \frac{b}{p} \ \Rightarrow \ a^{\frac{1}{p}} b^{\frac{1}{q}} \leq \frac{a}{p} + \frac{b}{q}.$$

Подставив в полученное неравенство

$$a = \frac{|\xi|^p}{\mathbf{E}(|\xi|^p)} \, , \, b = \frac{|\eta|^q}{\mathbf{E}(|\eta|^q)},$$

получим

$$\frac{|\xi\eta|}{\{\mathbf{E}(|\xi|^p)\}^{\frac{1}{p}} \{\mathbf{E}(|\eta|^q)\}^{\frac{1}{q}}} \leq \frac{q|\xi|^p \mathbf{E}(|\eta|^q) + p|\eta|^q \mathbf{E}(|\xi|^p)}{pq\mathbf{E}(|\xi|^p)\,\mathbf{E}(|\eta|^q)} \ \Rightarrow$$

$$\Rightarrow\ |\xi\eta| \le \frac{q|\xi|^p \mathbf{E}\left(|\eta|^q\right) + p|\eta|^q \mathbf{E}\left(|\xi|^p\right)}{pq\left\{\mathbf{E}\left(|\xi|^p\right)\right\}^{\frac{1}{q}}\left\{\mathbf{E}\left(|\eta|^q\right)\right\}^{\frac{1}{p}}}.$$

Возьмём математическое ожидание от обеих частей: знак неравенства (по свойствам) не изменится, $\mathbf{E}\{\mathbf{E}(h(\zeta))\} = \mathbf{E}(h(\zeta))$ т.к. $\mathbf{E}(h(\zeta)) = const$. В результате:

$$\mathbf{E}\left\{|\xi\eta|\right\} \le \frac{(p+q)\mathbf{E}\left(|\xi|^p\right)\mathbf{E}\left(|\eta|^q\right)}{pq\left\{\mathbf{E}\left(|\xi|^p\right)\right\}^{\frac{1}{q}}\left\{\mathbf{E}\left(|\eta|^q\right)\right\}^{\frac{1}{p}}}.$$

$$\frac{p+q}{pq} = \frac{1}{p} + \frac{1}{q} = 1\ \Rightarrow\ \mathbf{E}\left\{|\xi\eta|\right\} \le \left\{\mathbf{E}\left(|\xi|^p\right)\right\}^{\frac{1}{p}}\left\{\mathbf{E}\left(|\eta|^q\right)\right\}^{\frac{1}{q}}.$$

Замечание. При $p = q = 2$ получим хорошо известное **неравенство Коши-Буняковского**:

$$\mathbf{E}\left\{|\xi\eta|\right\} \le \sqrt{\mathbf{E}\left\{\xi^2\right\}\mathbf{E}\left\{\eta^2\right\}}.$$

2. **Неравенство Минковского**. Пусть $p > 1$. Тогда

$$\left\{\mathbf{E}\left(|\xi+\eta|^p\right)\right\}^{\frac{1}{p}} \le \left\{\mathbf{E}\left(|\xi|^p\right)\right\}^{\frac{1}{p}} + \left\{\mathbf{E}\left(|\eta|^p\right)\right\}^{\frac{1}{p}}.$$

Доказательство. Заметим прежде всего:

$$\frac{1}{p} + \frac{1}{q} = 1\ \Rightarrow\ \frac{1}{q} = \frac{p-1}{p}\,,\ q = \frac{p}{p-1}.$$

Из цепочки неравенств

$$|\xi+\eta| \le |\xi|+|\eta| \Rightarrow |\xi+\eta|^p = |\xi+\eta||\xi+\eta|^{p-1} \le |\xi||\xi+\eta|^{p-1}+|\eta||\xi+\eta|^{p-1}$$

получаем, взяв математические ожидания обеих частей:

$$\mathbf{E}\left(|\xi+\eta|^p\right) \le \mathbf{E}\left(|\xi||\xi+\eta|^{p-1}\right) + \mathbf{E}\left(|\eta||\xi+\eta|^{p-1}\right).$$

Применяя к слагаемым правой части неравенство Гёльдера, получим:

$$\mathbf{E}\left(|\xi+\eta|^p\right) \le \left\{\mathbf{E}\left(|\xi|^p\right)\right\}^{\frac{1}{p}}\left\{\mathbf{E}\left(|\xi+\eta|^{(p-1)\frac{p}{p-1}}\right)\right\}^{\frac{p-1}{p}} +$$

$$+\left\{\mathbf{E}\left(|\eta|^p\right)\right\}^{\frac{1}{p}}\left\{\mathbf{E}\left(|\xi+\eta|^{(p-1)\frac{p}{p-1}}\right)\right\}^{\frac{p-1}{p}} =$$

$$= \left(\left\{\mathbf{E}\left(|\xi|^p\right)\right\}^{\frac{1}{p}} + \left\{\mathbf{E}\left(|\eta|^p\right)\right\}^{\frac{1}{p}}\right)\left\{\mathbf{E}\left(|\xi+\eta|^p\right)\right\}^{1-\frac{1}{p}},$$

откуда следует неравенство Минковского.

3. **Неравенство Йенсена**. Пусть $\mathbf{E}\xi$ конечно и функция $g(x)$ выпукла вниз. Тогда

$$g(\mathbf{E}\xi) \leq \mathbf{E}(g(\xi)).$$

Доказательство. Из геометрических соображений очевидно: если $g(x)$ выпукла вниз, то для любого y существует число $\widetilde{g}(y)$ такое, что

$$g(x) \geq g(y) + (x-y)\widetilde{g}(y).$$

Положим $x = \xi$, $y = \mathbf{E}\xi$ и возьмём математическое ожидание обеих сторон:

$$\mathbf{E}(g(\xi)) \geq \mathbf{E}\left\{g(\mathbf{E}\xi)\right\} + \mathbf{E}\left\{(\xi - \mathbf{E}\xi)\widetilde{g}(\mathbf{E}\xi\right\} = g(\mathbf{E}\xi).$$

Следствие 2. *Пусть* $0 < v < u$, ζ *- случайная величина. . Тогда*

$$\left\{\mathbf{E}\left(|\zeta|^v\right)\right\}^{\frac{1}{v}} \leq \left\{\mathbf{E}\left(|\zeta|^u\right)\right\}^{\frac{1}{u}}.$$

Доказательство. Сразу отметим, что неравенство показывает: из существования абсолютного момента большего порядка следует существование абсолютного момента меньшего порядка; если не существует момент меньшего порядка, то не существует и абсолютный момент большего порядка.

Пусть $\xi = |\zeta|^v$, $g(x) = |x|^{u/v}$ - выпукла вниз, т.к. $u/v > 1$. Используя неравенство Йенсена, получим:

$$g(\mathbf{E}\xi) = \left\{\mathbf{E}\left(|\zeta|^v\right)\right\}^{\frac{u}{v}} \leq \mathbf{E}(g(\xi)) = \mathbf{E}\left(|\zeta|^u\right),$$

откуда следует требуемое неравенство.

4. **Неравенство Чебышева**.

$$\mathbf{P}(\xi \geq 0) = 1 \;\Rightarrow\; \forall \varepsilon > 0 : \; \mathbf{P}(\xi \geq \varepsilon) \leq \frac{\mathbf{E}\xi}{\varepsilon}.$$

Доказательство. Пусть $F_\xi(x)$ - функция распределения ξ. Заметим, что $F_\xi(x) = 0$, $x \leq 0$.

$$\mathbf{E}\xi = \int_{-\infty}^{\infty} x dF_\xi(x) = \int_0^{\infty} x dF_\xi(x) = \int_0^{\varepsilon} x dF_\xi(x) + \int_{\varepsilon}^{\infty} x dF_\xi(x) \geq$$

$$\geq \int_{\varepsilon}^{\infty} x dF_\xi(x) \geq \varepsilon \int_{\varepsilon}^{\infty} dF_\xi(x) = \varepsilon \mathbf{P}(\xi \geq \varepsilon),$$

откуда следует неравенство Чебышева. Отметим что в условии не требуется существование $\mathbf{E}\xi$. Дело в том, что в силу неотрицательности ξ её математическое ожидание не существует только если равно $+\infty$, что делает неравенство тривиальным.

Следствие 3.

1. *Пусть* $g(x)$ *- возрастающая борелевская функция,* $g(x) \geq 0$*. Тогда*

$$\mathbf{P}(\xi \geq 0) = 1 \;\Rightarrow\; \forall \varepsilon > 0: \; \mathbf{P}(\xi \geq \varepsilon) \leq \frac{\mathbf{E}(g(\xi))}{g(\varepsilon)}.$$

2. *Для произвольной случайной величины* ξ*, имеющей математическое ожидание, имеет место неравенство*

$$\mathbf{P}(|\xi - \mathbf{E}\xi| \geq \varepsilon) \leq \frac{\mathbf{D}\xi}{\varepsilon^2}.$$

Доказательство.

1. Из свойств $g(x)$ следует, что

$$\{g(\xi) \geq g(\varepsilon)\} = \{\xi \geq \varepsilon\},$$

после чего нужно применить неравенство Чебышева к случайной величине $\eta = g(\xi)$:

$$\mathbf{P}(\eta \geq g(\varepsilon)) = \mathbf{P}(g(\xi) \geq g(\varepsilon)) = \mathbf{P}(\xi \geq \varepsilon) \leq \frac{\mathbf{E}\eta}{g(\varepsilon)} = \frac{\mathbf{E}(g(\xi))}{g(\varepsilon)}.$$

2. Требуемое неравенство получается применением пункта 1 данного следствия к случайной величине $\eta = |\xi - \mathbf{E}\xi| \geq 0$ и функции

$$g(x) = x^2\,,\; x \geq 0.$$

Приложение 1. Множество меры 0

Введём вспомогательные обозначения: под символом Δ будем понимать произвольный (ограниченный) интервал, $|\Delta|$ - его длина:

$$\Delta = (a,b)\,,\ -\infty < a < b < \infty\,, |\Delta| = b - a.$$

Определение 1. *Рассмотрим множество E на прямой. Будем говорить, что множество E имеет меру* 0 *($\mu(E) = 0$), если*

$$\forall \varepsilon > 0\, \exists \{\Delta_\alpha^{(\varepsilon)}\}_\alpha : \Delta_\alpha^{(\varepsilon)} \subset \mathbb{R}\,,\ \sum_\alpha |\Delta_\alpha^{(\varepsilon)}| < \varepsilon \Rightarrow \bigcup_\alpha \Delta_\alpha^{(\varepsilon)} \supseteq E.$$

Приведем простейшие примеры множеств меры 0 на прямой.

1. **Конечные множества**. Множество E имеет вид:

$$E = \{x_1\} \bigcup \{x_2\} \bigcup \cdots \bigcup \{x_n\}\,,\ x_i \in \mathbb{R}\,,\ x_i \neq x_j\,,\ i \neq j.$$

   Фиксируем произвольное $\varepsilon > 0$. Возьмем в качестве интервала, "покрывающего" x_i,

$$\Delta_i^{(\varepsilon)} = (x_i - \frac{\varepsilon}{4n}, x_i + \frac{\varepsilon}{4n})\,,\ |\Delta_i^{(\varepsilon)}| = \frac{\varepsilon}{2n}.$$

   Тогда по построению

$$E \subset \bigcup_{i=1}^{n} \Delta_i^{(\varepsilon)}\,,\ \sum_{i=1}^{n} |\Delta_i^{(\varepsilon)}| = \frac{\varepsilon}{2} < \varepsilon\ \Rightarrow\ \mu(E) = 0.$$

2. **Счетные множества**. Множество E имеет вид:

$$E = \bigcup_{n \in \mathbb{N}} \{x_n\}\,,\ x_n \in \mathbb{R}\,,\ x_i \neq x_j\,,\ i \neq j.$$

   Фиксируем произвольное $\varepsilon > 0$. Возьмем в качестве интервала, "покрывающего" x_n,

$$\Delta_n^{(\varepsilon)} = (x_n - \frac{\varepsilon}{2^{n+2}}, x_n + \frac{\varepsilon}{2^{n+2}})\,,\ |\Delta_n^{(\varepsilon)}| = \frac{\varepsilon}{2^{n+1}}.$$

   Тогда по построению

$$E \subset \bigcup_{n \in \mathbb{N}} \Delta_n^{(\varepsilon)}\,,\ \sum_{n \in \mathbb{N}} |\Delta_n^{(\varepsilon)}| = \frac{\varepsilon}{2} < \varepsilon\ \Rightarrow\ \mu(E) = 0.$$

3. **Канторово множество**. Возьмем отрезок $[0,1]$ и разделим его на 3 равные части $[0,1/3]$, $[1/3,2/3]$, $[2/3,1]$, после чего "выкинем" из отрезка *внутренность средней части*, т.е. *интервал* $(1/3,2/3)$. С каждым из оставшихся отрезков $[0,1/3]$ и $[2/3,1]$ проделаем ту же процедуру: каждый из них вновь делится на 3 равные части и из каждого из них "выкидывается" *внутренность средней части*, в данном случае *интервалы* $(1/9,2/9)$ и $(7/9,8/9)$; каждый из оставшихся отрезков (теперь их уже 4) вновь делится на 3 равные части и из каждого из них опять "выкидывается" *внутренность средней части* и т.д., процедура продолжается до бесконечности. То, что остаётся в "пределе", называется *Канторовым множеством.* Нетрудно показать, что оно имеет мощность континуума. Тем не менее оно имеет меру 0, т.к. суммарная длина всех отброшенных интервалов равна

$$\frac{1}{3}+\frac{2}{9}+\frac{4}{27}+\cdots=\frac{1}{3}\sum_{k=0}^{\infty}\left(\frac{2}{3}\right)^k=\frac{1}{3}\frac{1}{1-\frac{2}{3}}=1.$$

Определение 2. *Будем говорить что какое-либо свойство выполняется в почти всех точках множества $A\subseteq\mathbb{R}$ (п.в. на A), если существует такое подмножество $E\subset A$, что во всех точках $A\backslash E$ свойство выполняется и $\mu(E)=0$.*

Иными словами, свойство может не выполняться только на подмножестве, имеющем меру ноль. Например, "функция $y=|x|$ дифференцируема п.в. на $\mathbb{R}$", "функция $y=tgx$ непрерывна п.в. на $\mathbb{R}$", "$|\cos x|<1$ п.в. на $\mathbb{R}$" и т.д..

Замечание. Аналогично можно ввести понятие множества нулевой меры в $\mathbb{R}^n$, $n>1$. В определении вместо интервала Δ обычно берут открытые шары (или параллелепипеды), а роль длины $|\Delta|$ играют площадь в $\mathbb{R}^2$, объём в $\mathbb{R}^3$ и т.д..

Приложение 2. Интеграл Стилтьеса

Мы приведём здесь определение и основные свойства интеграла Стилтьеса, не останавливаясь при этом на доказательствах (акцентируя внимание на нужных нам фактах).

Зададим на конечном интервале $(a,b)\subset\mathbb{R}$ функцию $g(x)$ и неубывающую функцию $F(x)$. Для наших нужд будет достаточно (так в дальнейшем и будем считать), чтобы $F(x)$ была произвольной функцией распределения на $\mathbb{R}$ (с суженной на (a,b) областью определения), хотя определение интеграла Стилтьеса шире. Если интервал (a,b) ограничен, то разобьём его произвольным образом выбранным конечным множеством точек $a = x_0 < x_1 < ... < x_n = b$ и образуем сумму

$$\sum_{i=1}^{n} g(\widetilde{x_i})(F(x_i) - F(x_{i-1}))\ ,\ \widetilde{x_i} \in (x_{i-1}, x_i),$$

причём $\widetilde{x_i}$ выбирается в своём интервале также произвольным образом.Теперь, увеличивая число точек разбиения $\{x_i\}$, одновременно устремим наибольшую из длин промежутков к нулю $max_i|x_i - x_{i-1}| \to 0$. Если при этом приведенная выше сумма стремится к определённому пределу (не зависящему от выбора $\{x_i\}$ и $\{\widetilde{x_i}\}$)

$$I = \lim_{n\to\infty} \sum_{i=1}^{n} g(\widetilde{x_i})(F(x_i) - F(x_{i-1}))\ ,\ \widetilde{x_i} \in (x_i, x_{i-1}),$$

то этот предел называется интегралом Стилтьеса от функции $g(x)$ с интегрирующей функцией $F(x)$ и обозначается

$$I = \int_a^b g(x)dF(x).$$

Несобственный интеграл Стилтьеса (например, по всей прямой), определяется обычным способом: рассматривается интеграл по произвольному конечному интервалу (a,b); величины a и b произвольным образом заставляют стремиться к $-\infty$ и ∞ соответственно; если при этом существует предел

$$\lim_{a\to-\infty, b\to\infty} \int_a^b g(x)dF(x),$$

то этот предел называется интегралом Стилтьеса от функции $g(x)$ с интегрирующей функцией $F(x)$ (по функции $F(x)$) по всей прямой и обозначается

$$I = \int_{-\infty}^{\infty} g(x)dF(x).$$

Аналогично определяется интеграл на луче.Можно доказать, что заведомо будут существовать интегралы Стилтьеса (как по конечным, так и по бесконечным промежуткам) для непрерывных и ограниченных функций $g(x)$. На самом деле можно существенно ослабить требования к $g(x)$, при необходимости мы будем это каждый раз обосновывать.

Заметим, что при установке пределов интегрирования необходимо указывать, включается в промежуток интегрирования или нет тот или иной его конец. Пусть концы a и b конечны, и символ a_- обозначает, что a включено в промежуток интегрирования, а символ a_+, соответственно, что a исключено из него.

$$\int_{a_-}^{b} g(x)dF(x) = \lim_{n\to\infty} \sum_{i=1}^{n} g(\widetilde{x}_i)(F(x_i) - F(x_{i-1})) =$$

$$\lim_{n\to\infty} \sum_{i=2}^{n} g(\widetilde{x}_i)(F(x_i) - F(x_{i-1})) + \lim_{x_1\to x_0=a} g(\widetilde{x}_1)(F(x_1) - F(x_0)) =$$

$$= \int_{a_+}^{b} g(x)dF(x) + g(a)(F(a+0) - F(a)).$$

Т.е., если $g(a) \neq 0$, и функция $F(x)$ в точке a разрывна (испытывает скачок), то

$$\int_{a_-}^{b} g(x)dF(x) - \int_{a_+}^{b} g(x)dF(x) = g(a)(F(a+0) - F(a)) \neq 0.$$

В частности, интеграл Стилтьеса, распространённый на одноточечные множества (например как предел по стягивающимся в точку промежуткам), может оказаться ненулевым. Условимся в дальнейшем правый конец промежутка исключать, а левый включать в промежуток интегрирования. Для нас это вполне естественно, т.к. в качестве интегрирующей функции $F(x)$ по уговору берётся функция распределения, т.е. непрерывная слева по определению: $F(b) = F(b-0)$. Пусть $F(x)$ - функция распределения случайной величины ξ. Тогда, учитывая всё вышесказанное,

$$\int_{a}^{b} dF(x) = \lim_{n\to\infty} \sum_{i=1}^{n} (F(x_i) - F(x_{i-1})) = \lim_{n\to\infty} (F(x_n) - F(x_0)) =$$

$$= F(b) - F(a) = \mathbf{P}(a \leq \xi < b),$$

$$\int_{-\infty}^{b} dF(x) = \mathbf{P}(\xi < b) \text{ и } \int_{-\infty}^{\infty} dF(x) = \mathbf{P}(\xi \in (-\infty, \infty)) = 1.$$

Пусть $F(x)$ абсолютно непрерывна, т.е. имеет производную $f(x) = F'(x)$ и восстанавливается интегралом от неё. Тогда из формулы конечных приращений

$$F(x_i) - F(x_{i-1}) = f(\widetilde{x}_i)(x_i - x_{i-1}) \,,\; x_{i-1} < \widetilde{x}_i < x_i$$

следует равенство:

$$\int_{a}^{b} g(x)dF(x) = \lim_{n\to\infty} \sum_{i=1}^{n} g(\widetilde{x}_i)(F(x_i) - F(x_{i-1})) =$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} g(\widetilde{x}_i) f(\widetilde{x}_i)(x_i - x_{i-1}) = \int_a^b g(x) f(x) dx.$$

Т.е. в этом случае интеграл Стилтьеса сводится к интегралу Римана.

Если $F(x)$ имеет скачок в точке $x = c$, то, выбрав разбиение промежутка интегрирования точками $\{x_i\}$ так, что при некоторых значениях индекса $x_k < c < x_{k+1}$, получим:

$$\int_a^b g(x)dF(x) = \lim_{n\to\infty} \sum_{i=1}^{k} g(\widetilde{x}_i)(F(x_i) - F(x_{i-1})) + g(c)(F(x_{k+1}) - F(x_k)) +$$

$$+ \lim_{n\to\infty} \sum_{i=k+2}^{n} g(\widetilde{x}_i)(F(x_i) - F(x_{i-1})) = \int_a^c g(x)dF(x) + \int_{c+}^b g(x)dF(x) +$$

$$+ g(c)(F(c+0) - F(c_0)).$$

Если $F(x)$ такова, что меняется только в точках $\{c_n\}_{n\in\mathbb{N}}$ (функция распределения дискретной случайной величины со значениями в этих точках), то:

$$\int_a^b g(x)dF(x) = \sum_{n\in\mathbb{N}} g(c_n)(F(c_n+0) - F(c_n-0)) = \sum_{n\in\mathbb{N}} g(c_n)(F(c_n+0) - F(c_n))$$

(в последнем равенстве мы использовали непрерывность слева функции $F(x)$; в случае бесконечных границ результат не изменится). Здесь, конечно, предполагалось, что $\forall n:\ c_n \in [a, b)$.

Здесь интеграл Стилтьеса свёлся к ряду. В частности, если $F(x)$ - функция распределения атомической(вырожденной) случайной величины ($\mathbf{P}(\xi = c) = 1$), то

$$\int_{-\infty}^{\infty} g(x)dF(x) = g(c)(F(c+0) - F(c-0)) = g(c)(F(c+0) - F(c)) = g(c).$$

В заключение сформулируем некоторые свойства интеграла Стилтьеса. Мы опустим доказательства, поскольку они очевидным образом следуют из определения интеграла Стилтьеса и стандартных методов, используемых при доказательстве аналогичных свойств интеграла Римана. Во всех свойствах границы могут быть как конечными, так и бесконечными.

1. Пусть $a = c_0 < c_1 < c_2 < ... < c_n < c_{n+1} = b$. Тогда

$$\int_a^b g(x)dF(x) = \sum_{k=0}^{n} \int_{c_k}^{c_{k+1}} g(x)dF(x).$$

2.
$$\int_a^b \alpha g(x)dF(x) = \alpha \int_a^b g(x)dF(x).$$

3.
$$\int_a^b \sum_{k=0}^n g_k(x)dF(x) = \sum_{k=0}^n \int_a^b g_k(x)dF(x).$$

4. Пусть $g(x) \geq 0$ и $a < b$. Тогда
$$\int_a^b g(x)dF(x) \geq 0.$$

5. Пусть $g(x) \leq h(x)$. Тогда
$$\int_a^b g(x)dF(x) \leq \int_a^b h(x)dF(x).$$

6. Если $F(x) \equiv C$ при $a \leq x \leq b$, то
$$\int_a^b g(x)dF(x) = 0.$$

7.
$$\left| \int_a^b g(x)dF(x) \right| \leq \int_a^b |g(x)|dF(x).$$

8.
$$\int_a^b g(x)d(\alpha_1 F_1(x) + \alpha_2 F_2(x)) = \alpha_1 \int_a^b g(x)dF_1(x) + \alpha_2 \int_a^b g(x)dF_2(x).$$

Замечание. Если потребовать $\alpha_1 \geq 0$, $\alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$, то функция $\alpha_1 F_1(x) + \alpha_2 F_2(x)$ также является функцией распределения (разумеется, если функциями распределения являлись $F_1(x)$ и $F_2(x)$). На самом деле эти условия необязательны.

Литература

1. А.А. Боровков, *Теория вероятностей*, Наука, **1986**.
2. Е.С. Вентцель, *Теория вероятностей*, Высшая школа, **2001**.
3. В.Г. Воинов, М.С. Никулин, *Несмещенные оценки и их применения*, Наука, **1989**.
4. Б.В. Гнеденко, *Курс теории вероятностей*, Наука, **1965**.
5. В.А. Тутубалин, *Теория вероятностей*, Издательство Московского Университета, **1972**.
6. В.А. Тутубалин, *Теория вероятностей и случайных процессов*, Издательство Московского Университета, **1992**.

ОГЛАВЛЕНИЕ