МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. ЛОМОНОСОВА
Механико-математический факультет

Курс лекций по теории вероятностей

Лектор — Холево Александр Семенович

II курс, 4 семестр, поток математиков

Москва, 2006 г.

Оглавление

Введение

Предисловие

Текст набирался в различное время тремя наборщиками: Д. Колосовым, Д. Мануйловым и С. Кузнецовым. Потом он был существенно отредактирован *DMVN Corporation*.

Большая просьба к читателям сообщать об ошибках и опечатках авторам.

На данном этапе возможны «дыры» в материале курса, которые образовались из-за слияния работы трёх наборщиков. Отнеситесь синсходительно и лучше чётко скажите, чего не хватает, иже такое заметите.

Раздел «Суровые вопросы» пока остаётся таким, как был. Он немного перевёрстан, часть формул сделана выключными, чтобы было легче читать и сложнее списывать :) .

В этой редакции исправлена огромная куча опечаток, замеченных при подготовке одним из студентов, коему большое спасибо (фамилия выясняется).

Последняя компиляция: `11 июня 2006 г.`
Обновления документа — на сайтах `http://dmvn.mexmat.net`, `http://dmvn.mexmat.ru`.
Об опечатках и неточностях пишите на `dmvn@mccme.ru`.

1. Основные понятия

1.1. Элементарные понятия теории вероятностей

1.1.1. События и их вероятности

Определение. *Детерминированное явление* A (событие) — это событие, которое всегда выполняется.

Определение. *Недетерминированные события* — случайные события, такие события и будут нас интересовать.

Рассмотрим n повторений опыта, в котором может произойти событие A. Пусть $n(A)$ — количество тех опытов, где A выполнилось, тогда частота события A: $\nu(A) = \frac{n(A)}{n}$.

Замечание. Замечено, что $\nu(A)$ сгущаются вокруг некоторого конечного значения $p(A)$.

Пример 1.1. Бросание монет. Под событием в данном случае можно понимать выпадание решки.

n	n(a)	$\nu(A)$
4040	2048	0.508
24000	12012	0.5005

В данном случае $p(A) = \frac{1}{2}$.

Вероятность события A — теоретическое модельное значение, к которому приближается $\nu(A)$ при бесконечно большом n.

Свойства частот:

1. $\nu_n(A) \geqslant 0$;
2. $\nu_n(\Omega) = 1$, где Ω — детерминированное событие;
3. $\nu_n(A \cup B) = \nu_n(A) + \nu_n(B)$, если $A \cap B = \varnothing$.

1.1.2. Примеры вероятностных моделей

Пример 1.2. Опыт с конечным числом равновероятных событий. Пусть $\omega_1, \omega_2, \ldots, \omega_n$ — элементарные события, и в опыте может произойти одно и только одно из них, тогда $\Omega = \{\omega_1, \ldots, \omega_n\}$ — пространство элементарных событий, или вероятностное пространство.

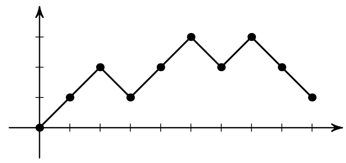

Рис. 1

Будем бросать монету (наше вероятностное пространство будет иметь всего два элемента) и рассматривать *кривую случайного блуждания* — график, который поднимается на 1 вверх, если выпадает орёл, и опускается на 1 вниз, если выпадает решка. Пример такого графика изображён на рис. 1.

Пример 1.3. При игре в рулетку $\Omega = \{0, 1, \ldots, 36\}$.

Определение. A называется *событием*, если $A \subset \Omega$, в частности, элементарное событие — это событие.

Определение. Событие $A \subset \Omega$, $\Omega = \{\omega_1, \ldots, \omega_n\}$ *осуществимо* $\Leftrightarrow$ осуществимо одно из элементарных событий, составляющих A.

Определение. *Достоверное событие* — Ω. *Невозможное событие* — $\varnothing$ (событие, которое не может произойти).

Вероятность события A в опыте с равновероятными исходами определяется по формуле $P = \frac{|A|}{|\Omega|}$.

Пример 1.4. На рулетке $P(\text{чет}) = P(\text{нечет}) = \frac{18}{37} < \frac{1}{2}$.

Задача 1.1 (Д'Аламбера). *Какова вероятность того, что при двух бросаниях монеты орёл выпадет хотя бы один раз?*

Решение. Пусть «O» — орёл, а «P» — решка. Тогда $\Omega = \{OO, PO, OP, PP\}$. Нас устраивают 3 исхода из 4, значит, $P(A) = \frac{3}{4}$. ■

1.1.3. Комбинаторные формулы

Определение. Числом размещений из N по n называется число способов разместить n различных элементов на N местах. Обозначается оно через A_N^n.

Легко видеть, что

$$A_N^n = N(N-1)\ldots(N-n+1) = \frac{N!}{(N-n)!}. \tag{1}$$

При $N = n$ получаем количество перестановок из N элементов: $S_N = N!$

Число сочетаний из N по n отличается от числа размещений тем, что мы не различаем элементы:

$$C_N^n = \frac{A_N^n}{n!} = \frac{N!}{(N-n)!n!}. \tag{2}$$

В частном случае $C_N^0 = A_N^0 = 0! = 1$.

Задача 1.2. *Выборочный контроль качества. Партия из N изделий, среди которых M бракованных. Наугад выбирается n изделий. Какова вероятность того, что среди выбранных изделий окажется ровно m бракованных?*

Решение. Элементарное событие в нашем случае — произвольная выборка, значит $|\Omega| = C_N^n$. Надо найти вероятность события $A_m = \{$из n выбранных ровно m бракованных$\}$. Выбираем из M бракованных изделий m, а из $(N-M)$ нормальных изделий — $(n-m)$. Тогда среди n выбранных изделий будет ровно m бракованных и $|A_m| = C_M^m \cdot C_{N-M}^{n-m} \implies P(A_m) = \dfrac{C_M^m \cdot C_{N-M}^{n-m}}{C_N^n}$ — гипергеометрическое распределение вероятности. ■

Задача 1.3. *Лотерея. Есть N билетов, из которых M выигрышных. Какова вероятность выигрыша у того, кто купил n билетов?*

Решение. Вероятность того, что не будет выигрышных билетов: $P(m=0) = P(A_0) = \dfrac{C_M^0 \cdot C_{N-M}^n}{C_N^n}$. Тогда вероятность того, что будут выигрышные: $P(m>0) = 1 - P(m=0) = 1 - \dfrac{C_M^0 \cdot C_{N-M}^n}{C_N^n}$. ■

1.1.4. Опыт с непрерывным пространством элементарных событий

Элементарное событие $\omega = (x, y) \in \Omega$. $A \subset \Omega \implies P(A) = \frac{\operatorname{mes} A}{\operatorname{mes} \Omega}$. Отсюда следует, что A и Ω должны быть измеримы.

Задача 1.4. *Отрезок $\big[0, 1\big]$ разламывают в 2-х местах случайным образом. Какова вероятность того, что из полученных кусков можно составить треугольник?*

Решение. Пусть отрезок разбивается на $x, y-x, 1-y$. Рассмотрим случай $x \leqslant y$ (второй вариант симметричен). Запишем неравенство треугольника для всех трех сторон:

$$\begin{cases} x + (y-x) > 1 - y \\ x + (1-y) > y - x \\ (y-x) + (1-y) > x \end{cases} \Leftrightarrow \begin{cases} y > \dfrac{1}{2} \\ y - x < \dfrac{1}{2} \\ x < \dfrac{1}{2} \end{cases}.$$

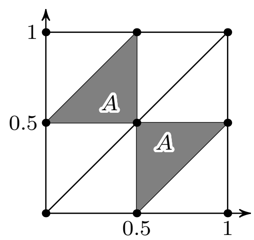

Рис. 2

Из графика видно, что $\operatorname{mes} A = 2 \cdot \frac{1}{8} = \frac{1}{4} \cdot \operatorname{mes} \Omega$. ■

1.2. Строгое определение вероятности. Аксиоматика Колмогорова

1.2.1. Основные определения

Теория вероятности	Теория множеств
Достоверное, невозможное событие	Ω, $\varnothing$
A_1 влечет A_2	$A_1 \subset A_2$
A_1 или(и) A_2	$A_1 \cup A_2$, $\bigcup\limits_{i=1}^{n} A_i$
A_1 и A_2	$A_1 \cap A_2$, $\bigcap\limits_{i=1}^{n} A_i$
A_1 и A_2 несовместны	$A_1 \cap A_2 = \varnothing$
не A	$\overline{A} = \{$совокупность ω, не входящих в A$\}$
Выполняется A_2, но не A_1	$A_2 \setminus A_1 = A_2 \cap \overline{A}_1$

Определение. Пусть Ω — некоторое множество. Тогда назовем класс $\mathscr{A} \in 2^\Omega$ подмножеств Ω *алгеброй множеств*, если:

1. $\Omega \in \mathscr{A}$;

2. $A_1, A_2 \in \mathscr{A} \implies A_1 \cup A_2 \in \mathscr{A}$ и $A_1 \cap A_2 \in \mathscr{A}$;
3. $A \in \mathscr{A} \implies \overline{A} \in \mathscr{A}$.

Свойства алгебры множеств таковы:

1. $\varnothing \in \mathscr{A}$;
2. $A_1, \ldots, A_n \in \mathscr{A} \implies \bigcup\limits_{i=1}^{n} A_i \in \mathscr{A}, \bigcap\limits_{i=1}^{n} A_i \in \mathscr{A}$.

Определение. Алгебра $\mathscr{A}$ называется *σ-алгеброй*, если выполнено условие:

2′. $A_1, A_2, \ldots \in \mathscr{A} \implies \bigcup\limits_{i=1}^{\infty} A_i \in \mathscr{A}, \bigcap\limits_{i=1}^{\infty} A_i \in \mathscr{A}$.

1.2.2. Вероятностное пространство и аксиомы Колмогорова

Определение. *Вероятностное пространство* — это тройка $(\Omega, \mathscr{A}, P)$, где Ω — пространство элементарных событий; элементы Ω называются элементарными событиями; $\mathscr{A}$ — σ-алгебра подмножеств Ω, элементы $\mathscr{A}$ называются событиями; P — числовая функция, определенная на $\mathscr{A}$ ($P\colon \mathscr{A} \to \mathbb{R}$). Для $A \in \mathscr{A}$, $P(A)$ — вещественное число, которое называется *вероятностью* события A.

Следующие свойства вероятности P на $(\Omega, \mathscr{A}, P)$ называются *аксиомами Колмогорова.*

1. $P(A) \geqslant 0,\ \forall\, A \in \mathscr{A}$;
2. $P(\Omega) = 1$;
3. Если A и B несовместные события ($A \cap B = \varnothing$), то $P(A \cup B) = P(A) + P(B)$;
4. Если $B_1 \supset B_2 \supset \ldots \supset B_n \supset \ldots$ — убывающая последовательность событий, причем $\bigcap\limits_{i=1}^{\infty} B_i = \varnothing$, то $\lim\limits_{n\to\infty} P(B_n) = 0$.

Пример 2.1. Дискретное пространство элементарных событий. $\Omega = \{\omega_1, \ldots, \omega_n, \ldots\}$, $\mathscr{A} = 2^{\Omega}$ — класс всех подмножеств в Ω. Пусть дана последовательность $\{p_1, \ldots, p_n, \ldots\}$, $p_n \geqslant 0$, $\sum\limits_{n=1}^{\infty} p_n = 1$. Тогда $P(A) = \sum\limits_{i\colon\, \omega_i \in A} p_i$.

Задача 1.5. *Доказать, что это действительно вероятностное пространство. Это некоторое обобщение на счетные множества.*

Следствия аксиом Колмогорова:

а. $P(\overline{A}) = 1 - P(A)$.
Это следует из того, что $A \cap \overline{A} = \varnothing, A \cup \overline{A} = \Omega$, откуда $1 = P(\Omega) = P(A \cup \overline{A}) = P(A) + P(\overline{A})$. В частности $P(\varnothing) = 0$.

б. Монотонность: Если $A_1 \subset A_2$, то $P(A_1) \leqslant P(A_2)$.
$A_2 = A_1 \cup (A_2 \setminus A_1) \implies P(A_2) = P(A_1) + P(A_2 \setminus A_1) \geqslant P(A_1)$, так как $A_1 \cap (A_2 \setminus A_1) = \varnothing$, $P(A_2 \setminus A_1) \geqslant 0$.
Отсюда, в частности, имеем $P \in [0, 1]$, так как $\varnothing \subseteq A \subseteq \Omega$.

в. $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$.
Докажем это: $A_1 = (A_1 \cap A_2) \cup (A_1 \setminus A_2)$; по аксиоме 3 $\Big((A_1 \cap A_2) \cap (A_1 \setminus A_2) = \varnothing\Big)$ имеем

$$P(A_1) = P(A_1 \cap A_2) + P(A_1 \setminus A_2). \tag{3}$$

Добавим к обеим частям $P(A_2)$, и получим, что

$$P(A_1) + P(A_2) = P(A_1 \cap A_2) + P(A_1 \setminus A_2) + P(A_2) = P(A_1 \cap A_2) + P(A_1 \cup A_2). \tag{4}$$

Откуда,

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2). \tag{5}$$

Задача 1.6. *Есть 2 кости. Какова вероятность того, что хотя бы на одной будет $\geqslant 5$ очков.*

Решение. Всего $6 \times 6 = 36$ элементарных событий, вероятность элементарного события равна $\frac{1}{36}$. Пусть $A_1 = \{\text{на 1–ой кости} \geqslant 5\}, A_2 = \{\text{на 2–ой кости} \geqslant 5\}$, тогда

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = \frac{12}{36} + \frac{12}{36} - \frac{4}{36} = \frac{5}{9}.$$

■

г. Конечная аддитивность: $A_1, \dots, A_n$ — попарно несовместные события, то есть $(A_i \cap A_j = \varnothing, i \neq j)$. Тогда $P\big(\bigcup\limits_{i=1}^{n} A_i\big) = \sum\limits_{i=1}^{n} P(A_i)$. Доказывается по индукции.

3*. σ-аддитивность: Пусть $A_1, \dots, A_n, \dots$ — попарно несовместные события. Тогда справедливо равенство:

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

Это мы докажем ниже.

1.2.3. Теорема равносильности систем аксиом

Теорема 1.1. *Система аксиом 1–4 (Аксиомы Колмогорова) равносильна системе аксиом* 1, 2, 3*.

□ $\boxed{3, 4 \Rightarrow 3^*}$ Пусть $A_1, \dots, A_n$ — попарно несовместные события. Обозначим $B_n = \bigcup\limits_{i=n+1}^{\infty} A_i$ — остаток ряда. Тогда $\bigcup\limits_{i=1}^{\infty} A_i = A_1 \cup \ldots \cup A_n \cup B_n$. Используя конечную аддитивность, получаем

$$P(\bigcup_{i=1}^{\infty} A_i) = P(A_1) + \ldots + P(A_n) + P(B_n). \tag{6}$$

Ясно, что $B_1 \supset B_2 \supset \ldots \supset B_n \supset \ldots$ и $\bigcap\limits_{n=1}^{\infty} B_n = \varnothing$ — это следует из того, что $\omega \in \bigcap\limits_{n=1}^{\infty} B_n \Leftrightarrow \omega \in B_n$ для $\forall n \in \mathbb{N}$; значит $\exists i_0 \colon \omega \in A_{i_0}$, но тогда $\forall n > i_0$ выполняется $\omega \notin B_n$, так как события A_i попарно несовместны. Тогда из 4 имеем $P(B_n) \to 0 \implies P(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} P(A_i)$.

$\boxed{3^* \Rightarrow 3}$ — очевидно.

$\boxed{3^* \Rightarrow 4}$ Рассмотрим $B_1 \supset \ldots \supset B_n \supset \ldots$ и $\bigcap\limits_{n=1}^{\infty} B_n = \varnothing$. Пусть $A_n = B_n \setminus B_{n+1}$. Легко видеть, что A_n несовместны и $B_n = \bigcup\limits_{k=n}^{\infty} A_k$. Используя счетную аддитивность, получаем

$$P(B_1) = P(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k), \quad P(B_n) = P(\bigcup_{k=n}^{\infty} A_k) = \sum_{k=n}^{\infty} P(A_k). \tag{7}$$

Значит, $P(B_n)$ — остаток сходящегося ряда $\sum\limits_{k=1}^{\infty} P(A_k)$, и, следовательно, $P(B_n) \to 0$. ■

1.3. Условные вероятности. Формула полной вероятности. Формула Байеса

1.3.1. Условная вероятность

Пример 3.1. Пусть A — множество курящих, B — множество больных. Если эти множества пересекаются, то вводится понятие *относительной частоты* $\nu_n(B|A) = \frac{n(B\cap A)}{n(A)}$. Если $n \gg 1$, то

$$n(A) \sim P(A) \cdot n, \quad n(B \cap A) \sim P(B \cap A) \cdot n \implies \nu_n \sim \frac{P(B \cap A)}{P(A)}.$$

Определение. Пусть $(\Omega, \mathscr{A}, P)$ — вероятностное пространство. Пусть $A \in \mathscr{A}$ и $P(A) > 0$. Тогда *условной вероятностью* $B \in \mathscr{A}$ при условии A называют $P(B|A) := \dfrac{P(B \cap A)}{P(A)}$.

Замечание. Если $B \subset A$,то $P(B|A) = \frac{P(B)}{P(A)}$. Если $B \cap A = \varnothing$, то $P(B|A) = 0$.

Пример 3.2. Распад радиоактивного атома. Пусть мы знаем, что:

1. Вероятность того, что атом не распадется до $t_0 + t$, при условии, что он не распался до t_0, зависит только от t и не зависит от t_0.
2. Эта вероятность стремится к 1 при $t \to 0$.

Найти закон распада.

Решение. $A(t) := \{$Атом не распался до момента времени $t\}$. Нам известно, что:

1. $p(t) = P(A(t+t_0)|A(t_0))$;
2. $\lim\limits_{t\to 0} p(t) = 1$.

$A(t_0+t) \subset A(t_0)$, поэтому

$$p(t+s) = \frac{P(A(t_0+t+s))}{P(A(t_0))} = \frac{P(A(t_0+t+s))}{P(A(t_0+t))} \cdot \frac{P(A(t_0+t))}{P(A(t_0))} = p(s)\cdot p(t) \implies p(t+s) = p(s)\cdot p(t). \tag{8}$$

С учётом условия 2 получаем экспоненциальный закон распада. ■

Определение. Семейство A_i $\{A_i \in \mathscr{A} \mid i = 1, 2, \ldots, n\}$ называется *разбиением* Ω, если $\bigsqcup\limits_{i=1}^{n} A_i = \Omega$.

1.3.2. Формула полной вероятности

Теорема 1.2 (Формула полной вероятности). *Пусть* $\{A_i\}$ *— разбиение* Ω, $P(A_i) > 0, B \in \mathscr{A}$. *Тогда*

$$P(B) = \sum_{i=1}^{n} P(B|A_i)\cdot P(A_i). \tag{9}$$

□ $B = \Big(B\cap A_1\Big) \cup \ldots \cup \Big(B \cap A_n\Big)$. Так как $\{A_i\}$ — разбиение Ω, то все $(B\cap A_i)$ несовместны, поэтому $P(B) = \sum\limits_{i=1}^{n} P(B\cap A_i) = \sum\limits_{i=1}^{n} P(B|A_i)\cdot P(A_i)$. Тут мы использовали формулу условной вероятности. ■

1.3.3. Формула Байеса

Теорема 1.3 (Формула Байеса). *Пусть* $\{A_i\}$ *— разбиение* Ω, $P(A_i) > 0$, $B \in \mathscr{A}$, $P(B) > 0$. *Тогда выполняется:*

$$P(A_i|B) = \frac{P(B|A_i)\cdot P(A_i)}{\sum\limits_{k=1}^{n} P(B|A_k)\cdot P(A_k)}. \tag{10}$$

□ Имеем

$$P(A_i|B) = \frac{P(A_i\cap B)}{P(B)}. \tag{11}$$

Применим формулу полной вероятности для $P(B)$, а $P(A_i \cap B)$ запишем как $P(B|A_i)\cdot P(A_i)$. Тогда наша условная вероятность запишется в виде:

$$P(A_i|B) = \frac{P(B|A_i)\cdot P(A_i)}{\sum\limits_{k=1}^{n} P(B|A_k)\cdot P(A_k)}. \tag{12}$$

■

Пример 3.3. Партия состоит из $N_1+N_2+N_3$ изделий, выпускаемых соответственно 1-м, 2-м и 3-м заводами. Каждому заводу соответствует процент брака P_1, P_2, P_3. Наугад выбранное изделие оказывается бракованным. Найти вероятность p_j того, что оно было выпущено j-м заводом.

Решение. Пусть $B = \{$Изделие бракованное$\}$, $A_j = \{$Изделие выпущено j-м заводом$\}$. Тогда

$$p_j = P(A_j|B) = \frac{N_j\cdot P_j}{N_1\cdot P_1 + N_2\cdot P_2 + N_3\cdot P_3}. \tag{13}$$

■

1.4. Независимость. Схема Бернулли

Определение. События $A, B \in \mathscr{A}$ называются *независимыми*, если $P(A\cap B) = P(A)\cdot P(B)$.

Замечание. Названо так, потому что если A и B независимы и $P(A) > 0$, тогда $P(B|A) = P(B)$, то есть A не влияет на B.

Замечание. Если A и B независимы, то A и $\overline{B}$; $\overline{A}$ и B; $\overline{A}$ и $\overline{B}$ — тоже независимы.

Пример 4.1. Из колоды в 36 карт вытягивается 1 карта. A — вытянули пику; B — вытянули даму. Проверим зависимость A и B.

Решение. $P(\text{элементарного события}) = \frac{1}{36}$, $P(A \cap B) = \frac{1}{36}$, $P(A) = \frac{1}{4}$, $P(B) = \frac{1}{9} \implies P(A \cap B) = P(A) \cdot P(B)$, откуда следует, что события A и B независимы. ■

Случай нескольких событий $A_1, \ldots, A_n$:

Определение. *Попарно независимы*: если A_i, A_j $(\forall\, i, j; i \neq j)$ независимы.

Определение. *Независимы в совокупности*: $\forall\, A_{i_1}, \ldots, A_{i_k}$ имеем $P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot \ldots \cdot P(A_{i_k})$.

Задача 1.7. *Покажем, что из попарной независимости* **не следует** *независимость в совокупности.*

Решение. Пусть монета бросается 2 раза. Рассмотрим события A={в первый раз выпал орёл}, B={во второй раз выпал орёл}, C={орёл выпал ровно 1 раз}. Тогда $P(A) = P(B) = P(C) = \frac{1}{2}$, $P(A \cap C) = P(A \cap B) = P(B \cap C) = \frac{1}{4}$, значит события A, B, C попарно независимы, но $P(A \cap B \cap C) = 0$, следовательно они *не независимы* в совокупности. ■

Определение. *Схема Бернулли*: последовательность n одинаковых испытаний, в каждом из которых с вероятностью p происходит успех, а с вероятностью $(1-p)$ — неудача.

Пример 4.2. Пусть в мишень производится n независимых выстрелов. Какова вероятность попадания в мишень хотя бы 1 раз?

Решение. Вероятность попадания в мишень будет $P = 1 - (1-p)^n$, так как $(1-p)^n$ — вероятность промаха во всех выстрелах. ■

Вероятностное пространство схемы Бернулли:

«Успех»=1, «неудача»=0 $\implies$ элементарное событие $\omega = (\omega_1, \ldots, \omega_n)$, где $\omega_i \in \{0, 1\}$. $\Omega = \{\omega\}$, $|\Omega| = 2^n$, $\mathscr{A} = 2^\Omega$ (то есть $\mathscr{A}$ — множество всех подмножеств Ω). $P(\omega) = p^m(1-p)^{n-m}$, где m — количество успехов в ω. $P(A) = \sum\limits_{\omega \in A} P(\omega)$. Обозначим событие A_m ={в ω ровно m успехов}, тогда $P(A_m) = C_n^m p^m (1-p)^{n-m}$. Это будет *биномиальное распределение.* Проверим корректность такого определения вероятности:

$$\sum_{m=0}^{n} P(A_m) = \sum_{m=0}^{n} C_n^m p^m (1-p)^{n-m} \stackrel{!}{=} (p + (1-p))^n = 1. \tag{14}$$

Переход «!» следует из формулы бинома Ньютона.

1.5. Простейшие предельные теоремы

1.5.1. Теорема Бернулли

Теорема 1.4 (Бернулли (теор. аналог устойчивости частот)). *Для любого $\varepsilon > 0$ существует предел*

$$\lim_{n \to \infty} P\left\{A_m \colon \left|\frac{m}{n} - p\right| < \varepsilon\right\} = 1,$$

где n — число испытаний в схеме Бернулли, m — число успехов, p — вероятность успеха.

□ Очевидно, что

$$P\left\{A_m \colon \left|\frac{m}{n} - p\right| < \varepsilon\right\} = 1 - P\left\{A_m \colon \left|\frac{m}{n} - p\right| \geqslant \varepsilon\right\}.$$

Тогда

$$P\left\{A_m \colon \left|\frac{m}{n} - p\right| \geqslant \varepsilon\right\} = \sum_{m \colon \left|\frac{m}{n} - p\right| \geqslant \varepsilon} P(A_m) \stackrel{!}{\leqslant} \sum_{m=0}^{n} \left(\frac{m - np}{\varepsilon n}\right)^2 P(A_m) = \frac{1}{\varepsilon^2 n^2} \sum_{m=0}^{n} (m - np)^2 P(A_m). \tag{15}$$

Неравенство «!» обосновано тем, что $\dfrac{|m - np|}{\varepsilon n} = \dfrac{\left|\frac{m}{n} - p\right|}{\varepsilon} \geqslant 1$.

$$\sum_{m=0}^{n} (m - np)^2 P(A_m) = \sum_{m=0}^{n} m^2 P(A_m) - \sum_{m=0}^{n} (2mnp) P(A_m) + \sum_{m=0}^{n} n^2 p^2 P(A_m). \tag{16}$$

Рассмотрим

$$\sum_{m=0}^{n} x^m P(A_m) = \sum_{m=0}^{n} C_n^m x^m p^m (1-p)^{n-m} \stackrel{!}{=} (xp + 1 - p)^n. \tag{17}$$

Переход «!» следует из формулы бинома Ньютона.

Продифференцируем по x и подставим $x = 1$:

$$\sum_{m=0}^{n} mC_n^m p^m (1-p)^{n-m} = np \Leftrightarrow \sum_{m=0}^{n} mP(A_m) = np. \tag{18}$$

Продифференцируем еще раз и снова подставим $x = 1$:

$$\sum_{m=0}^{n} m(m-1)C_n^m p^m (1-p)^{n-m} = n(n-1)p^2 \Leftrightarrow \sum_{m=0}^{n} m(m-1)P(A_m) = n(n-1)p^2. \tag{19}$$

В итоге получаем

$$\begin{aligned}\sum_{m=0}^{n}(m-np)^2P(A_m) = \sum_{m=0}^{n} m^2P(A_m) - \sum_{m=0}^{n}(2mnp)P(A_m) + \sum_{m=0}^{n} n^2p^2P(A_m) = \\ = n(n-1)p^2 + np - 2n^2p^2 + n^2p^2 = np(1-p),\end{aligned} \tag{20}$$

значит $P\left\{m\colon \left|\frac{m}{n} - p\right| \geqslant \varepsilon\right\} \leqslant \dfrac{1}{\varepsilon^2 n^2} \sum\limits_{m=0}^{n} (m-np)^2 P(A_m) = \dfrac{p(1-p)}{\varepsilon^2 n} \to 0$ при $n \to \infty$, следовательно

$$\lim_{n\to\infty} P\left\{m\colon \left|\frac{m}{n} - p\right| < \varepsilon\right\} = \lim_{n\to\infty}\left(1 - P\left\{m\colon \left|\frac{m}{n} - p\right| \geqslant \varepsilon\right\}\right) = 1. \tag{21}$$

■

Принцип малых вероятностей: событие малой вероятности следует рассматривать как невозможное при единичном испытании.

1.5.2. Теоремы Муавра – Лапласа и Пуассона

Теорема 1.5 (Муавра – Лапласа). *Пусть в схеме Бернулли $0 < p < 1$. Тогда при $n \to \infty$,*

$$P(A_m) = \frac{1}{\sqrt{2\pi np(1-p)}} \cdot e^{-\frac{x^2}{2}}(1 + o(1)), \tag{22}$$

где $x = \dfrac{m-np}{\sqrt{np(1-p)}}$, а $o(1)$ — равномерно по $|x| \leqslant C$.

Эта теорема доказывается с использованием формулы Стирлинга. Она является следствием ЦПТ (центральная предельная теорема).

Теорема 1.6 (Пуассона). *(Полезна при $p \to 0, p \to 1$). Рассмотрим последовательность схем Бернулли, где $p_n = \frac{\lambda}{n}$ $(\lambda > 0)$. Тогда $P_n(A_m) \to \frac{\lambda^m}{m!}e^{-\lambda}$, при $n \to \infty$.*

□

$$\begin{aligned}P_n(A_m) = \frac{n!}{m!(n-m)!}\left(\frac{\lambda}{n}\right)^m\left(1 - \frac{\lambda}{n}\right)^{n-m} = \\ = \frac{\lambda^m}{m!} \cdot \frac{n(n-1)\cdot\ldots\cdot(n-m+1)}{n\cdot n\cdot\ldots\cdot n} \cdot \left(1 - \frac{\lambda}{n}\right)^{n-m} \to \frac{\lambda^m}{m!}e^{-\lambda} \text{ при } n \to \infty.\end{aligned} \tag{23}$$

■

$\dfrac{\lambda^m}{m!}e^{-\lambda}$ — распределение Пуассона редких событий.

В применениях n велико, а $np = \lambda \sim 1$, p — малое фиксированное.

Пример 5.1. Рассмотрим $n \gg 1$ независимых радиоактивных атомов. Вероятность распада за время t для атома: $p = 1 - e^{\lambda t} \sim \frac{\lambda T}{n}, t = \frac{T}{n}$. Вероятность того, что к моменту T распадется m атомов, равна $\frac{(\lambda T)^m}{m!}e^{-\lambda T}$.

2. Случайные величины. Функции распределения

2.1. Определения и примеры

2.1.1. Случайные величины

Случайная величина $X(\omega)$ — числовая функция от элементарного события.

Определение. Пусть $(\Omega, \mathscr{A}, P)$ — вероятностное пространство. *Случайной величиной* X на Ω называется функция $X\colon \Omega \to \mathbb{R}$, измеримая относительно $\mathscr{A}$, т.е. $\forall\, x \in \mathbb{R}$ имеем $X^{-1}(-\infty; x) = \{\omega\colon X(\omega) < x\} \in \mathscr{A}$.

Замечание. События $\{\omega\colon X(\omega) \leqslant x\}$, $\{\omega\colon X(\omega) \geqslant x\}$, $\{\omega\colon x_1 \leqslant X(\omega) \leqslant x_2\}$, $\{\omega\colon X(\omega) = x\}$ и тому подобные события тоже лежат в $\mathscr{A}$.

Решение.

1. $\{\omega\colon X(\omega) \leqslant x\} = \bigcap\limits_{n=1}^{\infty} \left\{\omega\colon X(\omega) < x + \frac{1}{n}\right\}$, переходя к пределу получаем нужный результат, по свойству σ-алгебры бесконечное пересечение принадлежит $\mathscr{A}$.
2. $\{\omega\colon X(\omega) \geqslant x\} = \overline{\{\omega\colon X(\omega) < x\}} \in \mathscr{A}$, поскольку со всяким событием в σ-алгебру входит и его дополнение.
3. $\{\omega\colon x_1 \leqslant X(\omega) \leqslant x_2\} = \{\omega\colon x_1 \leqslant X(\omega)\} \cap \{\omega\colon X(\omega) \leqslant x_2\}$.
4. Надо положить $x_1 = x = x_2$ и применить 3.

■

Определение. Случайная величина X называется *дискретной*, если она может принимать только конечное или счетное число различных значений.

Пусть $X(\omega)$ — дискретная случайная величина. Перенумеруем образ: $X(\Omega) = \{x_1, x_2, \ldots\}$. Положим $A_i = \{\omega\colon X(\omega) = x_i\}$; $A_1, \ldots, A_n, \ldots$ попарно не пересекаются; $\{A_n\}$ — (счетное) разбиение Ω; $\bigcup\limits_{n=1}^{\infty} A_n = \Omega$; $p_i = P(A_i) = P\{\omega\colon X(\omega) = x_i\}$; $\sum\limits_{i=1}^{\infty} p_i = 1$; $p_i \geqslant 0$. Тогда $\Big\{(x_i, p_i)\Big\}$ называется *распределением вероятностей* дискретной случайной величины X.

Пример 1.1. *Биномиальное распределение.* $X(\omega)$ — число успехов в n испытаниях.

$$P\{\omega\colon X(\omega) = m\} = C_n^m p^m (1-p)^{n-m}. \tag{1}$$

2.1.2. Функции распределения

Напомним, что на $(\Omega, \mathscr{A}, P)$ была введена функция $X : \Omega \to \mathbb{R}$, такая что $\forall\, x \in \mathbb{R}$ имеем $X^{-1}(-\infty; x) \in \mathscr{A}$.

Определение. *Функцией распределения* случайной величины X называется $F(x) = P\big\{\omega\colon X(\omega) < x\big\}$ или $F(x) := P\big\{X^{-1}(-\infty; x)\big\}$.

Теорема 2.1. *$F(x)$ — функция распределения обладает свойствами:*

1° *$F(x)$ возрастает на $\mathbb{R}$;*

2° *$F(-\infty) = 0$, $F(+\infty) = 1$;*

3° *$F(x)$ непрерывна слева.*

□

1° Пусть $x_1 < x_2$, $A_1 := X^{-1}(-\infty; x_1)$, $A_2 := X^{-1}(-\infty; x_2)$. Так как $x_1 < x_2$, то $A_1 \subset A_2$, поэтому $P(A_1) \leqslant P(A_2) \Leftrightarrow F(x_1) \leqslant F(x_2)$.

2° Пусть $x_n \searrow -\infty$. Покажем, что $F(x_n) \to 0$. Рассмотрим $A_n := X^{-1}(-\infty; x_n)$. Очевидно, что $A_1 \supset A_2 \supset A_3 \ldots$ и $\bigcap\limits_{n=1}^{\infty} A_n = \varnothing$, так как если $\omega \in \bigcap\limits_{n=1}^{\infty} A_n$, то $\forall\, x_n$ верно $X(\omega) < x_n$, но это невозможно $\Longrightarrow$ по аксиоме непрерывности вероятности получаем $0 = \lim\limits_{n\to\infty} P(A_n) = \lim\limits_{n\to\infty} F(x_n)$.

Аналогично рассмотрим $x_n \nearrow +\infty$, $A_1 \subset A_2 \subset \ldots$ и $\bigcup\limits_{n=1}^{\infty} A_n = \Omega$, откуда $\overline{A}_1 \supset \overline{A}_2 \supset \ldots \Longrightarrow \bigcap\limits_{n=1}^{\infty} \overline{A}_n = \varnothing$, и поэтому $\lim\limits_{n\to\infty} P(\overline{A}_n) = 0 = 1 - \lim\limits_{n\to\infty} P(A_n)$. Следовательно, $\lim\limits_{n\to\infty} P(A_n) = 1$.

3° Сначала докажем ряд вспомогательных фактов.

Замечание.

а. Пусть $A_1 \supset A_2 \supset \ldots$ Рассмотрим $A := \bigcap\limits_{n=1}^{\infty} A_n$. Тогда $P(A) = \lim\limits_{n\to\infty} P(A_n)$. В самом деле, рассмотрим $A_n^0 := A_n \setminus A$. К ним применим аксиому непрерывности.

б. Если $A_1 \subset A_2 \subset \ldots$, то $P(\bigcup\limits_{n=1}^{\infty} A_n) = \lim\limits_{n\to\infty} P(A_n)$.

Теперь докажем пункт **3°** теоремы.

Рассмотрим $x_n \nearrow x_0$ и положим $A_n := X^{-1}(-\infty; x_n)$, $A_0 := X^{-1}(-\infty; x_0)$. Ясно, что $\bigcup\limits_{n=1}^{\infty} A_n = A_0$ (несложно видеть из того, что $x_n \nearrow x_0$).

Из пункта *б* замечания имеем $\lim\limits_{n\to\infty} P(A_n) = P(A_0)$, где $P(A_n) = F(x_n)$, $P(A_0) = F(x_0)$. Заметим далее, что $P\left(X^{-1}(-\infty; x_0]\right) = \lim\limits_{x_n \to x_0+0} F(x_n) = F(x_0+0)$. Положим $A_0 := X^{-1}(-\infty; x_0]$. Очевидно $\bigcap\limits_{n=1}^{\infty} A_n = A_0 (x_n \downarrow x_0) \implies$ осталось перейти к пределу. Итак, теперь ясно, что $P(X^{-1}[a,b]) = P(X^{-1}(-\infty; b]) - P(X^{-1}(-\infty; a)) = F(b+0) - F(a)$. В итоге получаем: $P\{\omega\colon X(\omega) \in [a,b]\} = F(b+0) - F(a)$, в частности $P\{\omega\colon X(\omega) = a\} = F(a+0) - F(a) =: \Delta F(a)$.

■

Задача 2.1. *Как выглядит функция распределения дискретной случайной величины?*

Определение. Функция распределения $F(X)$ называется абсолютно непрерывной, если $\exists\, p(x)\colon \forall\, a,b (a<b)$ имеем $F(b)-F(a) = \int\limits_a^b p(x)\,dx$ (интегрируемость по Лебегу). Функция $p(x)$ называется плотностью распределения вероятности.

Замечание. Если F абсолютно непрерывна, то она непрерывна.

1. $F'(x) = p(x)$ почти всюду. (Если $p(x) \in$ **SegC** (кусочно-непрерывные), то $F'(x) = p(x)$ во всех точках непрерывности).
2. $a \to -\infty \implies F(x) = \int\limits_{-\infty}^{x} p(\tau)\,d\tau$
3. $p(x) \geqslant 0$ почти всюду.

При $a \to -\infty, b \to +\infty: \int\limits_{-\infty}^{+\infty} p(\tau)\,d\tau = 1.$

Пример 1.2.

1. Равномерное распределение на $[a,b]$: $p(x) = \frac{1}{b-a}\chi_{[a,b]}$;
2. Экспоненциальное распределение: $F(t) = \begin{cases} 0, & t<0; \\ 1-e^{-\lambda t}, & t \geqslant 0. \end{cases}$

Нормальное распределение:
$\int\limits_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,dx = I \implies I^2 = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} e^{-\frac{x^2+y^2}{2}}\,dx\,dy = \int\limits_0^{2\pi} d\varphi \cdot \int\limits_0^{+\infty} e^{-\frac{r^2}{2}} r\,dr = 2\pi \implies I = \sqrt{2\pi} \implies p(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ — плотность распределения. $\Phi(x) := \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tau^2}{2}}\,d\tau$ — функция нормального распределения. Стандартное нормальное распределение: $\mathscr{N}(0,1)$. Еще рассматриваются функции вида

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad \sigma \neq 0, \quad m \in \mathbb{R}.$$

Обозначение таких функций: $\mathscr{N}(m,\sigma)$.

Пример 1.3. Пусть X имеет плотность $p(x)$. Найти распределение $Y = X^2$.

Решение.

$$F_Y(y) = P\{\omega\colon Y(\omega) < y\} = P\{\omega\colon X(\omega)^2 < y\} = P\{\omega\colon -\sqrt{y} < X(\omega) < \sqrt{y}\} = \\ = \int\limits_{-\sqrt{y}}^{\sqrt{y}} p_X(x)\,dx \implies p_Y(y) = \begin{cases} 0, & y \leqslant 0; \\ p_X(\sqrt{y})\dfrac{1}{2\sqrt{y}} + p_X(-\sqrt{y})\dfrac{1}{2\sqrt{y}}, & y \geqslant 0. \end{cases} \tag{2}$$

■

Замечание. Кроме дискретных и абсолютно непрерывных, бывают сингулярные случайные величины. Например функция распределения Кантора.

Теорема Лебега: произвольная функция распределения разлагается как $F(x) = C_1F_1(x) + C_2F_2(x) + C_3F_3(x)$, где $F_i(x)$ — соответственно функции распределения абсолютно непрерывной, сингулярной и дискретной величин.

2.2. Семейства случайных величин. Независимость

Пусть $X_i(\omega)$, $i = 1, \ldots, n$ — случайные величины на $(\Omega, \mathscr{A}, P)$. Иначе говоря, задан случайный вектор $X(\omega) = (X_1(\omega), \ldots, X_n(\omega)) \in \mathbb{R}^n$.

Определение. Функцией совместного распределения $X_1, \ldots, X_n$ называется $F(x_1, \ldots, x_n) = P\big\{\omega \colon X_1(\omega) < x_1; \ldots; X_n(\omega) < x_n\big\} = P(A)$.

Замечание. Функция определена корректно, так как A есть пересечение: $A = \bigcap X_i^{-1}(-\infty; x_i)$, а так как X_i — случайные величины, то каждый элемент пересечения лежит в $\mathscr{A}$.

2.2.1. Основные свойства

1) $F(x_1, \ldots, x_n)$ является неубывающей по каждому аргументу x_j, при фиксированных остальных аргументах, поскольку вероятность возрастающего семейства множеств возрастает.

2) $\lim\limits_{x_j \to -\infty} F(x_1, \ldots, x_j, \ldots, x_n) = 0$ при фиксированных $x_1, \ldots, \widehat{x_j}, \ldots x_n$. Доказательство этого свойства аналогично. $\lim\limits_{x1, \ldots, x_n \to +\infty} F(x_1, \ldots, x_j, \ldots, x_n) = 1$, поскольку любая точка будет захвачена.

3) F непрерывна слева по каждой из переменных. $\lim\limits_{x_2, \ldots, x_n \to +\infty} F(x_1, \ldots, x_n) = F(x_1)$ — функция распределения случайной величины x_1. Можно считать что

$$\begin{aligned} x_2, \ldots, x_n \uparrow +\infty \to \lim_{x_2, \ldots, x_n \uparrow +\infty} F(x_1, \ldots, x_n) = \\ = \lim_{x_2, \ldots, x_n \uparrow +\infty} P\big\{\omega : X_1(\omega) < x_1, \ldots, X_n(\omega) < x_n\big\} = \\ = P\big\{\omega : X_1(\omega) < x_1, X_2 < +\infty, \ldots, X_n(\omega) < +\infty\big\} = P\big\{\omega : X_1 < x_1\big\} = F_1(x_1) \end{aligned} \tag{3}$$

— функция распределения x_1.

4) Вероятность попадания в n-мерный полуинтервал (это $\big\{(x_1, \ldots, x_n) : a_1 \leqslant x_1 < b_1, \ldots, a_n \leqslant x_n < b_n\big\}$)

Рассмотрим $P\big\{\omega : x_1 \leqslant X_1(\omega) < x_1 + h_1, \ldots, x_n \leqslant X_n(\omega) < x_n + h_n\big\} =: \Delta_{h_1}^{(1)} \ldots \Delta_{h_n}^{(n)} F(x_1, \ldots, x_n)$, где $\Delta_{h_j}^{(j)} F(x_1, \ldots, x_n) = F(x_1, \ldots, x_j + h_j, \ldots, x_n) - F(x)$.

$\square$ Доказательство для $n = 2$.

$$\begin{aligned} P\big\{\omega : x_1 \leqslant X_1(\omega) < x_1 + h_1, x_2 \leqslant X_2(\omega) < x_2 + h_2\big\} = \\ = P\big\{\omega : X_1(\omega) < x_1 + h_1, X_2(\omega) < x_2 + h_2\big\} - P\big\{\omega : X_1(\omega) < x_1, X_2(\omega) < x_2 + h_2\big\} - \\ - P\big\{\omega : X_1(\omega) < x_1 + h_1, X_2(\omega) < x_2\big\} + P\big\{\omega : X_1(\omega) < x_1, X_2(\omega) < x_2\big\} = \\ = F(x_1 + h_1, x_2 + h_2) - F(x_1 + h_1, x_2) - F(x_1, x_2 + h_2) + F(x_1, x_2) = \Delta_{h_1}^{(1)} \Delta_{h_2}^{(2)} F(x_1, x_2). \end{aligned} \tag{4}$$

$\blacksquare$

Определение. Пусть $\mathcal{K}$ — некоторый класс подмножеств из $X : \mathcal{K} \subset 2^X$. Пересечение всех σ-алгебр, содержащих $\mathcal{K}$ называется σ-алгеброй, порожденной $\mathcal{K}$. Обозначается как $\sigma(\mathcal{K})$. В частности, *борелевской* σ-алгеброй называется σ-алгебра, порожденная всеми открытыми[1] множествами в $\mathbb{R}^n$. (Обозначается $\mathcal{B}(\mathbb{R}^n)$)

Теорема 2.2. *Пусть $(X_1, \ldots, X_n)$ — случайный вектор в $\mathbb{R}^n$. Пусть B — событие из σ-алгебры $\mathcal{B}(\mathbb{R}^n)$. Тогда $\big\{\omega : (X_1(\omega), \ldots, X_n(\omega)) \in B\big\} \in \mathscr{A}$.*

$\square$ Обозначим через $\sum$ все $B \subset \mathbb{R}^n$: для них выполнено $\big\{\omega : (X_1(\omega), \ldots, X_n(\omega)) \in B\big\} \in \mathscr{A}$.

1) $\sum$ содержит все полуинтервалы в $\mathbb{R}^n$.

2) $\sum$ является σ-алгеброй. Тогда ясно, что $\mathcal{B}(\mathbb{R}^n) \subseteq \sum$.

Пусть $X(\omega)$ — тот самый случайный вектор. $X(\omega) := (X_1(\omega), \ldots, X_n(\omega))$. Тогда $\big\{\omega : (X_1(\omega), \ldots, X_n(\omega)) \in B\big\}$ имеет вид $\big\{\omega : X(\omega) \in B\big\}$.

1) $\mathbb{R}^n \in \sum : \big\{\omega : X(\omega) \in \mathbb{R}^n\big\} = \Omega \in \mathscr{A}$.

2) Пусть $\big\{\omega : X(\omega) \in B\big\} \in \mathscr{A}$, тогда $\big\{\omega : X(\omega) \in \overline{B}\big\} = \big\{\omega : X(\omega) \notin B\big\}$, то есть если $B \in \sum \implies \overline{B} \in \sum$.

3) Если $B_i \in \sum \implies$ покажем, что $\bigcup\limits^{\infty} B_i \in \sum$, $\bigcap\limits^{\infty} B_i \in \sum$.

$\big\{\omega : X(\omega) \subset \bigcup\limits^{\infty} B_i\big\} = \bigcup\limits^{\infty} \big\{\omega : X(\omega) \in B_i\big\} \in \mathscr{A}$, т.к. $\mathscr{A}$ — σ-алгебра. Аналогично доказывается и для $\bigcap\limits^{\infty} B_i$. тем самым доказано, что $\sum$ — σ-алгебра. $\blacksquare$

Определение. $P_X(B) = P\big\{\omega : X(\omega) \in B\big\}$ называется распределением вероятности случайного вектора $X = (X_1, \ldots, X_n)$, либо совместным распределением вероятности. Здесь $B \in \mathcal{B}(\mathbb{R}^n)$.

[1] Можно считать, что это все открытые прямоугольники (так устроена топология в $\mathbb{R}^n$).

В частности, если $B=[x_1,x_1+h_1)\times[x_2,x_2+h_2)\times\ldots\times[x_n,x_n+h_n)$, то $P_X(B)=\Delta^{(1)}_{h_1}\ldots\Delta^{(n)}_{h_n}F(x_1,\ldots,x_n)$. В силу теоремы о продолжении меры (о единственности), $P_X(B)$ однозначно определяется по $F(x_1,\ldots,x_n)$

Определение. Совместное распределение $P_X(B)$ называется *абсолютно непрерывным*, если

$$P_X(B)=\int\limits_B p(x)\,dx,$$

где $B\in\mathcal{B}(\mathbb{R}^n)$, $p(x)$ — плотность совместного распределения, а x — n-мерный вектор.

Если B — полуинтервал, то

$$\Delta_h F(x)=\int\limits_{x_1}^{x_1+h_1}\ldots\int\limits_{x_n}^{x_n+h_n}p(y_1\ldots y_n)dy_1\ldots dy_n.$$

Ясно, что при $n=1$ получается немного более широкое определение, чем для одномерного классического случая (там для полуинтервала из $\mathbb{R}^1$, а здесь — для произвольного $B\in\mathcal{B}(\mathbb{R}^1)$). Тогда :

1) $\frac{\partial^n F(x)}{\partial x_1\ldots\partial x_n}=p(x_1\ldots x_n)$ для всех $(x_1\ldots x_n)$. В частности, для кусочно-непрерывных функций имеем это равенство во всех точках непрерывности.

2) $F(x_1\ldots x_n)=\int\limits_{-\infty}^{x_1}\ldots\int\limits_{-\infty}^{x_n}p(y_1\ldots y_n)dy_1\ldots dy_n$.

3)$P(x)\geqslant 0$ почти всюду (это следует из монотонности F), а $\int\limits_{-\infty}^{+\infty}\ldots\int\limits_{-\infty}^{+\infty}p(y_1\ldots y_n)dy_1\ldots dy_n=1$ — единичная масса, размазанная по всему пространству.

2.2.2. Независимость случайных величин

Определение. Случайные величины $X_1\ldots X_n$ называются независимыми, если их совместное распределение обладает свойствами: $\forall\,B_1,\ldots B_n\in\mathcal{B}(\mathbb{R})$, имеем $P\{\omega: X(\omega)\in B_1\times\ldots\times B_n\}=\prod\limits_{j=1}^{n}P(\omega: X_j(\omega)\in B_j)$.(н.1)

Упростим это определение: Пусть $B_j=[x_j,x_j+h_j)$. Тогда $\Delta_h F(x)=\prod\limits_{j=1}^{n}(\Delta^j_{h_j}F_j(x_j))$ (н.2)

$F(x_1\ldots x_n)=F_1(x_1)\cdot\ldots\cdot F_n(x_n)$ (н.3).

н.2 $\Longrightarrow$ н.1 следует из единственности продолжения меры

н.3 $\Longrightarrow$ н.2 несложно доказывается.

Получаем, что н.1 $\Leftrightarrow$ н.2 $\Leftrightarrow$ н.3

В частности если совместное распределение $X_1\ldots X_n$ абсолютно непрерывно, то н.1 н.2 н.3 $\Leftrightarrow p(x_1\ldots x_n)=p(x_1)\ldots p(x_n)$

Пример 2.1. 1) Пусть $V\subset\mathbb{R}^n$ — ограниченное борелевское множество, Пусть $p(x):=\frac{1}{\operatorname{mes}V}\chi(V)$, тогда $p\geqslant 0$, и $\int\limits_{\mathbb{R}^n}p\,dx=1$, а $P_X(B)=P\{\omega: X(\omega)\in B\cap V\}=\frac{\operatorname{mes}(B\cap V)}{\operatorname{mes}V}$. Здесь $X=(X_1,\ldots,X_n)$. Это равномерное распределение на множестве V. Если $V=[a_1,b_1]\times\ldots\times[a_n,b_n]$, то $X_1\ldots X_n$ — независимые случайные величины.

2)Многомерное нормальное распределение:

$$p(t)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(t-m)^2}{2\sigma^2}},$$

и тогда $p(x_1\ldots x_n)=p_1(x_1)\ldots p_n(x_n)$.

Задача 2.2. *Пусть $Y=f(X_1\ldots X_n)$, f — измеримая функция, а $X_1\ldots X_n$ имеют плотность совместного распределения $p(x_1\ldots x_n)$. Найти плотность распределения Y.*

Решение. $F_y(y)=P\{\omega: Y(\omega)<y\}=P\{\omega: f(X_1(\omega),\ldots X_n(\omega))<y\}=P\{\omega:(X_1(\omega),\ldots,X_n(\omega))\in B_y\}$, где $B_y=\{(x_1,\ldots,x_n): f(x_1,\ldots,x_n)<y\}$. Тогда в силу абсолютной непрерывности

$$F_Y(y)=\int\limits_{B_y}p(x_1,\ldots,x_n)\,dx_1\ldots\,dx_n,$$

откуда $P_Y(y)=\frac{dF_Y(y)}{dy}=\frac{d}{dy}\int\limits_{B_y}p(x_1,\ldots,x_n)\,dx_1\ldots\,dx_n$. ■

Пример 2.2. Рассмотрим конкретный пример: Найти плотность и функцию распределения $Y=X_1+X_2$, где X_1 и X_2 имеют совместную плотность $p(x_1,x_2)$.

Решение. Имеем

$$F_Y(y) = P\{\omega : X_1(\omega) + X_2(\omega) < y\} = \int\limits_{x_1+x_2<y} p(x_1, x_2)\, dx_1\, dx_2 =$$

$$= \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{y-x_1} p(x_1, x_2)\, dx_1\, dx_2 = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{y} p(x_1, x_2 - x_1)\, dx_1\, dx_2, \quad (5)$$

откуда, дифференцируя по верхнему пределу, получаем $P_Y(y) = \int\limits_{-\infty}^{+\infty} p(x_1, y - x_1)\, dx_1$. Получаем свертку функций с плотностями p_1, p_2 : $P_Y(y) = \int\limits_{-\infty}^{+\infty} p_1(x_1) p_2(y - x_1)\, dx_1$ — в случае независимости X_1 и X_2. ■

2.3. Математическое ожидание случайных величин

2.3.1. Интеграл Лебега по вероятностной мере

Пусть $X(\omega)$ — случайная величина на $(\Omega, \mathscr{A}, P)$, и пусть пока X — дискретна $\implies$ X принимает значения $\{x_1, x_2, \ldots\}$ с вероятностью $p_1, p_2, \ldots$

Пример 3.1. Математическим ожиданием дискретной случайной величины $X(\omega)$ называется число $M\{X\} = E\{X\} := \sum\limits_i x_i p_i$, при условии абсолютной сходимости этого ряда.

Замечание. M — от слова «Mean Value», E — «expectation».

Замечание. В случае $X(\omega) \geqslant 0, (x_j \geqslant 0, \forall j)$, если ряд расходится, то $M\{X\} = +\infty$.

Для сокращения $\sum\limits_i x_i p_i = \sum\limits_x x p_x$, где $p_x := P\{\omega : X(\omega) = x\}$.

Вероятностный смысл $M\{X\}$ — среднее ожидаемое значение случайной величины. Если $n(x)$ — количество испытаний, в которых происходит выигрыш x $\implies$ $\overline{x} = \frac{1}{n} \sum\limits_x x n(x) = \sum\limits_x x (\frac{n(x)}{n}) \approx \sum\limits_x x p_x = M\{X\}$, отметим что $(\frac{n(x)}{n})$ — частота x.

Пример 3.2. $X = \begin{cases} 1, \text{ с вероятностью } p \\ 0, \text{ с вероятностью } 1 - p \end{cases} \implies M\{X\} = 1p + 0(1 - p) = p.$

Пример 3.3. Рассмотрим биномиальное распределение $X(\omega) = m$ с вероятностью $p_m = C_n^m p^m (1-p)^{n-m}, m = 0, 1, \ldots, n$. Имеем $M\{X\} = \sum\limits_{m=0}^{n} m C_n^m p^m (1 - p)^{n-m} = \{$ выкладки в теореме Бернулли $\} = np$.

Пример 3.4. Пуассоновское распределение с показателем λ $X(\omega) = m$, $m = 0, 1, 2, \ldots$,

$$p_m = \frac{\lambda^m}{m!} e^{-\lambda} (\lambda > 0), \quad M\{X\} = \sum_{m=0}^{+\infty} m \frac{\lambda^m}{m!} e^{-\lambda} = \sum_{m=1}^{+\infty} \frac{\lambda^m}{(m-1)!} e^{-\lambda} = \lambda (\sum_{m=0}^{+\infty} \frac{\lambda^m}{m!}) e^{-\lambda} = \lambda.$$

Теорема 2.3. *(Основное свойство $M\{X\}$) Пусть $Y = f(X_1, \ldots, X_n)$, где $X_1, \ldots, X_n$ — дискретные случайные величины. Тогда, в предположении абсолютной сходимости ряда,*

$$M\{Y\} = \sum_{x_1, \ldots, x_m} f(x_1, \ldots, x_m) P\{\omega : X_1(\omega) = x_1, \ldots, X_m(\omega) = x_m\}. \quad (6)$$

□ Докажем для $m = 2$, в общем случае — аналогично. По определению имеем

$$M\{Y\} = \sum_y y P\{\omega : Y(\omega) = y\}, \quad y = f(x_1, x_2),$$

тогда

$$P\{\omega : Y(\omega) = y\} = \sum_{(x_1, x_2) : f(x_1, x_2) = y} P\{\omega : X_1(\omega) = x_1, X_2(\omega) = x_2\},$$

посмотрим, как через это можно записать $M\{Y\}$:

$$\sum_y f(x_1, x_2) \sum_{(x_1, x_2) : f(x_1, x_2) = y} P\{\omega : X_1(\omega) = x_1, X_2(\omega) = x_2\} =$$

$$\sum_{(x_1, \ldots, x_m)} f(x_1, \ldots, x_m) P\{\omega : X_1(\omega) = x_1, \ldots, X_m(\omega) = x_m\}. \quad (7)$$

Поскольку ряд сходится абсолютно, то порядок сумм можно изменять. ■

Замечание. Если $m = 1 \implies Mf(x) = \sum\limits_{X} f(x)P\{\omega : X(\omega) = x\}$, тогда $M\{|X|\} = \sum\limits_{X} |x|P\{\omega : X(\omega) = x\}$, то есть условие абсолютной сходимости ряда для $M\{X\} \leftrightarrow M\{|X|\} < +\infty$.

2.3.2. Свойства математического ожидания

Теорема 2.4. *Пусть $M\{X\}$ существует, тогда:*

1° $X(\omega) \geqslant 0 \implies M\{X\} \geqslant 0$
2° *Если $c = \mathrm{const}$ — неслучайная величина $\implies M\{c\} = c$;*
3° $M\{(cX)\} = cM\{X\}$;
4° $M\{(X_1 + X_2)\} = M\{X_1\} + M\{X_2\}$;
5° *Если $X_1 \leqslant X_2 \implies M\{X_1\} \leqslant M\{X_2\}$. Кроме того, $|M\{X\}| \leqslant M\{|X|\}$;*
6° *Если $X_1, \dots X_m$ независимы $\implies M(X_1 \dots X_m) = MX_1 \cdot \ldots \cdot MX_m$.*

□ **1°** Очевидно.
2° $Mc = c \cdot 1$ (остальные члены ряда — нули).
3° Применим теорему $f(x) = cx : M(cX) = \sum\limits_{x}(cx)P\{\omega : X(\omega) = x\} = cM\{X\}$ (Константу мы просто вынесли из-под знака суммы)
4°

$$
\begin{aligned}
M(X_1 + X_2) &= \sum_{(x_1, x_2)} (x_1 + x_2)P\{X_1(\omega) = x_1, X_2(\omega) = x_2\} = \\
&= \sum_{(x_1, x_2)} x_1 P\{X_1(\omega) = x_1, X_2(\omega) = x_2\} + \sum_{(X_1, X_2)} x_2 P\{X_1(\omega) = x_1, X_2(\omega) = x_2\} = MX_1 + MX_2, \quad (8)
\end{aligned}
$$

так как $\sum\limits_{x_i} P(X_i(\omega) = x_i) = 1$.

5° Следует из того, что $X_2(\omega) - X_1(\omega) \geqslant 0 \implies$ из 1) $M\{X_2 - X_1\} \geqslant 0$, по 3) 4) $MX_1 - MX_2 \geqslant 0$

Второе утверждение: имеем $-|X| \leqslant X \leqslant |X|$, поэтому $-M|X| \leqslant M\{X\} \leqslant M|X| \implies |M\{X\}| \leqslant M|X|$.

6° Рассмотрим функцию $f(x_1 \dots x_m) = x_1 \cdot \ldots \cdot x_m$. Имеем

$$
\begin{aligned}
M(X_1 \dots X_m) &= \sum_{(x_1, \dots x_m)} f(x_1, \dots, x_m) P\{X_1 = x_1, \dots, X_m = x_m\} \overset{!}{=} \\
&\overset{!}{=} \sum_{(X_1, \dots X_m)} (x_1 \cdot \ldots \cdot x_m) P\{X_1 = x_1\} \cdot \ldots \cdot P\{X_m = x_m\} \overset{!!}{=} MX_1 \cdot \ldots \cdot MX_m. \quad (9)
\end{aligned}
$$

Переход, отмеченный «!», верен в силу независимости величин $X_1, \dots X_m$, а «!!» — в силу абсолютной сходимости рядов. ■

Лемма 2.5. *Произвольная случайная величина $X(\omega)$ на σ-алгебре может быть равномерно аппроксимирована последовательностью дискретных случайных величин $X_n(\omega)$.*

□ Покажем, что можно аппроксимировать линейную функцию. Пусть $x_k^{(n)}$ — разбиение $\mathbb{R}$. $\forall\, n = 1, 2, \dots$ пусть $\mathbb{R}$ $x_k^{(n)} \in \mathbb{R} : \{x_k^{(n)} \uparrow\}$ при возрастании k, и $0 < x_{k+1}^{(n)} - x_k^{(n)} \leqslant \varepsilon_n$. Введем случайную функцию:

$$
\varphi^{(n)}(x) \sum_k x_k^{(n)} \chi_{[x_k^{(n)}, x_{k+1}^{(n)}]}, \ k \in \mathbb{Z} \implies
$$

ясно,что $0 < x - \varphi^{(n)}(x) \leqslant \varepsilon_n$, осталось положить $X_n(\omega) := \varphi^{(n)}(X(\omega))$ — это годится для любой случайной величины, т.к. $x - \varphi^{(n)}(x) \leqslant \varepsilon_n$. ■

Лемма 2.6. *Пусть $X_n(\omega)$ — последовательность дискретных случайных величин: $X_n \rightrightarrows X$ и MX_n существует для всех $n \implies$ существует $\lim MX_n$, и этот предел одинаков для $\forall\, \{X_n\}$.*

□ $|MX_n - MX_m| \leqslant M|X_n - X_m| \leqslant M|X_n(\omega) - X(\omega)| + M|X(\omega) - X_m(\omega)| \to 0$ (в силу равномерной сходимости). То что предел единствен легко доказывается. ■

Определение. $M\{X\} := \lim MX_n$, если $X_n \rightrightarrows X$.

Отсюда имеем

$$M\{X\} = \lim_{n\to\infty} M\varphi^{(n)}(X) = \lim_{n\to\infty} \sum_k x_k^{(n)} P\{x_k^{(n)} \leqslant X(\omega) \leqslant x_{k+1}^{(n)}\} =$$

$$= \lim_{n\to\infty} \sum_k x_k^{(n)} \Big[F\big(x_{k+1}^{(n)}\big) - F\big(x_k^{(n)}\big)\Big] =: \int\limits_{-\infty}^{+\infty} x\,dF(x) \text{ — интеграл Лебега–Стилтьеса.} \quad (10)$$

Свойства 1–6 имеют место для произвольных случайных величин, при условие существования математического ожидания. Доказываются с помощью предельного перехода от дискретных случайных величин. Докажем свойство

4)Если $\exists\, M\{X\}, M\{Y\} \implies M\{X\} + M\{Y\} = M\{(X+Y)\}$. В самом деле, найдутся дискретные величины $X_n \rightrightarrows X$, $Y_n \rightrightarrows Y$. Тогда $MX_n \to M\{X\}$; $MY_n \to MY$. Поскольку $X_n + Y_n \rightrightarrows X + Y$, а для дискретных это все выполняется, получаем что $M(X_n + Y_n) = MX_n + MY_n$. Следовательно $\exists \lim\limits_n M(X_n + Y_n) =: M(X+Y)$, но по формуле он равен сумме пределов.

Пусть X — случайная величина с функцией распределения $F(x)$. Тогда

$$M\{X\} = \lim_n \sum_{k=-\infty}^{+\infty} x_k^{(n)} \big[F(x_{k+1}^{(n)}) - F(x_k^{(n)})\big],$$

здесь $|x_k^{(n)} - x_{k+1}^{(n)}| \leqslant \varepsilon_n$, $\varepsilon_n \to 0$, и $x_k^{(n)} \to \pm\infty$ при $k \to \pm\infty$. Тогда $M\{X\} = \int\limits_{-\infty}^{+\infty} x\,dF(x)$.

Задача 2.3. *Если X — дискретная случайная величина, такая что событие x_j происходит с вероятностью p_j, то $\int\limits_{-\infty}^{+\infty} x\,dF(x) = \sum\limits_j p_j x_j$.*

Рассмотрим случай абсолютно непрерывных случайных величин с плотностью $p(x)$. Тогда

$$M\{X\} = \int\limits_{-\infty}^{+\infty} xp(x)\,dx.$$

Теорема 2.7. *Пусть $\int\limits_{+\infty}^{-\infty} |x|p(x)\,dx < \infty$. Тогда $M\{X\} = \int\limits_{-\infty}^{+\infty} xp(x)\,dx$.*

□ Приблизим X равномерно случайными величинами: $\varphi^{(n)} = x_k^{(n)}$, если $x \in [x_k^{(n)}, x_{k+1}^{(n)})$. Докажем, что $M\varphi^{(n)}(x)$ существует и стремится к $\int\limits_{-\infty}^{+\infty} xp(x)\,dx$. В самом деле,

$$M\varphi^{(n)}(x) = \sum_{k=-\infty}^{+\infty} x_k^{(n)} P\{x_k^{(n)} \leqslant X(\omega) < x_{k+1}^{(n)}\}, P\{x_k^{(n)} \leqslant X(\omega) < x_{k+1}^{(n)}\} = \int\limits_{x_k^{(n)}}^{x_{k+1}^{(n)}} p(x)\,dx. \quad (11)$$

Тогда

$$\sum_{k=-\infty}^{+\infty} |x_k^{(n)}| \int\limits_{x_k^{(n)}}^{x_{k+1}^{(n)}} p(x)\,dx \leqslant \sum_{k=-\infty}^{+\infty} \int\limits_{x_k^{(n)}}^{x_{k+1}^{(n)}} (|x| + \varepsilon_n)p(x)\,dx = \int\limits_{-\infty}^{+\infty} (|x| + \varepsilon_n)p(x)\,dx = \int\limits_{-\infty}^{+\infty} |x|p(x)\,dx + \varepsilon_n. \quad (12)$$

Значит сходится. ■

Теперь покажем совпадение математических ожиданий. Имеем

$$M\varphi^{(n)}(x) = \sum_{k=-\infty}^{+\infty} \int\limits_{x_k^{(n)}}^{x_{k+1}^{(n)}} \varphi^{(n)}(x)p(x)\,dx = \int\limits_{-\infty}^{+\infty} \varphi^{(n)}(x)p(x)\,dx, \quad (13)$$

а потому

$$\Big|M\varphi^{(n)}(x) - \int\limits_{-\infty}^{+\infty} xp(x)\,dx\Big| = \int\limits_{-\infty}^{+\infty} (\varphi^{(n)}(x) - x)p(x)\,dx \leqslant \int\limits_{-\infty}^{+\infty} \big|\varphi^{(n)}(x) - x\big|p(x)\,dx \leqslant \varepsilon_n \int\limits_{-\infty}^{+\infty} p(x)\,dx \to 0 (n\to\infty). \quad (14)$$

Пример 3.5. Равномерное распределение: $p(x) = \begin{cases} \frac{1}{b-a}, x \in [a,b] \\ 0, x \notin [a,b] \end{cases}$

$$M\{X\} = \frac{1}{b-a}\int\limits_a^b x dx = \frac{a+b}{2}$$

.

Пример 3.6. Показательное распределение: $p(x) = \begin{cases} \lambda e^{-\lambda x}, x \geqslant 0 \\ 0, x < 0 \end{cases}$, причем $\lambda > 0$.

$$M\{X\} = \int\limits_0^{+\infty} x\lambda e^{-\lambda x}\, dx = \lambda^{-1}.$$

Пример 3.7. $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$.

$$M\{X\} = \int\limits_{-\infty}^{+\infty} \frac{x}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}} = m.$$

Теорема 2.8. *Пусть* $Y = f(x_1, \ldots, x_n)$ f *— измерима.* $X_1 \ldots X_n$ *имеют плотность совместного распределения* $p(x_1 \ldots x_n)$. *Тогда*

$$MY = \int\limits_{-\infty}^{+\infty} \ldots \int\limits_{-\infty}^{+\infty} f(x_1, \ldots, x_n) p(x_1 \ldots x_n)\, dx_1 \ldots x_n \tag{15}$$

(при условной сходимости интеграла).

$$Mf(x) = \int\limits_{-\infty}^{+\infty} f(x)p(x)\, dx.$$

□

$$MY = \lim_{n\to\infty} M\varphi^{(n)}(Y) = \lim_{n\to\infty} \sum_{k=-\infty}^{+\infty} y_k^{(n)} P\{\omega : y_k^{(n)} \leqslant Y(\omega) < y_{k+1}^{(n)}\} =$$

$$= \lim_{n\to\infty} \sum_{k=-\infty}^{+\infty} y_k^{(n)} P\{\omega : y_k^{(n)} \leqslant f(X_1(\omega), \ldots, X_n(\omega)) < y_{k+1}^{(n)}\} =$$

$$\lim_{n\to\infty} \sum_{k=-\infty}^{+\infty} y_k^{(n)} \int\limits_{(x_1,\ldots,x_n):f(x_1,\ldots,x_m)\in[y_k^{(n)}, y_{k+1}^{(n)}]} \ldots\ldots\ldots \int p(x_1, \ldots, x_m)\, dx_1 \ldots\, dx_m =$$

$$= \lim_{n\to\infty} \int\limits_{-\infty}^{+\infty} \ldots \int\limits_{-\infty}^{+\infty} \varphi^{(n)}(f(x_1, \ldots, x_n)) p(x_1, \ldots, x_m)\, dx_1 \ldots\, dx_m = \int\limits_{-\infty}^{+\infty} \ldots \int\limits_{-\infty}^{+\infty} f(x_1, \ldots, x_n) p(x_1, \ldots, x_m)\, dx_1 \ldots\, dx_m. \tag{16}$$

Остается только оценить разность интегралов: это делается абсолютно также:

$$\Big|\int fp - \int \varphi^{(n)} p\Big| \leqslant \int \big|f - \varphi^{(n)}\big| p \leqslant \varepsilon_n \cdot 1.$$

■

2.4. Дисперсия. Неравенство Чебышева. Закон больших чисел

2.4.1. Дисперсия и моменты

Определение. Абсолютным моментом k—ого порядка для случайной величины X называется $M|X|^k, k \in \mathbb{N}$. Если $M|X|^k < \infty \implies k$—й момент равен MX^k

Определение. Центральным k-м моментом называется $M(X - MX)^k$.

Определение. Дисперсией случайной величины X называется $DX = M(X - MX)^2$.

Утверждение 2.9. *Если $MX^2 < \infty$, то существует MX и $DX < \infty$, причем $DX = MX^2 - (MX)^2$.*

□ Имеем $(|x| - 1)^2 \geqslant 0$, поэтому $|x| \leqslant \frac{1}{2} + \frac{1}{2}|x|^2, MX^2 < \infty \implies M|X| < \infty \implies \exists MX \implies DX = M(x^2 - 2XMX + (MX)^2) = MX^2 - 2(MX)^2 + (MX)^2 = MX^2 - (MX)^2$. ■

Пусть X — дискретная случайная величина, принимающая значение x_j с вероятностью p_j. Тогда $D\{X\} = \sum(x_j - M\{x\})^2 p_j$

Пусть X — абсолютно непрерывная случайная величина с плотностью $p(x)$. Тогда

$$D\{X\} = \int\limits_{-\infty}^{+\infty} (x - MX)^2 p(x)\,dx = \int\limits_{-\infty}^{+\infty} x^2 p(x)\,dx - \left(\int\limits_{-\infty}^{+\infty} xp(x)\,dx\right)^2$$

Пример 4.1. $x = \begin{cases} 1 \text{ } \textit{с вероятностью } p \\ 0 \text{ } \textit{с вероятностью } 1 - p \end{cases}$

Тогда $D\{X\} = p(1 - p)$

Пример 4.2. Пуассоновское распределение. $x_j = j$ с вероятностью $p_j = \frac{\lambda^j}{j!}e^{-\lambda}, j = 0, 1, \ldots, \lambda > 0$

Тогда

$$D\{X\} = \sum_{k=0}^{+\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} - \lambda^2 = \lambda.$$

Пример 4.3. Нормальное распределение.

$$D\{X\} = \int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} (x - m)^2 e^{-\frac{(x-m)^2}{2\sigma^2}} = \sigma^2.$$

2.4.2. Неравенство Чебышева

Теорема 2.10 (Неравенство Чебышева). *Пусть X — случайная величина, $M\{X^2\} < \infty$. Тогда для любого $\varepsilon > 0$ выполняется неравенство*

$$P\{\omega : |X(\omega) - M\{X\}| \geqslant \varepsilon\} \leqslant \frac{D\{X\}}{\varepsilon^2}.$$

□ Положим $\varphi(x) := \begin{cases} 0, & |x| < \varepsilon \\ \varepsilon, & |x| \geqslant \varepsilon \end{cases}$. Тогда $\varphi(x)^2 \leqslant x^2$, поэтому

$$\varphi(X(\omega)) = \begin{cases} 0, & \text{с вероятностью } P\{\omega : |X(\omega)| < \varepsilon\} \\ \varepsilon, & \text{с вероятностью } P\{\omega : |X(\omega)| \geqslant \varepsilon\}. \end{cases}$$

Тогда $M\{\varphi(x)^2\} = 0 \cdot P\{\omega : |X(\omega)| < \varepsilon\} + \varepsilon^2 \cdot P\{\omega : |X(\omega)| \geqslant \varepsilon\}$ Так как $\varphi^2(X(\omega)) \leqslant X^2(\omega)$, то

$$M\{\varphi^2(x)\} \leqslant M\{X^2\}, \quad \varepsilon^2 P\{\omega : |X(\omega)| \geqslant \varepsilon\} \leqslant M\{X^2\}.$$

Осталось заменить X на $X - M\{X\}$. ■

Определение. $X(\omega) = Y(\omega)$ с вероятностью 1, если $P\{\omega : X(\omega) = Y(\omega)\} = 1$ (в случае меры Бореля-Лебега $X = Y$ п.в.)

Теорема 2.11. *В предположении конечности дисперсий:*

1. *$DX \geqslant 0$, причём $DX = 0 \Leftrightarrow X(\omega) = MX$ с вероятностью 1;*

2. $D(cX) = c^2 DX$, где c — неслучайная величина;

3. Если $X_1, \ldots, X_n$ — попарно независимые случайные величины, то $D(X_1 + \cdots + X_n) = DX_1 + \cdots + DX_n$.

$\square$

1. $DX \geqslant 0$ — очевидно.
   Пусть $0 = DX = M(X - MX)^2$. По неравенству Чебышёва:
   $$P\left\{\omega : |X(\omega) - MX| \geqslant \frac{1}{n}\right\} = 0 \quad \forall n$$
   $$\{\omega : X(\omega) \neq MX\} = \bigcup_{n=1}^{\infty} \left\{|X - MX| \geqslant \frac{1}{n}\right\}$$
   По свойству непрерывности вероятности:
   $$P\{\omega : X(\omega) \neq MX\} = \lim_{n\to\infty} P\left\{\omega : |X(\omega) - MX| \geqslant \frac{1}{n}\right\} = 0$$
   $$P\{\omega : X(\omega) = MX\} = 1 - P\{X \neq MX\} = 1$$
   $X(\omega) = MX$ почти наверное (с вероятностью 1).

2. очевидно

3. Аддитивность: $|XY| \leqslant \frac{1}{2}(X^2 + Y^2)$, $(X+Y)^2 \leqslant 2(X^2 + Y^2)$. Отсюда $DX, DY < \infty \Rightarrow D(X+Y) < \infty$.

$$\begin{aligned}
D(X_1 + \cdots + X_n) &= M\left(\sum_{i=1}^n X_i - M\sum_{i=1}^n X_i\right)^2 = M\left(\sum_{i=1}^n (X_i - MX_i)\right)^2 = \\
&= M\left(\sum_{i=1}^n (X_i - MX_i)\sum_{j=1}^n (X_j - MX_j)\right) = \sum_{i,j=1}^n M\big((X_i - MX_i)(X_j - MX_j)\big) = \\
&= \sum_{i=1}^n DX_i + \sum_{i\neq j} M\big((X_i - MX_i)(X_j - MX_j)\big) = \sum_{i=1}^n DX_i. \quad (17)
\end{aligned}$$

Последнее равенство верно в силу независимости. На самом деле требуется, чтобы ковариации
$$\operatorname{cov}(X, Y) := M\big((X - MX)(Y - MY)\big)$$
были равны нулю (что верно для независимых случайных величин, но условие независимости не является необходимым).

$\blacksquare$

2.4.3. ЗБЧ (Закон больших чисел)

Теорема 2.12. *Пусть $X_1, \ldots, X_n, \ldots$ — последовательность попарно независимых случайных величин, $M(X_i^2) < \infty$, $m = MX_i$, $\sigma^2 = DX_i$ (математические ожидания и дисперсии одинаковы). Рассмотрим выборочное среднее:*
$$\overline{X}^{(n)} = \frac{1}{n}\sum_{i=1}^n X_i.$$
Тогда
$$\lim_{n\to\infty} P\left\{\left|\overline{X}^{(n)} - m\right| \geqslant \varepsilon\right\} = 0 \qquad (\varepsilon > 0).$$

Говорят, что $\overline{X}^{(n)}$ стремится к m по вероятности.
$\square$
$$M\overline{X}^{(n)} = \frac{1}{n}\sum_{i=1}^n MX_i = m$$

$$D\overline{X}^{(n)} = \frac{1}{n^2}\sum_{i=1}^{n} DX_i = \frac{\sigma^2}{n}$$

В силу неравенства Чебышёва:

$$P\left\{\omega : \left|\overline{X}^{(n)}(\omega) - m\right| \geqslant \varepsilon\right\} \leqslant \frac{\sigma^2}{n\varepsilon^2} \underset{n\to\infty}{\longrightarrow} 0.$$

■

ЗБЧ — широкое обобщение теоремы Бернулли.

Пример 4.4. Схема Бернулли.

$$X_i(\omega) = \begin{cases} 1 & \text{с вероятностью } p \\ 0 & \text{с вероятностью } (1-p) \end{cases} \quad \text{— независимы и одинаково распределены}$$

$$\overline{X}^{(n)}(\omega) = \frac{1}{n}\sum_{i=1}^{n} X_i(\omega) = \nu \quad \text{— выборочная частота успехов}$$

$$m = MX_i = p, \qquad \sigma^2 = DX_i = p(1-p)$$

$$P\left\{\omega : |\nu(\omega) - p| \geqslant \varepsilon\right\} \leqslant \frac{p(1-p)}{n\varepsilon^2} \underset{n\to\infty}{\longrightarrow} 0$$

Точно такую оценку мы получили раньше.

2.4.4. Применение в статистике

$X_1, \ldots, X_n, \ldots$ — независимые с функцией распределения $F(x)$. $F(x)$ неизвестна и её надо оценить, наблюдая X_i. Выборочная функция распределения:

$$F_n(x) = \frac{1}{n}\,(\text{количество тех } X_j,\ j = 1, \ldots, n, \text{ для которых } X_j < x) = \frac{1}{n}N(x)$$

$$p = P\left\{\omega : X(\omega) < x\right\} = F(x)$$

$$\forall x \quad F_n(x) \underset{n\to\infty}{\longrightarrow} F(x) \quad \text{по вероятности}$$

3. Характеристические функции

3.1. Определения и примеры

Определение. Комплексной случайной величиной называется $Z(\omega) = X(\omega) + iY(\omega)$, где X и Y — вещественные случайные величины на $(\Omega, \mathscr{A}, P)$.

$$MZ := MX + iMY \quad (\text{при условии существования } MX \text{ и } MY)$$

Для комплексных случайных величин имеют место обычные свойства мат. ожидания.

Лемма 3.1. $|MZ| \leqslant M|Z|$.

□ Для дискретных комплексных случайных величин: Z принимает значение z_k с вероятностью p_k. $MZ = \sum\limits_k z_k p_k$. В силу обобщённого неравенства треугольника

$$|MZ| = \left|\sum_k z_k p_k\right| \leqslant \sum_k |z_k| p_k = M|Z|.$$

Произвольную комплексную случайную величину можно равномерно аппроксимировать дискретными. ■

Определение. Характеристической функцией вещественной случайной величины X называется

$$\varphi(t) = Me^{itX}, \quad t \in \mathbb{R}.$$

$$e^{i\alpha} = \cos\alpha + i\sin\alpha, \quad \alpha \in \mathbb{R} \quad (\text{формула Эйлера})$$

Отсюда

$$\left|e^{i\alpha}\right| = \sqrt{\cos^2\alpha + \sin^2\alpha} = 1$$

Следовательно, $\varphi(t) = Me^{itX}$ существует при всех $t \in \mathbb{R}$.

$$\varphi(t) = M(\cos tX) + iM(\sin tX)$$

Пусть X имеет функцию распределения F. Тогда для любой борелевской функции f такой, что $Mf(X)$ существует:

$$Mf(X) = \int\limits_{-\infty}^{+\infty} f(x)\,dF(x) \quad \text{— интеграл Лебега–Стилтьеса}$$

$$\varphi(t) = \int\limits_{-\infty}^{+\infty} e^{itx}\,dF(x) = \int\limits_{-\infty}^{+\infty} \cos tx\,dF(x) + i\int\limits_{-\infty}^{+\infty} \sin tx\,dF(x)$$

X — дискретная случайная величина (принимает значение x_k с вероятностью p_k).

$$\varphi(t) = \sum_k e^{itx}p_k = \sum_k (\cos tx_k)p_k + i\sum_k (\sin tx_k)p_k$$

— суперпозиция комплексных гармоник.

X имеет абсолютно непрерывное распределение с плотностью $p(x)$. $dF(x) = p(x)\,dx$

$$\varphi(t) = \int\limits_{-\infty}^{+\infty} e^{itx}p(x)\,dx =$$

$$= \int\limits_{-\infty}^{+\infty} (\cos tx)p(x)\,dx + i\int\limits_{-\infty}^{+\infty} (\sin tx)p(x)\,dx$$

— комплексное преобразование Фурье для $p(x)$. Преобразование Фурье обратимо: по $\varphi(t)$ можно восстановить $p(x)$. $\varphi(t)$ — равнозначный способ задания случайной величины.

Пример 1.1. $x = \begin{cases} 1 & \text{с вероятностью } p \\ 0 & \text{с вероятностью } (1-p) \end{cases}$ — случайный бит. $\varphi(t) = (1-p) + e^{it}\cdot p$.

Пример 1.2. Пуассоновское распределение. $X : n = 0, 1, 2, \ldots$ с вероятностью $p_k = \frac{\lambda^n e^{-\lambda}}{n!}$.

$$\varphi(t) = \sum_{n=0}^{\infty} e^{itn}\frac{\lambda^n e^{-\lambda}}{n!} = e^{-\lambda}\sum_{n=0}^{\infty}\frac{(\lambda e^{it})^n}{n!} = e^{\lambda(e^{it}-1)}$$

Пример 1.3. Стандартное нормальное распределение.

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

$$\varphi(t) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} e^{itx}e^{-\frac{x^2}{2}}\,dx = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} \cos tx\,e^{-\frac{x^2}{2}}\,dx \quad \text{(для sin получится ноль)}$$

$$\varphi'(t) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}(-x\sin tx)e^{-\frac{x^2}{2}}\,dx = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}\sin tx\,d\Big(e^{-\frac{x^2}{2}}\Big) = \underbrace{\frac{1}{\sqrt{2\pi}}\sin tx\,e^{-\frac{x^2}{2}}\Big|_{-\infty}^{+\infty}}_{0} - t\,\underbrace{\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\cos tx\,dx}_{\varphi(t)}$$

$$\varphi'(t) = -t\varphi(t) \quad \text{— линейное дифференциальное уравнение}$$

$$\varphi(t) = Ce^{-\frac{t^2}{2}}; \quad \varphi(0) = \int\limits_{-\infty}^{+\infty} p(x)\,dx = 1 \Rightarrow C = 1 \Rightarrow \varphi(t) = e^{-\frac{t^2}{2}}$$

$$p(x) = \frac{1}{\sqrt{2\pi}}\varphi(x).$$

3.2. Свойства характеристических функций

Теорема 3.2 (Основные свойства характеристических функций).

1° $\varphi(0) = 1$, $|\varphi(t)| \leqslant 1$ $\forall t \in \mathbb{R}$ *(нормировка);*

2° $\varphi(t)$ *является положительно определённой функцией, то есть*

$$\sum_{k,j=1}^{n} \overline{c_k} c_j \varphi(t_j - t_k) \geqslant 0$$

для всех $n \in \mathbb{N}$ *и любых наборов* $\{t_1, \dots, t_n\} \subset \mathbb{R}$, $\{c_1, \dots, c_n\} \subset \mathbb{C}$;

3° $\varphi(t)$ *равномерно непрерывна по* $t \in \mathbb{R}$;

4° $Y(\omega) = aX(\omega) + b$ *(линейное преобразование)* $\Rightarrow \varphi_Y(t) = e^{itb}\varphi_X(at)$;

5° *Если* $Y(\omega) = X_1(\omega) + \dots + X_n(\omega)$, $X_1, \dots, X_n$ *независимы, то* $\varphi_Y(t) = \varphi_{X_1}(t) \cdot \ldots \cdot \varphi_{X_n}(t)$.

□

1. $\varphi(t) = Me^{itX}$. $\varphi(0) = Me^{i\cdot 0\cdot X} = M1 = 1$. $\forall t \quad |\varphi(t)| = \left|Me^{itX}\right| \leqslant M\left|e^{itX}\right| = M1 = 1$.

2.

$$0 \leqslant M\left|\sum_{j=1}^{n} c_j e^{it_jX}\right|^2 = M\overline{\left(\sum_{k=1}^{n} c_k e^{it_kX}\right)}\left(\sum_{j=1}^{n} c_j e^{it_jX}\right) = \sum_{k,j=1}^{n} \overline{c_k} c_j M\underbrace{e^{-it_kX + it_jX}}_{e^{i(t_j - t_k)X}} = \sum_{k,j=1}^{n} \overline{c_k} c_j \varphi(t_j - t_k)$$

3. **Лемма 3.3.** $\left|e^{i\alpha} - 1\right| = 2\left|\sin\frac{\alpha}{2}\right|$, $\alpha \in \mathbb{R}$.

□

$$\left|e^{i\alpha} - 1\right| = \sqrt{(\cos\alpha - 1)^2 + \sin^2\alpha} = \sqrt{1 - 2\cos\alpha + \cos^2\alpha + \sin^2\alpha} = \sqrt{2(1 - \cos\alpha)} = \sqrt{4\sin^2\frac{\alpha}{2}} = 2\left|\sin\frac{\alpha}{2}\right|$$

■

$$|\varphi(t) - \varphi(s)| = \left|Me^{itX} - Me^{isX}\right| = \left|Me^{itX}\left(1 - e^{i(s-t)X}\right)\right| \leqslant M\left|e^{itX}\left(1 - e^{i(s-t)X}\right)\right| = M\left|e^{i(s-t)X} - 1\right| =$$

$$= \{\text{в силу леммы}\} = 2M\left|\sin\frac{\overbrace{(s-t)}^{\Delta}X}{2}\right| \xrightarrow{(?)} 0 \ (\Delta \to 0)$$

Обоснуем предельный переход (?). $f_\Delta(\omega) = \left|\sin\frac{\Delta\cdot X(\omega)}{2}\right|$. $|f_\Delta(\omega)| \leqslant 1$ — равномерное ограничение.
$\forall$ фикс. $\omega \in \Omega \quad f_\Delta(\omega) \to 0$ $(\Delta \to 0)$ — мажорированная сходимость.

3.3. Теорема Лебега и ее трагические последствия

Теорема 3.4 (Лебега о мажорированной сходимости). *Если* $X_n(\omega) \to X(\omega)$ *для всех* ω, *причём*

$$|X_n(\omega) - X(\omega)| \leqslant g(\omega),$$

где $Mg(\omega) < \infty$, *то* $MX_n \to MX$.

(Это теорема из курса действительного анализа)

Поэтому $|\varphi(t) - \varphi(s)| \to 0$ $(\Delta \to 0)$.

4. $\varphi_Y(t) = Me^{itY} = Me^{it(aX+b)} = M\big(e^{itaX}\overbrace{e^{itb}}^{\text{const}}\big) = e^{itb}Me^{i(ta)X} = e^{itb}\varphi_X(at)$.

Пример 3.1. Характеристическая функция $\mathscr{N}(m, \sigma^2)$.
$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-m)^2}{2\sigma^2}}$.
$X \sim \mathscr{N}(0,1) \Rightarrow Y(\omega) = \sigma X(\omega) + m \sim \mathscr{N}(m, \sigma^2)$.
$\varphi_X(t) = e^{-\frac{t^2}{2}} \Rightarrow \varphi_Y(t) = e^{itm - \frac{\sigma^2t^2}{2}}$.
При $\sigma^2 = 0$: $\varphi_Y(t) = e^{itm}$ — характеристическая функция неслучайной величины $X(\omega) \equiv m$ (а плотность при $\sigma \to 0$ не существует).

5. $X_1, \dots, X_n$ — независимые случайные величины $\Rightarrow e^{itX_1}, \dots, e^{itX_n}$ — тоже независимые.

$$\varphi_Y(t) = Me^{it(X_1+\dots+X_n)} = M\big(e^{itX_1}\dots e^{itX_n}\big) = \big(Me^{itX_1}\big)\dots\big(Me^{itX_n}\big) = \varphi_{X_1}(t)\dots\varphi_{X_n}(t)$$

■

Определение. Моментом n-го порядка случайной величины X называется MX^n (если существует). Абсолютным моментом n-го порядка называется $M|X|^n$ (может быть $=\infty$).

Теорема 3.5. *Пусть $M|X|^m < \infty$. Тогда $\varphi(t)$ m раз дифференцируема при всех $t \in \mathbb{R}$, причём $\varphi^{(k)}(0) = i^k MX^k$, $k \in \{1, \dots, m\}$. Имеет место разложение в ряд Тейлора:*

$$\varphi(t) = \sum_{k=0}^{m} \frac{(it)^k}{k!} MX^k + o(t^m) \text{ при } t \to 0.$$

□ $\varphi^{(k)}(t) = i^k Me^{itX}(x)^k, \quad k \in \{0, \dots, m\}$ — получается формальным дифференцированием под знаком математического ожидания. Нужно обосновать законность такого дифференцирования.

$$|X|^k \leqslant |X|^m + 1 \quad (k \in \{1, \dots, m\}) \quad \Rightarrow \quad M|X|^k < \infty$$

Индукция по k. При $k = 0$ формула тривиальным образом верна. Пусть верно для $k < m$. Докажем для $k+1$.

$$\left|\frac{\varphi^{(k)}(t+\Delta) - \varphi^{(k)}(t)}{\Delta} - i^{k+1}Me^{itX}(X)^{k+1}\right| = \{\text{предп. инд.}\} = \left|\frac{i^k M(X^k)\left(e^{i(t+\Delta)X} - e^{itX}\right)}{\Delta} - i^{k+1}MX^{k+1}e^{itX}\right| =$$

$$= \left|MX^k\left(\left(\frac{e^{i(t+\Delta)X} - e^{itX}}{\Delta}\right) - iXe^{itX}\right)\right| \leqslant M|X|^k \underbrace{\left|\frac{e^{i\Delta X} - 1}{\Delta} - iX\right|}_{Y_\Delta}$$

Нужно показать, что $|Y_\Delta(\omega) \leqslant C|X|$, $Y_\Delta(\omega) \to 0\ \forall\omega$. Тогда мы сможем воспользоваться теоремой о мажорированной сходимости. (Тогда $\big||X|^kY\big| \leqslant c|X|^{k+1}$.) $Y_\Delta \to 0$ поточечно (т.к. $\left(e^{i\Delta X}\right)'_{\Delta=0} = iX$).

$$\left|\frac{e^{i\Delta X} - 1}{\Delta} - iX\right| \leqslant \left|\frac{e^{i\Delta X} - 1}{\Delta}\right| + |X| = \frac{2\left|\sin\frac{\Delta X}{2}\right|}{|\Delta|} + |X| \leqslant \frac{2\left|\frac{\Delta X}{2}\right|}{|\Delta|} + |X| = 2|X|$$

■

3.4. Формула обращения

Теорема 3.6 (Формула обращения). *Пусть $\varphi(t)$ — характеристическая функция случайной величины X с функцией распределения $F(x)$. Тогда для любых точек непрерывности $x_1 < x_2$ функция распределения $F(x)$ удовлетворяет условию:*

$$F(x_2) - F(x_1) = \lim_{\sigma\to 0} \frac{1}{2\pi} \underbrace{\int_{-\infty}^{+\infty} \varphi(t) e^{-\frac{\sigma^2t^2}{2}} \frac{e^{-itx_2} - e^{-itx_1}}{-it}\, dt}_{(*)}$$

Докажем эту теорему мы несколько позже, а пока займёмся следствиями.

Лемма 3.7 (Следствие 1). *Характеристическая функция однозначно определяет функцию распределения $F(x)$. $\quad \varphi(t) \leftrightarrow F(x)$.*

Лемма 3.8 (Следствие 2). *Пусть характеристическая функция $\varphi(t)$ интегрируема (по Лебегу) на $\mathbb{R}$. Тогда $F(x)$ абсолютно непрерывна, и распределение имеет плотность*

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \varphi(t)\, dt.$$

(формула обратного преобразования Фурье)

□ **(Доказательство следствия 2)**

$$\varphi(t)e^{-\frac{\sigma^2t^2}{2}}\frac{e^{-itx_2}-e^{-itx_1}}{-it}\underset{\sigma\to 0}{\longrightarrow}\varphi(t)\frac{e^{-itx_2}-e^{-itx_1}}{-it}\quad\forall t\in\mathbb{R}\quad\text{(поточечная сходимость)}$$

$$\left|\varphi(t)e^{-\frac{\sigma^2t^2}{2}}\frac{e^{-itx_2}-e^{-itx_1}}{-it}\right|=|\varphi(t)|\left|e^{it(x_1-x_2)}-1\right|\frac{e^{-\frac{\sigma^2t^2}{2}}}{|t|}\leqslant|\varphi(t)|\frac{2|\sin(t(x_1-x_2)/2)|}{|t|}\leqslant\underbrace{|\varphi(t)|\cdot|x_1-x_2|}_{\text{интегрируемая мажоранта}}$$

По теореме Лебега можно перейти к пределу под знаком интеграла:

$$F(x_2)-F(x_1)=\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}\varphi(t)\frac{e^{-itx_2}-e^{-itx_1}}{it}\,dt=\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}\varphi(t)\left(\int\limits_{x_1}^{x_2}e^{-ity}\,dy\right)dt$$

По теореме Фубини меняем порядок интегрирования:

$$F(x_2)-F(x_1)=\dots=\int\limits_{x_1}^{x_2}\left(\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}e^{-ity}\varphi(t)\,dt\right)dy,$$

откуда F абсолютно непрерывна, и $p(y)=\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}e^{-itx}\varphi(t)\,dt$.

■

□ **(Доказательство формулы обращения)** Пусть X — случайная величина с функцией распределения $F(x)$ и характеристической функцией $\varphi(x)$. Рассмотрим сглаженную случайную величину $X+Y_\sigma$, где $Y_\sigma\sim\mathcal{N}(0,\sigma^2)$, X и Y независимы. Характеристическая функция $X+Y_\sigma$ есть $\varphi(t)e^{-\frac{\sigma^2t^2}{2}}=\varphi_\sigma(t)$. Функция $\varphi(t)$ превращается в интегрируемую функцию $\varphi_\sigma(t)$.

$p_\sigma(x)$ — плотность распределения $X+Y_\sigma$. Докажем формулу обращения для плотности:

$$p_\sigma(x)=\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}e^{-itx}\varphi_\sigma(t)\,dt$$

Далее получим:

$$F_\sigma(x_2)-F_\sigma(x_1)=\int\limits_{x_1}^{x_2}p_\sigma(t)\,dt=\frac{1}{2\pi}\underbrace{\int\limits_{-\infty}^{+\infty}\dots}_{(*)}$$

В силу обобщения формулы свёртки:

$$p_\sigma(x)=\int\limits_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(y-x)^2}{2\sigma^2}}\,dF(y)$$

С другой стороны,

$$\frac{1}{2\pi}\underbrace{\int\limits_{-\infty}^{+\infty}e^{-itx}\varphi_\sigma(t)\,dt}_{\text{существует (инт. Лебега)}}=\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}e^{-itx}e^{-\frac{\sigma^2t^2}{2}}\underbrace{\int\limits_{-\infty}^{+\infty}e^{ity}\,dF(y)}_{\varphi(t)}\,dt=\{\text{теорема Фубини}\}=$$

$$=\int\limits_{-\infty}^{+\infty}\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}e^{it(y-x)-\frac{\sigma^2t^2}{2}}\,dt\,dF(y)=\int\limits_{-\infty}^{+\infty}\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}e^{it(y-x)}\underbrace{\frac{1}{\sqrt{\frac{2\pi}{\sigma^2}}}e^{-\frac{\sigma^2t^2}{2}}}_{\text{плотность }\mathcal{N}\left(0,\frac{1}{\sigma^2}\right)}\,dt\,dF(y)=$$

$$=\int\limits_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(y-x)^2}{2\sigma^2}}\,dF(y)=p_\sigma(x)$$

Формула обращения для сглаженной случайной величины доказана:

$$F_\sigma(x_2) - F_\sigma(x_1) = \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} \frac{e^{-itx_2} - e^{-itx_1}}{-it} e^{-\frac{\sigma^2 t^2}{2}} \varphi(t)\, dt.$$

Осталось показать, что $\lim\limits_{\sigma\to 0} F_\sigma(x) = F(x)$ в точках непрерывности. Зафиксируем $\varepsilon > 0$.

$$F_\sigma(x) = P\{X + Y_\sigma < x\} = P\{X < x - Y_\sigma\} \qquad \forall x.$$

$$F_\sigma(x) \leqslant P\{X < x - Y_\sigma, |Y_\sigma| \leqslant \varepsilon\} + P\{|Y_\sigma| > \varepsilon\} \leqslant \{\text{по неравенству Чебышёва}\} \leqslant P\{X < x + \varepsilon\} + \frac{\sigma^2}{\varepsilon^2} = \\ = \{\text{выберем } \sigma^2 = \varepsilon^3\} = F(x+\varepsilon) + \varepsilon.$$

$$F_\sigma(x) \geqslant P\{X < x - Y_\sigma, |Y_\sigma| \leqslant \varepsilon\} \geqslant P\{X < x - \varepsilon\} - P\{|Y_\sigma| > \varepsilon\} \geqslant \{\text{по неравенству Чебышёва}\} \geqslant F(x-\varepsilon) - \frac{\sigma^2}{\varepsilon^2} = \\ = F(x-\varepsilon) - \varepsilon.$$

Имеем: $F(x-\varepsilon) - \varepsilon \leqslant F_\sigma(x) \leqslant F(x+\varepsilon) + \varepsilon$, $F(x-\varepsilon) - \varepsilon \to F(x)$, $F(x+\varepsilon) + \varepsilon \to F(x)$ (т.к. x — точка непрерывности F). Отсюда $F_\sigma(x) \to F(x)$. ■

4. Наиболее суровые вопросы теории вероятностей

4.1. Теоремы Хелли

Определение. Последовательность $\{F_n(x)\}$ функций распределения слабо сходится к $F(x)$, если $\forall\, x \in \mathbf{C}(F)$ имеем $\lim F_n(x) = F(x)$. Тогда можно считать, что $0 \leqslant F \leqslant 1$, если разрешить переопределять F в точках разрыва, чтобы F была непрерывна слева. Обозначение: $F_n \Rightarrow F$.

Теорема 4.1 (Хелли-I). *Из последовательности функций распределения можно выбрать слабо сходящуюся.*

□ Рассмотрим счётное всюду плотное $D \subset \mathbb{R}$ и занумеруем его элементы: $D = \left\{x^1, x^2, x^3, \dots\right\}$. Рассмотрим $\left\{F_n(x^1)\right\}$ и выберем (в силу её ограниченности) сходящуюся подпоследовательность $\left\{F_{1n}(x^1)\right\}$. Далее рассмотрим $\left\{F_{1n}(x^2)\right\}$ и в ней выберем сходящуюся $\left\{F_{2n}(x^2)\right\}$ и так далее. Таким образом, количество точек сходимости на n-м шаге увеличивается на 1. Очевидно, что последовательность $\{F_{nn}\}$ сходится на D к пределу (обозначим его F), поскольку $\forall\, k$ начиная с $n = k$ последовательность $\left\{F_{nn}(x^k)\right\}$ является подпоследовательностью сходящейся. Далее, покажем, что если $x \in \mathbf{C}(F)$, то $F_{nn}(x) \to F(x)$. Пусть $x', x'' \in D$, причём $x' < x < x''$, тогда $F(x') \leqslant F_{nn}(x) \leqslant F(x'')$. Осталось устремить $x' \to x \leftarrow x''$, тогда по теореме о двух милиционерах имеем $F_{nn}(x) \to F(x)$, ибо это число зажато между $F(x')$ и $F(x'')$. ■

Теорема 4.2 (Хелли-II). *Пусть $F_n \Rightarrow F$, где F — функция распределения. Тогда $\int\limits_{\mathbb{R}} f\, dF_n \to \int\limits_{\mathbb{R}} f\, dF$, если $f \in \mathbf{C}$ и ограничена на $\mathbb{R}$.*

□ Пусть $|f| \leqslant C$. Фиксируем $\varepsilon > 0$, пусть $a < b \in \mathbb{R}$ — достаточно далёкие точки непрерывности F, т. е. $F(a) < \varepsilon$, а $F(b) > 1 - \varepsilon$. Тогда, поскольку $F_n \Rightarrow F$, $\exists\, N\colon \forall\, n > N$ имеем $F_n(a) < \varepsilon$ и $F_n(b) > 1 - \varepsilon$. Тогда

$$\left|\int\limits_{-\infty}^{+\infty} f\, dF_n - \int\limits_a^b f\, dF_n\right| \leqslant \int\limits_{-\infty}^{a} |f|\, dF_n + \int\limits_b^{+\infty} |f|\, dF_n \leqslant 2C\varepsilon, \tag{1}$$

и

$$\left|\int\limits_{-\infty}^{+\infty} f\, dF - \int\limits_a^b f\, dF\right| \leqslant \int\limits_{-\infty}^{a} |f|\, dF + \int\limits_b^{+\infty} |f|\, dF \leqslant 2C\varepsilon. \tag{2}$$

Осталось показать, что $\int\limits_a^b f\, dF_n \to \int\limits_a^b f\, dF$. Разобьём отрезок $[a, b]$ точками непрерывности так мелко (N штук), что $\omega(f) < \varepsilon$ на каждом элементе разбиения. Так можно сделать в силу равномерной непрерывности f

на $[a,b]$. Приблизим f равномерно ступенчатой функцией g, для которой $\|f-g\|_{[a,b]} \leqslant \varepsilon$, тогда

$$\left|\int\limits_a^b f\,dF_n - \int\limits_a^b f\,dF\right| \leqslant \int\limits_a^b |f-g|\,dF_n + \left|\int\limits_a^b g\,dF_n - \int\limits_a^b g\,dF\right| + \int\limits_a^b |f-g|\,dF \leqslant$$

$$\leqslant \varepsilon + C\left(\sum_{k=1}^N \bigl(F_n(x_k) - F(x_k) - F_n(x_{k-1}) + F(x_{k-1})\bigr)\right) + \varepsilon, \quad (3)$$

а последнее слагаемое можно сделать маленьким при $n \to \infty$. ■

4.2. Предельные теоремы

Теорема 4.3 (Прямая предельная теорема). *Пусть F_n, F — функции распределения, а φ_n, φ — их характеристические функции. Пусть $F_n \Rightarrow F$, тогда $\varphi_n \to \varphi$.*

□ Применим вторую теорему Хелли к $f = e^{itx}$, тогда $\varphi_n(t) = \int e^{itx}\,dF_n \to \int e^{itx}\,dF = \varphi(t)$. ■

Лемма 4.4 (Оценка вероятности хвостов). $\mathsf{P}\bigl\{|X| \geqslant \frac{2}{u}\bigr\} \leqslant \frac{1}{u}\int\limits_{-u}^{u}\bigl(1-\varphi(t)\bigr)\,dt$, *где* $\varphi = \operatorname{char} X$.

□ Имеем

$$\frac{1}{2u}\int\limits_{-u}^{u}\bigl(1-\varphi(t)\bigr)\,dt = \frac{1}{2u}\int\limits_{-u}^{u}\left(1-\mathsf{M}e^{-itX}\right)\,dt = \bigl\{\text{Фубини}\bigr\} =$$

$$= 1 - \frac{1}{2u}\mathsf{M}\int\limits_{-u}^{u} e^{itX}\,dt = 1 - \mathsf{M}\frac{e^{iuX}-e^{-iuX}}{2iuX} = 1 - M\frac{\sin uX}{uX} = (*). \quad (4)$$

Имеем $1 - \frac{\sin x}{x} \geqslant g(x) := \frac{1}{2} - \frac{1}{2}\chi_{(-2,2)}$, поэтому

$$(*) \geqslant \mathsf{M}g(uX) = 0\cdot\mathsf{P}\bigl\{|uX|<2\bigr\} + \frac{1}{2}\cdot\mathsf{P}\bigl\{|uX|\geqslant 2\bigr\},$$

что и требуется. ■

Теорема 4.5 (Обратная предельная теорема). *Если $\{\varphi_n = \operatorname{char} F_n\}$ сходится к φ, непрерывной при $t=0$, то $F_n \Rightarrow F$, где F — функция распределения, для которой $\operatorname{char} F = \varphi$.*

□ По первой теореме Хелли выделим $F_{n_k} \Rightarrow F$. Беда в том, что она в общем случае не будет функцией распределения. Докажем, что $F(-\infty) = 0$, а $F(+\infty) = 1$. Фиксируем $\varepsilon > 0$, тогда, поскольку $\varphi \in \mathbf{C}(0)$, а $\varphi(0) = \lim\varphi_n(0)$, то $\exists\, U_u(0)\colon |1-\varphi(t)| \leqslant \varepsilon$. Тогда $\frac{1}{u}\int\limits_{-u}^{u}\bigl(1-\varphi(t)\bigr)\,dt < \varepsilon$, поэтому для достаточно больших n имеем $\frac{1}{u}\int\limits_{-u}^{u}\bigl(1-\varphi_n(t)\bigr)\,dt < \varepsilon$, ибо по теореме Лебега интегралы сходятся к интегралу предела.

Применим лемму 4.4, тогда $\mathsf{P}\bigl\{|X_n| \geqslant \frac{2}{u}\bigr\} < \varepsilon$. Имеем $\mathsf{P}\bigl\{X_n \leqslant -\frac{2}{u}\bigr\} + \mathsf{P}\bigl\{X_n \geqslant \frac{2}{u}\bigr\} = F_n\bigl(-\frac{2}{u}\bigr) + 1 - F_n\bigl(\frac{2}{u}\bigr)$. Выберем u так, чтобы $\pm\frac{2}{u}$ были точками непрерывности F, и устремим $n\to\infty$. Отсюда $\varepsilon \geqslant F\bigl(-\frac{2}{u}\bigr) + 1 - F\bigl(\frac{2}{u}\bigr) \geqslant F(-\infty) + 1 - F(+\infty)$.

По второй теореме Хелли имеем $\int\limits_{\mathbb{R}} e^{itx}\,dF_n \to \int\limits_{\mathbb{R}} e^{itx}\,dF$, откуда $\varphi_{n_k} \to \varphi = \operatorname{char} F$. Если теперь допустить, что $F_n \not\Rightarrow F$, то найдутся две подпоследовательности $F_{n'} \to F^*$ и $F_{n''} \Rightarrow F^{**}$, но по прямой теореме $\varphi_{n'} \to \varphi^*$ и $\varphi_{n''} \to \varphi^{**}$, но их предел общий, поэтому $\varphi^* = \varphi^{**} = \varphi$, но это противоречит формуле обращения. ■

Теорема 4.6 (ЦПТ). *Пусть $X_1, \ldots, X_n$ — независимые одинаково распределённые случайные величины, для которых $\mathsf{M}X_i^2 < \infty$. Положим $m := \mathsf{M}X_i$, а $\sigma^2 := \mathsf{D}X_i > 0$. Пусть F_n — функции распределения величин $S_n := \frac{1}{\sqrt{n}}\sum\limits_{i=1}^{n}\frac{X_i - m}{\sigma}$. Тогда*

$$\lim F_n = \Phi(x) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{y^2}{2}}\,dy. \quad (5)$$

□ Рассмотрим величины $\frac{X_i-m}{\sigma}$, тогда $\mathsf{M}\frac{X_i-m}{\sigma} = 0$, а $\mathsf{D}\frac{X_i-m}{\sigma} = 1$. Пусть им соответствует характеристическая функция $\varphi(t)$, тогда $\varphi \in \mathbf{D}^2$, кроме того, $\varphi(0) = 1$, $\varphi'(0) = 0$, $\varphi''(0) = i^2 = -1$. По формуле Тейлора имеем

$\varphi(t) = 1 - \frac{t^2}{2} + o\left(t^2\right)$, тогда по мультипликативному свойству имеем $\operatorname{char} S_n = \varphi^n\left(\frac{t}{\sqrt{n}}\right)$. При фиксированном t имеем $\varphi^n\left(\frac{t}{\sqrt{n}}\right) = \left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n \to e^{-\frac{t^2}{2}} = \operatorname{char}\mathcal{N}(0,1)$. Осталось применить теорему непрерывности. ■

Применения ЦПТ — статистика. Пусть имеются одинаково ошибочные наблюдения с двухсторонними погрешностями, т. е. $X_i = m + Y_i$, где Y_i — ошибки, а m — истинное значение. Предположение состоит в том, что $\mathsf{M}Y_i = 0$, а $MY_i^2 < \infty$, т. е. разброс конечен. Введём $S_n := \frac{1}{n}\sum\limits_{i=1}^{n} X_i$, тогда рассмотрим $\mathsf{P}\big\{\omega\colon |S_n - m| \geqslant \varepsilon\big\} = \mathsf{P}\Big\{\frac{1}{n}\sum\limits_{i=1}^{n}(X_i - m) < \frac{x\sigma}{\sqrt{n}}\Big\} \to \Phi(x)$. Отсюда $\mathsf{P}\Big\{\omega\colon \big|\frac{1}{n}\sum\limits_{i=1}^{n}(X_i - m)\big| \geqslant \frac{x\sigma}{\sqrt{n}}\Big\} \to 1 - \Phi(x) + \Phi(-x) = 1 - 2\Phi_0(x)$, где $\Phi_0(x) = \frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{y^2}{2}}\,dy$. Примерные данные показывают следующее:

$x \backslash \mathsf{P}_x$	ЗБЧ	ЦПТ
1	$\leqslant 1$	$\leqslant 0.317$
3	$\leqslant 0.1$	$\leqslant 0.0026$
10	$\leqslant 0.01$	~ 0

4.3. Теорема Ляпунова

Теорема 4.7 (Ляпунова). *Пусть $X_1, \ldots, X_n$ — независимые случайные величины. Пусть $\mathsf{M}|X_i|^3 < \infty$ и*

$$\frac{\sqrt[3]{\sum \mathsf{M}|X_i - \mathsf{M}X_i|^3}}{\sqrt{\sum \mathsf{M}|X_i - \mathsf{M}X_i|^2}} \to 0 \quad \Rightarrow \quad \mathsf{P}\left\{\frac{\sum(X_i - \mathsf{M}X_i)}{\sqrt{\sum \mathsf{M}|X_i - \mathsf{M}X_i|^2}} < x\right\} \to \Phi(x).$$

4.4. Закон 0 и 1 Колмогорова

Рассмотрим вероятностное пространство $(\Omega, \mathcal{A}, \mathsf{P})$ и на нём — $X_1, \ldots, X_n$ — случайные величины. Мы доказывали, что если $B \in \mathcal{B}(\mathbb{R}^n)$, то $\big\{\omega\colon \big(X_1(\omega), \ldots, X_n(\omega)\big) \in B\big\} \in \mathcal{A}$. Более того, если рассмотреть все $B \in \mathcal{B}$, то совокупность всех прообразов образуют σ-алгебру, поскольку полный прообраз уважает необходимые теоретико-множественные операции. Эта σ-алгебра называется порождённой $X_1, \ldots, X_n$ и обозначается $\mathcal{A}(X_1, \ldots, X_n)$, кроме того, ясно, что она содержится в $\mathcal{A}$.

Рассмотрим бесконечную последовательность $\{X_n\}$. Рассмотрим $\mathcal{A}(X_{n+1}, \ldots, X_m)$ и зафиксируем n. Рассмотрим $\mathcal{A}_n := \bigcup\limits_{m=n+1}^{\infty} \mathcal{A}(X_{n+1}, \ldots, X_m)$. Очевидно, что это алгебра, но не σ-алгебра, так что её надо расширить: возьмём $\mathcal{A}(X_{n+1}, \ldots) := \sigma(\mathcal{A}_n)$. Теперь пересечём их всех, и получим искомую *σ-алгебру остаточных событий*: $\mathcal{A}_\infty = \bigcap\limits_{n=0}^{\infty} \mathcal{A}(X_{n+1}, \ldots)$. Ясно, что при пересечении ничего не испортится. Каждое из событий $A \in \mathcal{A}_\infty$ называется *остаточным*, или *хвостовым*.

Определение. Пусть дана последовательность случайных величин $\{X_i\}$. Они называются независимыми, если $\forall\, n$ имеем $X_1, \ldots, X_n$ независимы.

Определение. Назовём σ-алгебры $\mathcal{A}_1$ и $\mathcal{A}_2$ независимыми, если $\forall\, A_1 \in \mathcal{A}_1$, $\forall\, A_2 \in \mathcal{A}_2$ имеем $\mathsf{P}(A_1 A_2) = \mathsf{P}(A_1)\mathsf{P}(A_2)$.

Теорема 4.8 (Закон «0 или 1» Колмогорова). *Пусть $X_1, \ldots$ — последовательность случайных величин, а $\mathcal{A}_\infty$ — порождённая ею σ-алгебра остаточных событий. Пусть $A \in \mathcal{A}_\infty$, тогда либо $\mathsf{P}(A) = 0$, либо $\mathsf{P}(A) = 1$.*

□ Пусть $A \in \mathcal{A}_\infty$, тогда $A \in \mathcal{A}(X_{n+1}, \ldots)$ для всякого $n \in \mathbb{N}_0$. Фиксируем n и пусть $A_1 \in \mathcal{A}(X_1, \ldots, X_n)$. Имеем $\mathcal{A}(X_1, \ldots, X_n)$ и $\mathcal{A}(X_{n+1}, \ldots)$ — независимые σ-алгебры, поскольку они порождены независимыми случайными величинами. Теперь рассмотрим все n, тогда получим, что A не зависит от любого $A_1 \in \mathcal{A}(X_1, \ldots)$ (теорема о единственном продолжении меры с сохранением σ-аддитивности), поэтому, в частности, A не зависит от самого себя, ибо $\mathcal{A}_\infty \subset \mathcal{A}(X_1, \ldots)$. Это означает, что (положим в формуле $A_1 := A$) $\mathsf{P}(AA) = \mathsf{P}(A)\mathsf{P}(A)$, откуда следует утверждение теоремы. ■

4.5. Усиленный закон больших чисел

Определение. Сходимость по вероятностной мере: $X_n \xrightarrow{\mathsf{P}} X$, если $\forall\, C > 0$ имеем $\mathsf{P}\left\{\omega\colon |X_n - X| \geqslant C\right\} \to 0$.

Определение. Сходимость распределений: $F_n \Rightarrow F$, т. е. $\forall\, x \in \mathbf{C}(F)$ имеем $F_n(x) \to F(x)$.

Лемма 4.9. *Сходимость по вероятности влечёт слабую сходимость.*

□ Имеем $F_n(x) = \mathsf{P}\left\{X_n < x\right\} = \mathsf{P}\left\{X_n < x,\, |X - X_n| < \varepsilon\right\} + \mathsf{P}\left\{X_n < x,\, |X_n - X| \geqslant \varepsilon\right\} \leqslant \mathsf{P}\left\{X_n < x + \varepsilon\right\} + \mathsf{P}\left\{|X - X_n| \geqslant \varepsilon\right\} = F(x+\varepsilon) + \mathsf{P}\left\{|X - X_n| \geqslant \varepsilon\right\}$. Аналогично $F_n(x) = \mathsf{P}\left\{X_n < x\right\} \geqslant \mathsf{P}\left\{X_n < x,\, |X - X_n| < \varepsilon\right\} \geqslant$

$\mathsf{P}\{X < x-\varepsilon\} - \mathsf{P}\{|X_n - X| \geqslant \varepsilon\} = F(x-\varepsilon) - \mathsf{P}\{|X_n - X| \geqslant \varepsilon\}$. Если x — точка непрерывности F, то $F_n(x)$ никуда на денется. ■

Определение. Сходимость почти всюду по вероятностной мере называется сходимостью почти наверное.

Лемма 4.10. *Из сходимости почти всюду вытекает сходимость по вероятности. Обратное неверно.*

□ Пусть $X_n \to X$ почти всюду. Фиксируем $C > 0$. Рассмотрим $A_n := \bigcap\limits_{k \geqslant n} \{|X_k - X| < C\}$. Тогда $A_n \subset \subset A_{n+1} \subset \ldots$, и $\Omega = \bigcup A_n$ с точностью до множества меры 0, поэтому $\mathsf{P}A_n \to 1$, откуда $\mathsf{P}(\Omega \setminus A_n) \to 0$, но это и означает сходимость по мере. ■

Теорема 4.11 (Лемма Бореля–Кантелли). *Пусть $\{A_n\}$ — последовательность событий в $(\Omega, \mathcal{A}, \mathsf{P})$. Рассмотрим $B = \bigcap\limits_{k=1}^{\infty} \bigcup\limits_{n=k}^{\infty} A_n$. Тогда:*

1°. *Если $\sum \mathsf{P}(A_n)$ сходится, то $\mathsf{P}(B) = 0$.*

2°. *Если $\{A_n\}$ независимы и $\sum \mathsf{P}(A_n)$ расходится, то $\mathsf{P}(B) = 1$.*

□ Событие B содержит то, что происходит бесконечно много раз. Дополнение к B содержит то, что происходит лишь конечное число раз. Для доказательства первого утверждения достаточно заметить, что $\mathsf{P}(B) \leqslant \leqslant \mathsf{P}\{\bigcup\limits_{n=k}^{\infty} A_n\} \leqslant \sum\limits_{n=k}^{\infty} \mathsf{P}(A_n) \to 0$. Отсюда $\mathsf{P}(B) = 0$. Докажем второе утверждение. Имеем $\overline{\bigcup\limits_{n=k}^{\infty} A_n} = \bigcap\limits_{n=k}^{\infty} \overline{A}_n \subset \bigcap\limits_{n=k}^{K} \overline{A}_n$. Отсюда $\mathsf{P}\{\overline{\bigcup\limits_{n=k}^{\infty} A_n}\} \leqslant \mathsf{P}\{\bigcap\limits_{n=k}^{K} \overline{A}_n\} = \{\text{ независимость }\} = \prod\limits_{n=k}^{K} \mathsf{P}\left(\overline{A}_n\right) \leqslant \exp\left(-\sum\limits_{n=k}^{K} \mathsf{P}(A_n)\right)$, ибо $1-x \leqslant e^{-x}$ при $x > 0$. Устремим K к бесконечности, получим в пределе 0 в силу расходимости ряда. Далее остаётся воспользоваться счётной аддитивностью меры: объединение множеств меры нуль есть множество меры нуль. Итак, дополнение имеет меру нуль, откуда следует наше утверждение. ■

Определение. Индикатор (характеристическая функция) события $A \in \mathcal{A}$ — функция $I_A \colon \Omega \to \mathbb{R}$, равная единице на $\omega \in A$ и нулю на $\omega \notin A$.

Лемма 4.12 (Неравенство Колмогорова). *Пусть $X_1, \ldots, X_n$ — независимые случайные величины, для которых $\mathsf{M}X_i = 0$, а $\mathsf{D}X_i = \sigma_i^2$. Тогда $\forall\, a > 0$ имеем $\mathsf{P}\{\max\limits_{i \leqslant n} |X_1 + \ldots + X_i| \geqslant a\} \leqslant \frac{1}{a^2} \sum \sigma_i^2$. Физический смысл леммы: оценка вероятности того, что кривая случайного блуждания суммы X_i зашкалит за уровень a.*

□ Введём $S_i := X_1 + \ldots + X_i$. Рассмотрим искомое событие $A = \{\omega\colon \max\limits_{i \leqslant n} |S_i| \geqslant a\}$ и разобьём его на куски: $A_j := \{\omega\colon \forall\, i < j\; |S_i| < a,\; |S_j| \geqslant a\}$, т. е. это события, отвечающие за то, что зашкал произошёл именно на j-м шаге. Тогда $I_A = \sum\limits_{j=1}^{n} I_{A_j}$, кроме того, в силу независимости и того, что $\mathsf{M}X_i = 0$, имеем

$$\mathsf{M}(X_1 + \ldots + X_n)^2 = \mathsf{M}X_1^2 + \ldots + \mathsf{M}X_n^2,$$

откуда

$$\sum \sigma_i^2 = \mathsf{M}S_n^2 \geqslant \mathsf{M}S_n^2 \cdot I_A = \sum \mathsf{M}S_n^2 \cdot I_{A_j} = \sum \mathsf{M}(S_n - S_j + S_j)^2 \cdot I_{A_j} = \\ = \sum \left(\mathsf{M}(S_n - S_j)^2 \cdot I_{A_j} + 2\mathsf{M}(S_n - S_j)S_j \cdot I_{A_j} + \mathsf{M}S_j^2 \cdot I_{A_j}\right) = (*). \quad (6)$$

Второе слагаемое здесь равно нулю, ибо первый множитель выражается через $X_{j+1}, \ldots, X_n$, а второй — через $X_1, \ldots, X_j$. Первое слагаемое мы отбросим совсем, поскольку оно неотрицательно. Продолжаем: $(*) \geqslant \sum \mathsf{M}S_j^2 \cdot I_{A_j} \geqslant \sum \mathsf{M}a^2 \cdot I_{A_j} = a^2 \mathsf{M} \sum I_{A_j} = a^2 \mathsf{M} I_A = a^2 \mathsf{P}(A)$, что и требовалось доказать. ■

Следствие 4.1. *В условиях неравенства Колмогорова и сходимости $\sum \sigma_i^2$ имеем $\sum X_i$ сходится п. н.*

□ Проверим выполнение критерия Коши почти всюду: $S_{N+i} - S_N = X_{N+1} + \ldots + X_{N+i}$. Рассмотрим $\varepsilon > 0$, тогда $\mathsf{P}\{\max\limits_{i \leqslant n} |S_{N+i} - S_N| \geqslant \varepsilon\} \leqslant \frac{1}{\varepsilon^2} \sum\limits_{j=1}^{n} \sigma_{N+j}^2$. Переходя к пределу при $n \to \infty$, получаем оценку вида $\leqslant \frac{1}{\varepsilon^2} \sum\limits_{j=N+1}^{\infty} \sigma_j^2$, а подходящим выбором N эту величину можно сделать сколь угодно маленькой. Отсюда всё следует, поскольку объединение множеств меры нуль есть множество меры нуль. ■

Лемма 4.13. *Из сходимости ряда $\sum \frac{a_i}{i}$ следует $\frac{1}{n} \sum\limits_{i=1}^{n} a_i \to 0$.*

□ Рассмотрим $b_i := \frac{a_i}{i}$ и переформулируем: из сходимости $\sum b_i$ следует $\frac{1}{n} \sum\limits_{i=1}^{n} i b_i \to 0$. Положим $S_n := b_1 + \ldots + b_n$, тогда S_n равномерно ограничены некоторым числом B. Положим $S_0 = 0$. Очевидно, что $\frac{1}{n} \sum\limits_{i=1}^{n} i b_i =$

$\frac{1}{n}\big((S_n - S_0) + (S_n - S_1) + \ldots + (S_n - S_{n-1})\big)$. В силу критерия Коши имеем $\forall\varepsilon > 0\,\exists N\colon \forall n \geqslant m \geqslant N$ имеем $|S_n - S_m| \leqslant \varepsilon$. Тогда при $n \geqslant N$ получаем $\frac{1}{n}\big|\sum\limits_{i=1}^{n} ib_i\big| \leqslant \frac{1}{n}\big(2BN + (n-N)\varepsilon\big) = 2BN\frac{1}{n} + \frac{n-N}{n}\varepsilon$. Теперь всё доказано, так как подходящим выбором n и ε можно сделать это число сколь угодно малым. ■

Теорема 4.14 (ЗБЧ 2.0β). *Пусть $X_1, \ldots, X_n$ — независимые случайные величины, $m_i = \mathsf{M}X_i$, а $\sigma_i^2 = \mathsf{D}X_i$. Пусть сходится $\sum \frac{\sigma_i^2}{i^2}$, тогда $\lim\limits_{n} \frac{1}{n}\sum\limits_{i=1}^{n}(X_i - m_i) = 0$ почти всюду.*

□ Рассмотрим $\frac{X_i - m_i}{i}$, тогда $\mathsf{M}\frac{X_i - m_i}{i} = 0$, а $\mathsf{D}\frac{X_i - m_i}{i} = \frac{\sigma_i^2}{i^2}$. В силу следствия, $\sum\limits_{i=1}^{n}\frac{X_i - m_i}{i}$ сходится почти всюду. Осталось применить лемму 4.13. ■

Лемма 4.15. *Пусть $\mathsf{M}|X| < \infty$, тогда*

$$\sum \frac{1}{n^2}\mathsf{M}X^2 \cdot I_{\{|X|<n\}} < \infty.$$

□ Положим $A_n = \{n - 1 \leqslant |X| < n\}$, тогда $\mathsf{M}|X| = \sum \mathsf{M}|X| \cdot I_{A_n}$. Обозначим $b_n = \mathsf{M}|X| \cdot I_{A_n}$, откуда $\mathsf{M}X^2 \cdot I_{\{|X|<n\}} = \sum\limits_{k=1}^{n}\mathsf{M}X^2 \cdot I_{A_k} \leqslant \sum\limits_{k=1}^{n} k\mathsf{M}|X| \cdot I_{A_k} = \sum\limits_{k=1}^{n} kb_k$. Осталось доказать сходимость ряда $\sum\limits_{n=1}^{\infty}\frac{1}{n^2}\sum\limits_{k=1}^{n} kb_k = \sum\limits_{k=1}^{\infty} kb_k \sum\limits_{n=k}^{\infty}\frac{1}{n^2} = b_1\cdot\frac{\pi^2}{6} + \sum\limits_{k=2}^{\infty} kb_k\sum\limits_{n=k}^{\infty}\frac{1}{n^2}$. Используем оценку для величины остатка гармонического ряда с помощью интеграла: $\sum\limits_{k=n}^{\infty}\frac{1}{n^2} \leqslant \frac{1}{k-1}$, поэтому второе слагаемое оценивается так: $\sum\limits_{k=2}^{\infty} kb_k\sum\limits_{n=k}^{\infty}\frac{1}{n^2} \leqslant \sum\limits_{k=2}^{\infty}\frac{k}{k-1}b_k$, а последний ряд сходится, поскольку $\sum b_k$ сходится. ■

Теорема 4.16 (УЗБЧ). *Пусть $X_1, \ldots, X_n$ — независимые одинаково распределённые случайные величины. Тогда $\lim\limits_{n}\frac{1}{n}\sum\limits_{i=1}^{n} X_i = m$ почти всюду равносильно $\mathsf{M}X_i = m$, при условии $\mathsf{M}|X_i| < \infty$.*

□ *Достаточность.* Переходя к $X_n' = X_n - m$, мы всё сведём к рассмотрению величин с нулевым M. Пусть X имеет то же распределение, что и X_i, причём $\mathsf{M}X = 0$. Рассмотрим $X_n = Y_n + Z_n$, где $Y_n = X_n \cdot I_{\{|X_n|<n\}}$, т. е. срезки. Тогда

$$\sum \mathsf{P}\{Z_n \neq 0\} = \sum \mathsf{P}\{|X_n| \geqslant n\} = \sum n\mathsf{P}\{n \leqslant |X_n| < n+1\} \leqslant \\ \leqslant \sum \mathsf{M}\{|X_n| \cdot I_{\{n\leqslant|X|<n+1\}}\} = \mathsf{M}\Big\{|X| \cdot \sum I_{\{n\leqslant|X|<n+1\}}\Big\} = \mathsf{M}|X| < \infty. \quad (7)$$

Осталось шарахнуть леммой Бореля – Кантелли.

Обозначим $M_i := \mathsf{M}Y_i$, и всё сведено к доказательству сходимости $\frac{1}{n}\sum Y_i \to 0$ почти наверное. Имеем $\mathsf{M}Y_n = \mathsf{M}X_n \cdot I_{\{|X_n|<n\}}$, поэтому, раз $\mathsf{M}X = 0$, а $\mathsf{M}|X| < \infty$, то $\mathsf{M}\big\{X_n \cdot I_{\{|X_n|<n\}}\big\} = \mathsf{M}X \cdot I_{\{|X|<n\}} \to \mathsf{M}X = 0$ по теореме о мажорируемой сходимости, ибо срезки $|X|$ сходятся к $|X|$ почти всюду. Из анализа известно, что если $m_i \to 0$, то и $\frac{1}{n}\sum\limits_{i=1}^{n} m_i \to 0$, поэтому $\mathsf{D}Y_n = \mathsf{M}Y_n^2 - \big(\mathsf{M}Y_n\big)^2 \leqslant \mathsf{M}Y_n^2 = \mathsf{M}X_n^2 \cdot I_{\{|X_n|<n\}} = \mathsf{M}X^2 \cdot I_{\{|X|<n\}}$. По лемме 4.15 получаем

$$\sum_{n=1}^{\infty}\frac{\mathsf{D}Y_n}{n^2} \leqslant \sum_{n=1}^{\infty}\frac{1}{n^2}\mathsf{M}X^2 \cdot I_{\{|X|<n\}} < \infty,$$

поэтому применима теорема 4.14, по которой $\frac{1}{n}\sum\limits_{i=1}^{n}(Y_i - m_i) \to 0$ почти всюду, что и требовалось доказать.

Необходимость. Имеем

$$\frac{x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{n-1}{n}\cdot\frac{1}{n-1}\sum_{i=1}^{n-1} X_i \to m - 1\cdot m = 0$$

по условию теоремы. С вероятностью 1 происходит лишь конечное число событий $A_n = \Big\{\frac{|X_n|}{n} \geqslant 1\Big\}$, поэтому в силу ЛБК и одинаковой распределённости имеем $\sum \mathsf{P}(A_n) = \sum \mathsf{P}\{|X| \geqslant n\} < \infty$.

Далее, легко видеть, что

$$\sum_{n=1}^{\infty}\mathsf{P}\{|X| \geqslant n\} = \sum_{n=1}^{\infty} n\mathsf{P}\{n \leqslant |X| < n+1\} = \sum_{n=0}^{\infty}(n+1)\mathsf{P}\{n \leqslant |X| < n+1\} - \sum_{n=0}^{\infty}\mathsf{P}\{n \leqslant |X| < n+1\} \geqslant \\ \geqslant \sum_{n=0}^{\infty}\mathsf{M}\{|X| \cdot I_{\{n\leqslant|X|<n+1\}}\} - 1 = \mathsf{M}|X|\sum_{n=0}^{\infty} I_{\{n\leqslant|X|<n+1\}} - 1 = \mathsf{M}|X| - 1,$$

поэтому существует последнее в цепочке математическое ожидание. Теперь можно воспользоваться достаточностью УЗБЧ, откуда следует утверждение теоремы. ■