Содержание

1 Основы теории вероятности.

Вероятностное пространство.

Определение 1. Пусть Ω — множество, тогда $\mathfrak{A} \in 2^\Omega$ называется **алгеброй**, если

1. $\Omega \in A$.
2. $\forall A \in \mathfrak{A}\ \overline{A} \in \mathfrak{A}$. Здесь и далее $\overline{A} = \Omega \setminus A$.
3. $\forall A, B \in \mathfrak{A}\ A \cup B \in \mathfrak{A}$.

При этом Ω называется **множеством элементов событий**, $\mathfrak{A}$ — **набор событий**, $A \in \mathfrak{A}$ — **событие**, $A \cup B = A + B$ — **сумма событий**, $\overline{A}$ — **противоположное событие**, $A \cap B = AB$ — **произведение событий**.

Определение 2. Алгебра является **сигма-алгеброй**, если она замкнута относительно объединения счётного количества своих элементов.

Определение 3. Пусть $\mathfrak{A}$ — сигма-алгебра на Ω. Пусть $P\colon \mathfrak{A} \to [0; +\infty)$ и

1. $P(\Omega) = 1$.
2. Если $\{A_i\}_{i=1}^\infty \subset \mathfrak{A}$ и $\forall A_i A_j = \varnothing$ то

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

Тогда $(\Omega; \mathfrak{A}; P)$ называется **вероятностным пространством**.

Определение 4. Пара событий называется **несовместной**, если их пересечение пусто. Набор событий **несовместен**, если они попарно несовместны.

Определение 5. Пусть $A \subset 2^\Omega$ — алгебра. Тогда минимальная по включению сигма-алгебра $\sigma(A) \supset A$ называется **минимальной сигма-алгеброй**.

Утверждение. *Таковая существует.*

Доказательство. Хотя бы одна такая существует (2^Ω), причём если пересечь сколько угодно сигма-алгебр, то получится искомая сигма-алгебра. □

Определение 6. Пусть $\mathfrak{A}$ — алгебра на Ω, $P\colon \mathfrak{A} \to [0; +\infty)$ и

- $P(\Omega) = 1$.
- Если $\{A_i\}_{i=1}^{\infty} \subset \mathfrak{A}$ и $\bigsqcup\limits_{i=1}^{\infty} A_i \in \mathfrak{A}$, то

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

Тогда $(\Omega; \mathfrak{A}; P)$ — **вероятностное пространство в широком смысле**.

Теорема 1 (О продолжении меры)**.** *Пусть $(\Omega; \mathfrak{A}; P)$ — вероятностное пространство в широком смысле. Тогда существует единственная функция вероятности $Q\colon \sigma(\mathfrak{A}) \to [0 : +\infty)$, такое что $Q\Big|_{\mathfrak{A}} \equiv P$.*
Без доказательства.

Замечание. Эта теорема позволяет нам сказать, например, что мы хотим задать вероятность на отрезках.

Определение 7. Борелевская сигма-алгебра — минимальная σ-алгебра, которая содержит все открытые множества.

Пример. Дискретное вероятностное пространство: $\Omega = \{\omega_i\}_{i=1}^{N}$, $A = 2^{\Omega}$, $P(\{\omega_i\}) = p_i$, $\sum p_i = 1$. Тогда $P(A)$ — сумма вероятностей элементов A.

Пример. Геометрическая вероятность: $\Omega \subset \mathbb{R}^n$, измеримо по Лебегу, $\mu A < +\infty$, $\mathfrak{A}$ состоит из измеримых по Лебегу множеств, $P(A) = \frac{\mu A}{\mu \Omega}$. Обычно при этом $\mathbb{R}^n$ не более чем трёхмерно.

Свойства вероятности.

Свойство 7.1.

$$\forall A, B \in \mathfrak{A}\ A \subset B \Rightarrow P(A) \leqslant P(B)$$

Доказательство. Понятно, что $B \setminus A \in \mathfrak{A}$. Тогда

$$B = A \sqcup (B \setminus A) \Rightarrow P(B) = P(A) + P(B \setminus A) \geqslant P(A)$$

□

Следствие 0.1.

$$\forall A \in \mathfrak{A}\ P(A) \leqslant 1$$

Свойство 7.2.

$$P(A) = 1 - P(\overline{A})$$

Свойство 7.3.

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

Доказательство.

$$B = (B \setminus AB) \sqcup AB \Rightarrow P(B) = P(B \setminus AB) + P(AB)$$

Тогда

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

□

Утверждение (Формула включений-исключений)**.**

$$P(A_1 + \dots + A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{\substack{i,j=1\\ i<j}}^{n} P(A_i A_j) + \sum_{\substack{i,j,k=1\\ i<j<k}}^{n} P(A_i A_j A_k) - \dots + (-1)^n P(A_1 \cdots A_n)$$

Доказательство. Мне лень это писать, докажите сами по индукции. □

Утверждение.

$$P\left(\bigcup_i A_i\right) \leqslant \sum_i P(A_i)$$

Доказательство. Пусть $B_1 = A_1$, $B_2 = A_2\overline{A_1}$, $B_3 = A_3\overline{A_1 \cup A_2}$ и так далее. Тогда

$$\bigcup_i A_i = \bigsqcup_i B_i$$

При этом $B_i \subset A_i$, а значит

$$\sum_i P(A_i) \geqslant \sum_i P(B_i)$$

□

Теорема 2. *Пусть $(\Omega; \mathfrak{A}; P)$ — вероятностное пространство. Тогда следующие утверждения равносильны:*

1. *P счётно-аддитивна.*
2. *P конечно-аддитивна и $\forall \{B_i\}_{i=1}^{\infty} : B_{i+1} \subset B_i,\ B = \bigcap\limits_{i=1}^{\infty} B_i\ \lim\limits_{n\to\infty} P(B_i) = P(B)$ (непрерывность сверху).*
3. *P конечно-аддитивна и $\forall \{C_i\}_{i=1}^{\infty} : C_{i+1} \supset B_i,\ C = \bigcup\limits_{i=1}^{\infty} C_i\ \lim\limits_{n\to\infty} P(C_i) = P(C)$ (непрерывность сверху).*

Доказательство. Равносильность двух непрерывностей тривиально из формул де Моргана.
Докажем, что из 1 следует 2. Конечная аддитивность есть, докажем непрерывность сверху. Пусть $A_1 = B_1\overline{B_2}$, $A_2 = B_2\overline{B_3}$ и так далее. Очевидно, A_i несовместны. Также очевидно, что A_i несовместны с B. Также заметим, что

$$B_n = B \sqcup \bigsqcup_{i=n+1}^{\infty} A_i$$

Отсюда $P(B_n) = P(B) + \sum\limits_{i=n+1}^{\infty} P(A_i)$, а справа остаток (очевидно, сходящегося) ряда, который стремится к нулю при $n \to \infty$.
Теперь из 2 докажем 1. Рассмотрим $\{A_i\}_{i=1}^{\infty}$ несовместные. Очевидно,

$$\sum_{i=1}^{\infty} P(A_i) = \lim_{n\to\infty} \sum_{i=1}^{n} P(A_i)$$

А ещё мы знаем, что

$$\bigsqcup_{i=1}^{\infty} A_i = \bigsqcup_{i=1}^{n} A_i \sqcup \bigsqcup_{i=n+1}^{\infty} A_i$$

То есть

$$\lim_{n\to\infty} P\left(\bigsqcup_{i=1}^{n} A_i\right) = \lim_{n\to\infty}\left(P\left(\bigsqcup_{i=1}^{\infty} A_i\right) - P\left(\bigsqcup_{i=n+1}^{\infty} A_i\right)\right) = P\left(\bigsqcup_{i=1}^{\infty} A_i\right) - \lim_{n\to\infty} P\left(\bigsqcup_{i=n+1}^{\infty} A_i\right)$$

Второе слагаемое — ноль но непрерывности меры, а отсюда счётная аддитивность. □

Условная вероятность.

Замечание. Пусть $|\Omega| = n$, $|A| = k$, $|B| = m$, $|AB| = l$. Если мы знаем, что B произошло, как узнать вероятность того, что произошло A? Ну, это

$$\frac{l}{m} = \frac{l/n}{m/n} = \frac{P(AB)}{P(B)}$$

Определение 8. Пусть $(\Omega; \mathfrak{A}; P)$ — вероятностное пространство, $B \in \mathfrak{A}$, $P(B) > 0$. Тогда **условной вероятностью** A при условии B называется

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

Также обозначается $P_B(A)$.

Свойство 8.1. *Несложно проверить, что условная вероятность является вероятностью.*

Утверждение (Произведение вероятностей)**.** *Несложно по определению проверить*

$$P(A_1 \cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2) \cdots P(A_n|A_1A_2 \cdots A_{n-1})$$

Теорема 3 (Формула полной вероятности)**.** *Пусть $A \in \mathfrak{A}$, $B_i \in \mathfrak{A}$ несовместны, $A \subset \bigsqcup_{i=1}^{n} B_i$ (обычно объединение равно Ω), и $\forall i \in [0 : n]$ $P(B_i) > 0$. Тогда*

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

Доказательство.

$$P(A) = P\left(A \cap \bigsqcup_{i=1}^{n} B_i\right) = P\left(\bigsqcup_{i=1}^{n} A \cap B_i\right)$$

Всё. □

Теорема 4 (Формула Байеса)**.** *Пусть $A, B \in \mathfrak{A}$, $P(A) > 0$, $P(B) > 0$. Тогда*

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

Доказательство. Очевидно из определения. □

Определение 9. События $A, B \in \mathfrak{A}$ называются **независимыми**, если

$$P(AB) = P(A)P(B)$$

Определение 10. Говорят, что $A_1, A_2, \ldots, A_n \in \mathfrak{A}$ **независимы в совокупности**, если $P(A_1A_2 \cdots A_n) = P(A_1)P(A_2) \cdots P(A_n)$

Свойство 10.1. *Несложно проверить, что независимость событий A, B равносильна $P(A|B) = P(A)$.*

Свойство 10.2. *Независимые в совокупности события попарно независимы. Обратное неверно.*

Определение 11. Пусть у нас есть два вероятностных пространства: $(\Omega_1, \mathfrak{A}_1; P_1)$ и $(\Omega_2, \mathfrak{A}_2; P_2)$. Рассмотрим вот такое вероятностное пространство: $(\Omega, \mathfrak{A}; P)$, где $\Omega = \Omega_1 \times \Omega_2$, $\mathfrak{A}$ — минимальная σ-алгебра, включающая в себя $\mathfrak{A}_1 \times \mathfrak{A}_2$,

$$P((A_1; A_2)) = P_1(A_1)P_2(A_2)$$

Тогда $(\Omega_1, \mathfrak{A}_1; P_1)$ и $(\Omega_2, \mathfrak{A}_2; P_2)$ — независимые испытания.

Схема Бернулли.

Пример. Схема Бернулли: $\Omega_1 = \{0; 1\}$, $\mathfrak{A}_1 = 2^{\Omega_1}$, $P_1(1) = p$, $P_1(0) = 1 - p = q$. Хочется рассмотреть эту штуку в степени n (то есть n одинаковых независимых испытаний). Тогда что у нас получается для $\omega \in \Omega = \Omega_1^n$?

$$P(\omega) = \sum_{i=1}^{n} P_i(\omega_i) = p^{\sum \omega_i} q^{n - \sum \omega_i}$$

Посчитаем тут такую вероятность: пусть S_n — количество успехов в n испытаниях? Посчитаем вероятность того, что $S_n = k$? Очевидно, оно равно $\binom{n}{k} p^k q^{n-k}$.

Утверждение. *Пусть k^* — наиболее вероятное число успехов в Бернуллиевских испытаниях. Тогда*

$$k^* = \begin{cases} p(n-1) \text{ или } p(n-1)+1 & p(n-1) \in \mathbb{N} \\ \lceil p(n-1) - 1 \rceil & p(n-1) \notin \mathbb{N} \end{cases}$$

Доказательство. Давайте рассмотрим вот такое частное:

$$\frac{P(S_n = k+1)}{P(S_n = k)}$$

Чему оно равно?

$$\frac{P(S_n = k+1)}{P(S_n = k)} = \frac{\binom{n}{k+1} p^{k+1} q^{n-k-1}}{\binom{n}{k} p^k q^{n-k}} = \frac{p}{q} \cdot \frac{n-k}{k+1}$$

Нам хочется оценить, больше это чем 1 или меньше (это позволит нам найти K^*). Ну,

$$\frac{p}{q} \cdot \frac{n-k}{k+1} > 1 \Leftrightarrow p(n-k) > q(k+1) \Leftrightarrow pn - pk > k = pk + 1 - p \Leftrightarrow pn > k + 1 - p \Leftrightarrow pn + p - 1 > k$$

То есть возрастание достигается при $k < p(n-1) - 1$, а иначе убывание. Тогда где экстремум? Рассмотрим $k = p(n-1) - 1$. Если это целое число, то там $P(S_n = k+1) = P(S_n = k)$, и это самое k даёт значение больше остальных. То есть $k^* = p(n-1) - 1$ или $k^* = p(n-1)$.
А что если оно не целое? То надо куда-то округлить. А именно вверх, потому что тогда оно больше, чем следующее, а предыдущее меньше его. □

Пример. Пусть $n = 10000$, $p = \frac{1}{10000}$. Давайте посчитаем $P(S_n > 3)$. Ну, это

$$1 - P(S_n \leqslant 3) = 1 - q^{10000} - 10000pq^{10000-1} - \binom{10000}{2} p^2 q^{10000-2} - \binom{10000}{3} p^3 q^{10000-3}$$

Фиг мы такое посчитаем.

Пример. Или если взять $p = q = 0.5$, то при $n = 5 \cdot 10^3$ мы не сможем нормально посчитать $P(S_n = 2349)$.

Замечание. Ну и как такое считать?

Теорема 5 (Теорема Пуассона)**.** *Пусть у нас есть несколько схем Бернулли. В первой одно испытание и вероятность успеха p_1, во второй — 2 и вероятность успеха p_2, в n-ной n испытаний и вероятность p_n. Пусть $np_n \underset{n\to\infty}{\longrightarrow} \lambda > 0$. Тогда*

$$P(S_n = k) \underset{n\to\infty}{\longrightarrow} e^{-\lambda} \frac{\lambda^k}{k!}$$

Доказательство. Известно,

$$P(S_n = k) = \frac{1}{k!} n(n-1)\cdots(n-k+1) p_n^k (1-p_n)^{n-k}$$

Известно, что

$$np_n = \lambda + o(1) \Rightarrow p_n = \frac{\lambda}{n} + o\left(\frac{1}{n}\right)$$

Тогда

$$P(S_n = k) = \frac{1}{k!} \cancelto{1}{n(n-1)\cdots(n-k+1)\frac{1}{n^k}} \cancelto{\lambda^k}{(\lambda + o(1))^k} \frac{\cancelto{e^{-\lambda}}{\left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^n}}{\cancelto{1}{\left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^k}} \xrightarrow[n\to\infty]{} e^{-\lambda}\frac{\lambda^k}{k!}$$

□

Лемма 1. *Пусть* $p \in (0;1)$, $H(x) = x \ln \frac{x}{p} + (1-x)\ln\frac{1-x}{1-p}$. *Пусть* $p^* = \frac{k}{n}$. *Пусть* $k \to +\infty$, $n-k \to +\infty$. *Тогда*

$$P(S_n = k) \sim \frac{1}{\sqrt{2\pi n p^*(1-p^*)}} \exp(-nH(p^*))$$

Доказательство. Мы знаем формулу Стирлинга

$$n! \sim \sqrt{2\pi n} n^n e^{-n}$$

Тогда

$$P(S_n = k) = \frac{\sqrt{2\pi n} n^n e^{-n}}{\sqrt{2\pi \underbrace{k}_{np^*}} k^k e^{-k} \sqrt{2\pi \underbrace{(n-k)}_{n(1-p^*)}} (n-k)^{n-k} e^{-n+k}} p^k (1-p)^{n-k} =$$

$$= \frac{n^n p^k (1-p)^{n-k}}{\sqrt{2\pi n p^*(1-p^*)} k^k (n-k)^{n-k}} = \frac{1}{\sqrt{2\pi n p^*(1-p^*)}} \exp \underbrace{\ln \frac{n^n p^k (1-p)^{n-k}}{k^k (n-k)^{n-k}}}_{L}$$

При этом

$$L = \ln \frac{n^n p^k (1-p)^{n-k}}{k^k (n-k)^{n-k}} = \ln \frac{n^n p^k (1-p)^n (n-k)^k}{k^k (n-k)^n (1-p)^k} =$$

$$= \ln \left(\underbrace{\frac{n^n}{(n-k)^n}}_{(1-p^*)^{-n}} (1-p)^n \right) + \ln \frac{p^k (n-k)^k}{(np^*)^k (1-p)^k} =$$

$$= n \ln \frac{1-p}{1-p^*} + k \ln \frac{p}{p^*} + k \ln \frac{(n-k)}{n(1-p)} = n \ln \frac{1-p}{1-p^*} + k \ln \frac{p}{p^*} + k \ln \frac{1-p^*}{1-p} =$$

$$= -(n-k) \ln \frac{1-p^*}{1-p} - k \ln \frac{p^*}{p} = -n \underbrace{\left(p^* \ln \frac{p^*}{p} + (1-p^*) \ln \frac{1-p^*}{1-p} \right)}_{H(p^*)}$$

Это ли не то, что нам надо? □

Лемма 2.

$$H(x) = \frac{(x-p)^2}{2p(1-p)} + O((x-p)^3)$$

Доказательство.

$$H'(x) = \ln\frac{x}{p} + x\cdot\frac{p}{x}\cdot\frac{1}{p} - \ln\frac{1-x}{1-p} - 1 = \ln\frac{x}{p} - \ln\frac{1-x}{1-p}$$

$$H''(x) = \frac{1}{x} + \frac{1}{1-x}$$

Тогда $H'(p) = 0$, $H''(p) = \frac{1}{p(1-p)}$. По Тейлору получаем искомое. □

Теорема 6 (Локальная теорема Муавра — Лапласа). *Пусть* $p \in (0;1)$, $H(x) = x\ln\frac{x}{p} + (1-x)\ln\frac{1-x}{1-p}$. *Пусть* $p^* = \frac{k}{n}$. *Пусть* $k \to +\infty$, $n - k \to +\infty$.
Пусть $k - np = p(n^{2/3})$. *Тогда*

$$P(S_n = k) \sim \frac{1}{\sqrt{2\pi np(1-p)}}\exp\left(-\frac{(k-np)^2}{2np(1-p)}\right)$$

Доказательство. Известно

$$P(S_n = k) \sim \frac{1}{\sqrt{2\pi np*((1-p*)}}\exp(-nH(p^*))$$

Отсюда

$$P(S_n = k) \sim \frac{1}{\sqrt{2\pi np*(1-p*)}}\exp(-n\frac{(p^*-p)^2}{2p(1-p)} + n\cdot O((p^*-p)^3))$$

Заметим, что $\frac{k}{n} - p = o(n^{-1/3})$, поэтому $p* \approx p$. Тогда

$$\begin{aligned}P(S_n = k) &\sim \frac{1}{\sqrt{2\pi np(1-p)}}\exp\left(-n\frac{(p-k/n)^2}{2p(1-p)} + O(n(k/n-p)^3)\right)\\ &\sim \frac{1}{\sqrt{2\pi np(1-p)}}\exp\left(-n\frac{(np-k)^2}{2p(1-p)n^2} + o(1)\right)\end{aligned}$$

Что и требовалось доказать. □

Теорема 7 (Интегральная теорема Муавра — Лапласа). *Пусть*

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}\ \mathrm{d}t$$

Далее мы будем называть эту функцию функцией стандартного нормального распределения. Тогда

$$\sup_{-\infty<x_1<x_2<+\infty}\left|P\left(x_1 \leqslant \frac{S_n - np}{\sqrt{npq}} \leqslant x_2\right)(\Phi(x_2) - \Phi(x_1))\right| \underset{n\to\infty}{\longrightarrow} 0$$

Иными словами

$$P\left(x_1 \leqslant \frac{S_n - np}{\sqrt{npq}} \leqslant x_2\right) \approx \frac{1}{\sqrt{2\pi}}\int\limits_{x_1}^{x_2} e^{-\frac{t^2}{2}}\ \mathrm{d}t$$

Пока без доказательства.

Замечание. Оценка теоремы Пуассона.
Обычно в задачах np_n не стремится, а просто равно $\lambda > 0$. Тогда

$$\sum_{k=0}^{\infty}\left|P(S_n = k) - e^{-\lambda}\frac{\lambda^k}{k!} \leqslant \frac{\lambda}{n}\right| \leqslant \frac{2\lambda}{n}\min\{2;\lambda\}$$

Оценка локальной теоремы Лапласа. Если $|p^* - p| \leqslant \frac{1}{2} mn \min\{p; q\}$, то

$$P(S_n = k) = \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left(-\frac{(k-np)^2}{2np(1-p)}\right)(1 + \varepsilon(k;n))$$

Где

$$\varepsilon(k;n) = \exp\left(\theta \frac{|k-np|^3}{3n^2p^2q^2} + \frac{1}{npq}\left(\frac{1}{6} + |k-np|\right)\right) \qquad |\theta| < 1$$

Оценка интегральной теоремы Лапласа.

$$\sum_x \left| P\left(\frac{S_n - np}{\sqrt{npq}} \leqslant x\right) - \Phi(x)\right| \leqslant \frac{p^2+q^2}{\sqrt{npq}}$$

Пример. Пусть у нас есть два узла связи на 2000 пользователей в каждом. И у нас есть канал связи, который пропускает N. Хочется минимизировать N, но так, чтобы вероятность перегрузки была меньше $\frac{1}{100}$. Будем предполагать, что люди пользуются данным каналом связи в течение двух минут из одного часа, то есть каждый пользователь может пользоваться каналом в данный момент с вероятностью $p = \frac{1}{30}$.
Ну так и что мы хотим по сути? Мы хотим $P(S_{2000} > N) < \frac{1}{100}$, что равносильно $P(S_{2000} \leqslant N) \geqslant \frac{99}{100}$.
Используем Пуассона: $np \approx 6.67$.

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

Это мы хрен посчитаем, но, короче, получится $N = 87$.
А если применить интегральную теорему Муавра — Лапласа, то получим мы

$$N = \left\lceil q_{\frac{99}{100}} \sqrt{npq} + np \right\rceil = 86$$

Где $q_{\frac{99}{100}}$ — такое число, что $\Phi(q_{\frac{99}{100}}) = \frac{99}{100}$.

Определение 12. Если $\alpha \in (0;1)$ и $\Phi(q_\alpha) = \alpha$, то q_α называется **квантилем порядка** α.

2 Случайные величины.

2.1 Одномерные случайные величины.

Распределение случайных величин, функция распределения случайных величин.

Определение 1. **Борелевская сигма-алгебра** — это минимальная сигма-алгебра, содержащая все открытые множества.

Определение 2. Пусть $\Omega; \mathfrak{A}$ — множество с сигма-алгеброй. Тогда такое $\xi\colon \Omega \to \mathbb{R}$, что $\forall B \in \mathfrak{B}\ \xi^{-1}(B) \in \mathfrak{A}$, называется **случайной величиной**.

Определение 3. Пусть $(\Omega; \mathfrak{A}; P)$ — вероятностное пространство, ξ — случайная величина. Тогда распределение ξ — функция

$$P_\xi\colon \underset{B \mapsto P(\{\omega | \xi(\omega) \in B\})}{\mathfrak{B} \to \mathbb{R}}$$

Замечание. $P(\{\omega \mid \xi(\omega) \in B\})$ обозначается $P(\xi \in B)$.

Свойство 3.1. *P_ξ — вероятность на* $(\mathbb{R}; \mathfrak{B})$.

Определение 4. Пусть ξ — случайная величина. Тогда

$$F_\xi(t) = P(\xi \leqslant t)$$

называется **функцией распределения** ξ

Свойство 4.1. *Очевидно, функция распределения нестрого возрастает.*

Свойство 4.2. *Не менее очевидно, $F_\xi(+\infty) = 1$, $F_\xi(-\infty) = 0$;*

Свойство 4.3. *Функция распределения непрерывна справа.*

Доказательство. Возьмём $F(t + \varepsilon_n) - F(t)$. Она равна $P(t < \xi \leqslant t + \varepsilon_n)$. При $\varepsilon_n \to 0$, получим, что аргумент P стремится к $\varnothing$, а значит $P(t < \xi \leqslant t + \varepsilon_n)$ стремится к нулю. □

Лемма 1. *Пусть $P\colon \mathfrak{B} \to \mathbb{R}$ — некоторая функция. Тогда*

$$\forall \{B_n\}_{n=1}^{\infty} \subset \mathfrak{B} : B_{n+1} \subset B_n \; P(B_n) \underset{n\to\infty}{\to} P\left(\bigcap_{n=1}^{\infty} B_n\right) \Leftrightarrow$$

$$\Leftrightarrow \forall \{C_n\}_{n=1}^{\infty} \subset \mathfrak{B} : C_{n+1} \subset C_n, \bigcap_{n=1}^{\infty} C_n = \varnothing \; P(C_n) \underset{n\to\infty}{\to} 0$$

Доказательство. Следствие слева направо очевидно. Наоборот. Пусть

$$\bigcap_{n=1}^{\infty} B_n = B$$

Возьмём $C_n = B_n\overline{B}$. Тогда, очевидно, C_n подходят под условие справа, а значит $P(C_n) \longrightarrow 0$.
Также несложно заметить, что $P(B_n) = P(C_n) + P(B)$, а отсюда получим $P(B_n) \longrightarrow P(B)$. □

Теорема 1. *Пусть F — монотонно возрастающая непрерывная слева функция, равная нулю в $-\infty$ и единице в $+\infty$. Тогда существует вероятностное пространство и случайная величина в нём, что F — её функция распределения.*

Доказательство. Пусть $\Omega = \mathbb{R}$, $\mathfrak{A}$ — алгебра, состоящая из множеств вида $\bigsqcup\limits_{k=1}^{n} (a_k; b_k]$ или $(-\infty; b)$ или $(a; \infty)$ или $\mathbb{R}$.

$$P\left(\bigsqcup_{k=1}^{n} (a_k; b_k]\right) = \sum_{k=1}^{n} F(b_k) - F(a_k)$$
$$P((-\infty; b)) = F(b) \qquad P((a; +\infty)) = 1 - F(a) \qquad P(\mathbb{R}) = 1$$

Получим вероятностное пространство в широком смысле (разве что непрерывность сверху надо проверить). Ну, проверим её, используя лемму. Пусть, не умаляя общности, $A_n = (a_{n,1}; a_{n,2}]$, $A_{n+1} \subset A_n$ и пересечение всех пусто.
Из непрерывности F справа следует, что существует $B_n = (b_{n,1}; b_{n,2}]$, где $\operatorname{Cl} B_n \subset A_n$ и $P(A_n) - P(B_n) \leqslant \varepsilon 2^{-n}$. Тогда пересечение всех B_n также пусто.
Предположим, что существует M, такое что $\forall n \; A_n \in [-M; M]$. $[-M; M]$ — компакт, следовательно. Заметим, что

$$[-M; M] = \bigcup_{k=1}^{\infty} [-M; M] \setminus \operatorname{Cl} B_n$$

Справа — открытое покрытие компакта, значит из него можно вытащить конечное подпокрытие, то есть пересечение какого-то конечного числа B_n пусто. Пусть это пересечение от 1 до n_0. Тогда

$$P(A_{n_0}) = P(A_{n_0} \setminus \bigcap_{k=1}^{n_0} B_k) + P\left(\bigcap_{k=1}^{n_0} B_k\right)$$

Отсюда $P\left(\bigcap\limits_{k=1}^{n_0} B_k\right) = 0$. А

$$P(A_{n_0} \setminus \bigcap_{k=1}^{n_0} B_k) = P(\bigcap_{k=1}^{n_0} A_{n_0} \setminus B_k) \leqslant P(\bigcap_{k=1}^{n_0} A_k \setminus B_k) \leqslant \sum_{k=1} P(A_n) - P(B_k) \leqslant \varepsilon \sum k = 1^{n_0} 2^{-n} < \varepsilon$$

Если же мы не находимся в промежутке $[-M; M]$, то можно указать такие M_1 и M_2, что $P((-\infty; M_1) \cup (M_2; +\infty)) < \frac{\varepsilon}{2}$. Тогда

$$P(A_n) = P(A_n[M_1; M_2]) + P(A_n\overline{[M_1; M_2]})$$

Левую часть суммы мы разобрали, а правая мала т.к. M_2, что $P((-\infty; M_1) \cup (M_2; +\infty)) < \frac{\varepsilon}{2}$. Осталось предъявить случайную величину $\xi(\omega) = \omega$. □

Типы распределений.

Дискретные случайные величины и распределения.

Определение 5. Случайная величина ξ называется **дискретной**, если существует такое не более чем счётное множество E, что $P_\xi(E) = 1$.

Пример. Вырожденное: $P(\xi = c) = 1$. Обозначают $I(c)$ или I_c.

Пример. Распределение Бернулли: $P(\xi = 0) = p$, $P(\xi = 1) = q = 1 - p$. Обозначение: $\mathrm{Bern}(p)$.

Пример. Биномиальное распределение: $P(\xi = k) = \binom{n}{k} p^k q^{n-k}$. Обозначение: $\mathrm{Bin}(n; p)$,

Пример. Отрицательное биномиальное распределение: $\xi = (\min n : S_n = r) - r$, где $r \in \mathbb{n}$. То есть

$$P(\xi = k) = \binom{k + r - 1}{r - 1} p^r q^k$$

Обозначение: $\mathrm{NB}(r; p)$. Также это обобщается на произвольное r при помощи гамма-функции.
В случае $r = 1$ распределение называется геометрическим. Геометрическое распределение — количество неудач до первого успеха.

Пример. Распределение Пуассона:

$$P(\xi = k) = e^{-\lambda} \frac{\lambda^k}{k!} \qquad k \in \mathbb{Z}_+$$

Обозначение: $\mathrm{Pois}(\lambda)$.

Определение 6. Носителем случайной величины ξ называется минимально по включению замкнутое множество E, удовлетворяющее условию $P(\xi \in E) = 1$.

Абсолютно непрерывные случайные величины и распределения.

Определение 7. Величина ξ (или её случайно распределение) называется **абсолютно непрерывной**, если существует $p \in L(\mathbb{R} \to \mathbb{R})$ с неотрицательными значениями такая что $P(\xi \in B) = \int_B P(x)\ \mathrm{d}x$. В таком случае p называется **плотностью** ξ.

Свойство 7.1. *В таком случае*

$$F(t) = \int\limits_{-\infty}^{t} p(x)\ \mathrm{d}x$$

То есть плотность почти всюду равна производной функции распределения.

Свойство 7.2. $P(\xi = c) = 0$.

Свойство 7.3. *Функция распределения абсолютно непрерывной случайной величины непрерывна на* $\mathbb{R}$.

Свойство 7.4.

$$P(x_0 \leqslant \xi \leqslant x_0 + h) = P(x_0 < \xi < x_0 + h) = F(x_0 + h) - F(x_0) = p(x_0)h + o(h)$$

Свойство 7.5. *Пусть* $E = \operatorname{supp} p$. *Тогда* E *является носителем по нашему прошлому определению.*

Пример.

$$p(x) = \frac{1}{b-a}\chi_{[a;b]}$$

Тогда

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a;b) \\ 1 & x \geqslant b \end{cases}$$

Обозначение: $U[a;b]$.

Утверждение. *Пусть* $\xi = U[a;b]$, $c > 0$. *Тогда* $\eta = c\xi + d = U[ac+d; bc+d]$

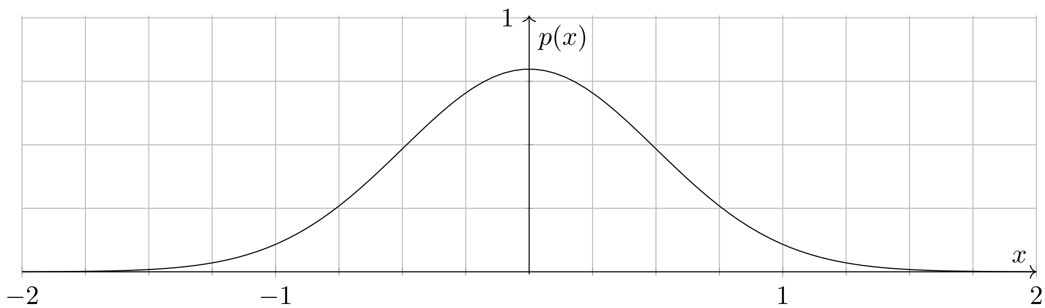
Доказательство.

$$P(\eta \leqslant t) = P(c\xi + d \leqslant t) = P\left(\xi \leqslant \frac{t-d}{c}\right)$$

Несложно проверить, что это именно $U[ac+d; bc+d]$. □

Пример. Нормальное (гауссовское) распределение: $\mathrm{N}(\mu; \sigma^2)$:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$



$\mathrm{N}(0;1)$ — стандартное нормальное распределение. Ещё оно обозначается $\Phi(x)$.

Утверждение. *Пусть* $\xi = \mathrm{N}(\mu; \sigma^2)$, $\eta = a\xi + b$. *Тогда* $\eta = \mathrm{N}(a\mu_b; a^2\sigma^2)$.

Доказательство. Пусть $a > 0$. Обозначим $y = ax + b$. Тогда

$$P(\eta \leqslant t) = P(a\xi + b \leqslant t) = P(\xi \leqslant \frac{t-b}{a}) = F_\xi(\frac{t-b}{a}) = \frac{1}{2\pi\sigma^2} \int\limits_{-\infty}^{\frac{t-b}{a}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\, \mathrm{d}x =$$

$$= \frac{1}{\sqrt{2\pi\sigma^2 a^2}} \int\limits_{-\infty}^{t} \exp\left(-\frac{(\frac{y-b}{a} - \mu)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi\sigma^2 a^2}} \int\limits_{-\infty}^{t} \exp\left(-\frac{(y-b-a\mu)^2}{2a^2\sigma^2}\right)$$

Что и требовалось доказать. При $a < 0$ аналогично. □

Следствие 1.1. *Если* $\xi = \mathrm{N}(0;1)$, *то* $\sigma\xi + \mu = \mathrm{N}(\mu; \sigma^2)$.

Пример. Распределение Коши: $\mathrm{Cauchy}(x_0; \gamma)$.

$$p(x) = \frac{1}{\pi\gamma} \cdot \frac{1}{1 + \left(\frac{x - x_0}{\gamma}\right)^2}$$

Тогда

$$F(t) = \frac{1}{\pi\gamma} \int\limits_{-\infty}^{t} \frac{\mathrm{d}x}{1 + \left(\frac{x - x_0}{\gamma}\right)^2} = \frac{1}{\pi} \tan^{-1}\left(\frac{t - x_0}{\gamma}\right) + \frac{1}{2}$$

Пример. Экспоненциальное распределение: $\mathrm{Exp}(\lambda)$.

$$p(x) = \lambda e^{-\lambda x} \chi_{\mathbb{R}_+}$$

Тогда

$$F(t) = 1 - e^{-\lambda t} \chi_{\mathbb{R}_+}$$

Пример. Г-распределение: $\Gamma(k, \lambda)$. Сначала для $k \in \mathbb{N}$, $\lambda > 0$.

$$p(x) = \frac{\lambda^k x^{k-1}}{(k-1)!} e^{-\lambda x} \qquad x \geqslant 0$$

Для $k = \alpha > 0$ заменим $(k-1)!$ на $\Gamma(\alpha)$.

Утверждение. *Пусть ξ — абсолютно непрерывная случайная величина с плотностью p_ξ. Пусть $g \in C^{(1)}(\mathbb{R} \to \mathbb{R})$ — строго монотонна. Пусть $\eta = g(\xi)$. Тогда*

$$p_\eta(y) = p_\xi(g^{-1}(y)) \left|\frac{\mathrm{d}}{\mathrm{d}y} g^{-1}(y)\right| = p_\xi(g^{-1}(y)) \left|\frac{1}{g'(g^{-1}(y))}\right|$$

Доказательство. Во-первых, у условиях теоремы g^{-1} существует. Также второе равенство следует из первого в силу теоремы об обратном отображении.
Не умаляя общности, g строго возрастает. Тогда

$$F_\eta(y) = P(g(\xi) \leqslant y) \overset{g\uparrow}{=} P(\xi \leqslant g^{-1}(y)) = \int\limits_{-\infty}^{g^{-1}(y)} p_\xi(x)\ \mathrm{d}x$$

Продифференцировав это равенство по y получим искомое равенство. □

Сингулярные случайные величины и распределения.

Определение 8. x называется **точкой роста** монотонно возрастающей функции f, если $\forall \varepsilon > 0$ $f(x + \varepsilon) - f(x - \varepsilon) > 0$.

Определение 9. Случайная величина ξ (и её распределение) называются **сингулярными**, если $F_\xi \in C(\mathbb{R})$ и мера множества точек F_ξ роста равна нулю.

Пример. Функция Кантора — функция, которая выглядит так: в нуле она равна нулю, в единице — единице, а во всех остальных точках строится так: отрезок $[0; 1]$ делится на три части и в средней части равна среднему значению краёв (т.е. $\frac{1}{2}$). И так далее.
Оня является функцией распределения сингулярной случайной величины.

Теорема 2 (Теорема Лебега)**.** *Пусть F — функция распределения. Тогда существуют единственные F_{disc}, F_{ac} и F_{sing}, которые в сумме дают F.*
Без доказательства.

2.2 Многомерные случайные величины.

Распределение многомерных случайных величин.

Определение 10. Вектор ξ называется **случайным вектором** или **многомерной случайной величиной**, если ξ_i — случайная величина.

Определение 11. Распределением случайного вектора называется функция P_ξ, определённая на $\mathfrak{B}^n$, заданная так:

$$P_\xi(B_1;\dots;B_n) = P(\{(\omega_1;\dots;\omega_n) \mid \xi_1(\omega_1) \in B_1 \wedge \dots \wedge \xi_n(\omega_n) \in B_n\})$$

Последнее обычно обозначается так: $P(\xi_1 \in B_1, \xi_2 \in B_2;\dots;\xi_n \in B_n)$.

Определение 12. Функцией распределения случайного вектора ξ называется функция

$$F_\xi(t_1;\dots;t_n) = P_\xi(\forall i \in [1:n]\ \xi_i \leqslant t_i)$$

Утверждение.

$$\begin{aligned} P(\forall i \in [1:n]\ a_i < \xi_i \leqslant b_i) = &\\ &F(b_1;b_2;\dots;b_n)\\ &- F(a_1;b_2;\dots;b_n) - F(b_1;a_2;\dots;b_n) - \dots - F(b_1;b_2;\dots;a_n)\\ &+ \vdots\\ &\pm F(a_1;a_2;\dots;a_n)\end{aligned}$$

То есть в этой сумме участвует F от a_i и b_i в произвольном сочетании, при этом минус стоит там, где нечётное количество a_i.

Доказательство.

$$P(\forall i \in [1:n]\ a_i < \xi_i \leqslant b_i) = P(\xi_1 \leqslant b_1, \forall i \in [2:n]\ a_i < \xi_i \leqslant b_i) - P(a_1 \geqslant \xi_1, \forall i \in [2:n]\ a_i < \xi_i \leqslant b_i)$$

Проведя такую операцию несколько раз, получим искомое. □

Замечание. Если ввести обозначение $\Delta_{a_i;b_i}F = F(\cdot;\dots;b_i;\cdot;\dots;\cdots) - F(\cdot;\dots;a_i;\cdot;\dots;\cdots)$, то арифметическая сумма из утверждения выше записывается как

$$\Delta_{a_1;b_1}\Delta_{a_2;b_2}\cdots\Delta_{a_n;b_n}F$$

Свойство 12.1.

$$F(+\infty;\cdots;+\infty) = 1$$
$$F(-\infty;\cdots;-\infty) = 0$$

F непрерывна справа.

Теорема 3. *Если функция распределения удовлетворяет трём свойствам выше, то она является функцией распределения некоторого случайного вектора.*

Доказательство. Аналогично одномерному случаю. □

Определение 13. Случайный вектор (и его распределение) называется **дискретным**, если существует не более чем счётное множество E такое что $P(\xi \in E) = 1$.

Пример. Полиномиальное распределение. Пусть $p = (p_1;\dots;p_m)$, где $\sum p_i = 1$. Обозначается $\mathrm{Poly}(n;p)$. Физическая интерпретация такая: мы бросаем кубик с m гранями n раз. Пусть $S_{n,j}$ — количество исходов типа j в n независимых испытаниях. Тогда искомая случайная величина обладает распределением

$$P(S_{n,1} \in B_1; S_{n,2} \in B_2;\dots;S_{n,m} \in B_m)$$

Рассмотрим $P(S_{n,1} = k_1; S_{n,2} = k_2;\dots;S_{n,m} = k_m)$, где $\sum k_m = n$. Чему равна такая вероятность? Ну,

$$\frac{n!}{k_1!k_2!\cdots k_m!}p_1^{k_1}p_2^{k_2}\cdots p_m^{k_m}$$

Замечание. Несложно заметить, что штука справа — слагаемые в сумме $(p_1 + p_2 + \dots + p_m)^n$.

Определение 14. Случайный вектор ξ (и его распределение) называется **абсолютно непрерывным**, если существует $p \in L(\mathbb{R}^n \to \mathbb{R})$ с неотрицательными значениями такая что $P(\xi \in B) = \int_B P \, \mathrm{d}\mu_n$ (тут интеграл n-кратный). В таком случае p называется **плотностью** ξ.

Пример. Случайный вектор ξ имеет стандартное многомерное нормальное распределение $\mathrm{N}(\mathbb{0}_n; E_n)$, если его плотность равна

$$p(x_1; \dots; x_n) = \frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{\sum\limits_{i=1}^n x_i^2}{2}\right)$$

Свойство 14.1. *Несложно заметить, что это произведение плотностей одномерных стандартных нормальных распределений.*

Пример. Случайный вектор η имеет многомерное нормальное распределение $\mathrm{N}(\mu; \Sigma)$ (где $\mu \in \mathbb{R}^n$, Σ — симметричная матрица $n \times n$ с неотрицательными собственными числами), если он равен $\sqrt{\Sigma}\xi + \mu$, где ξ — стандартный многомерный гауссовский вектор.

Замечание. В случае $\Sigma > 0$ можно написать плотность:

$$p(y) = \frac{1}{\sqrt{2\pi}^n \sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}(y-\mu)^T \Sigma^{-1} (y-\mu)\right)$$