Теория вероятностей и математическая статистика

С. М. Ананьевский *

1 ноября 2016 г.

Содержание

*Конспект подготовлен студентом Яскевичем С. В.

1 Условные вероятности

Пусть у нас есть вероятностное пространство $(\Omega, \mathfrak{F}, P)$ и два случайных события $A, B \in \mathfrak{F}$, причём будем считать, что $P(B) \neq 0$.

Определение 1.1. *Условной вероятностью события* A *при условии события* B называется число $P(A/B) = \frac{P(A \cap B)}{P(B)}$ (иногда обозначается $P_B(A)$).

Пример 1.2. Если у нас есть игральный кубик, то вероятность выпадения нечётной грани при условии, что количество очков не превосходит 3, равна

Теорема 1.3 (Свойства условной вероятности). *Условная вероятность является вероятностью.*

Доказательство. Проверим аксиомы вероятности:

1. $P_B(A) = \frac{P(A \cap B)}{P(B)}$. Так как числитель и знаменатель неотрицательны, то и дробь не отрицательна.
2. $P_B(\Omega) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$
3. Пусть $A_1, A_2 \ldots, \in \mathfrak{F}$; $A_i \cap A_j = \varnothing$ (при $i \neq j$). Тогда:

$$P_B(\bigcup_i A_i) = \frac{P(\bigcup_i B \cap A_i)}{P(B)} = \frac{\sum_i P(B \cap A_i)}{P(B)} =$$

$$= \sum_i \frac{P(B \cap A_i)}{P(B)} = \sum_i P_B(A_i)$$

□

Следствие 1.4. *Все свойства вероятности для условной вероятности выполнены.*

2 Формула полной вероятности. Формула Байеса

Определение 2.1. Пусть $A_1, A_2 \ldots, \in \mathfrak{F}$; $A_i \cap A_j = \varnothing$ (при $i \neq j$), $\bigcup_i A_i = \Omega$. Тогда $A_1, A_2 \ldots$ — полная система событий.

Теорема 2.2 (Формула полной вероятности). *Пусть* $A_1, A_2 \ldots$ *— полная система событий и* $\forall i \; P(A_i) > 0$. *Тогда вероятность любого случайного события* $B \in \mathfrak{F}$ *можно вычислить по формуле:*

$$P(B) = \sum_i P(B/A_i) \cdot P(A_i)$$

Доказательство.

$$P(B) = P(B \cap \Omega) = P(B \cap (\bigcup_i A_i)) = P(\bigcup_i (B \cap A_i)) = \sum_i P(B \cap A_i) = \sum_i \frac{P(B \cap A_i) \cdot P(A_i)}{P(A_i)} =$$

$$= \sum_i P(B/A_i) \cdot P(A_i)$$

□

Теорема 2.3 (Байеса). *Пусть* $A_1, A_2 \ldots$ *— полная система событий,* $\forall i\ P(A_i) > 0$*,* $P(B) > 0$*. Тогда* $\forall k \geqslant 1\ P(A_k/B) = \frac{P(B/A_k)\cdot P(A_k)}{\sum\limits_i P(B/A_i)\cdot P(A_i)}$

Доказательство. Правая часть равна:

$$\frac{P(B/A_k)\cdot P(A_k)}{P(B)} = \frac{P(B\cap A_k)\cdot P(A_k)}{P(A_k)\cdot P(B)} = \frac{P(B\cap A_k)}{P(B)} = P(A_k/B)$$

□

Примеры 2.4.

1. Пусть с завода №1 поставлено 5 ящиков деталей, с завода №2 — 3 ящика, а с завода №3 — 2 ящика. Предположим также, что завод №1 допускает 2% брака, завод №2 — 5% брака, а завод №3 — 10%. Какова вероятность выбрать хорошую деталь? Какова вероятность того, что деталь изготовлена на заводе №1 при условии, что она хорошая?

   Ящики считаем одинаковыми. Рассмотрим события C_1, C_2, C_3, где C_i означает выбрать ящик с завода №i, и событие B, означающее выбор хорошей детали. Ясно, что $C1, C_2, C_3$ — полная система событий. Чтобы вычислить вероятность события B, можно воспользоваться формулой полной вероятности:

   $$P(B) = \sum_{i=1}^{3} P(B/C_i)\cdot P(C_i) = 0.98\cdot 0.5 + 0.95\cdot 0.3 + 0.9\cdot 0.2$$

   Чтобы ответить на второй вопрос, мы можем воспользоваться формулой Байеса:

   $$P(C_1/B) = \frac{0.98\cdot 0.5}{0.98\cdot 0.5 + 0.95\cdot 0.3 + 0.9\cdot 0.2}$$

2. Представим, что у нас имеется ящик с шестью белыми и четырьмя чёрными шариками. Сначала мы потеряли один шарик из этого ящика (какой — неизвестно), а затем из оставшихся мы вытащили два шарика. Какова вероятность вытащить два белых шарика? Какова вероятность того, что был потерян чёрный шар, при условии, что мы вытащили два белых шара?

   Введём два события, описывающие первый этап эксперимента: $C_б, C_ч$ — потеря белого и чёрного шаров соответственно. $C_б, C_ч$ — полная система событий. Пусть B означает "вытащить два белых шарика".

   $$P(B) = P(B/C_б)\cdot P(C_б) + P(B/C_ч)\cdot P(C_ч) = \frac{C_5^2}{C_9^2}\cdot\frac{6}{10} + \frac{C_6^2}{C_9^2}\cdot\frac{4}{10}$$

   $$P(C_ч/B) = \frac{C_6^2\cdot 4}{C_5^2\cdot 6 + C_6^2\cdot 4}$$

3 Независимые события. Пример Бернштейна

Важно: нельзя путать понятия *независимости* событий и *несовместности.*

Пусть имеется эксперимент, описываемый с помощью вероятностного пространства $(\Omega, \mathfrak{F}, P)$, и даны случайные события $A, B \in \mathfrak{F}$.

Независимость событий можно было бы рассматривать как выполнение равенств

$$P(A/B) = P(A/\overline{B}) = P(A).$$

Однако здесь нарушена симметрия - логично, что если событие A независимо от B, то и обратное тоже верно — B независимо от A.

Определение 3.1. A и B независимы, если $P(A \cap B) = P(A)P(B)$.

Таким образом, если $P(B) > 0$, то независимость A и B равносильна $P(A/B) = P(A)$.

Предложение 3.2 (Свойства независимых событий). *1. A, B независимы $\iff$ A, $\overline{B}$ независимы $\iff$ $\overline{A}$, B независимы $\iff$ $\overline{A}$, $\overline{B}$ независимы.*

2. $\forall A \in \mathfrak{F}$ A и Ω независимы.

3. $\forall A \in \mathfrak{F}$ A и $\varnothing$ независимы.

Доказательство.

$$P(A) = P(A \cap \Omega) = P(A \cap (B \cup \overline{B})) = P((A \cap B) \cup (A \cap \overline{B}))$$

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

$\square$

Упражнение 3.3. *Пусть A и B независимы, A и C независимы. Верно ли, что A и $B \cup C$ независимы? Верно ли, что A и $B \cap C$ независимы?*

Можем ли мы определить понятие независимости для числа событий, большего 2? Для событий $A_1, A_2, \ldots, A_n$ мы можем выделить попарную независимость:

$$\forall i \neq j \quad A_i, A_j \text{ независимы.}$$

Или же независимость в совокупности (совместную): $A_1, \ldots, A_n$ независимы в совокупности, если:

1. $\forall i \neq j$ A_i, A_j — независимы;
2. $\forall i_1 < i_2 < i_3 s$ $P(\bigcap_{j=1}^{3} A_{ij}) = \prod_{j=1}^{3} P(A_{ij})$
3. $P(\bigcap_{j=1}^{n} A_{ij}) = \prod_{j=1}^{n} P(A_{ij})$ и так далее.

Это равносильно:

$$\forall 2 \leqslant k \leqslant n \quad \forall i_1 < i_2 < \ldots < i_k \quad P(\bigcap_{j=1}^{k} A_{ij}) = \prod_{j=1}^{k} P(A_{ij})$$

Пример 3.4 (Берштейна). Рассмотрим эксперимент: будем подбрасывать тетраэдр с белой, синей, красной и разноцветной (бело-сине-красной) гранями. Введём три события:

$$Б = \{\text{внизу присутствует белый цвет}\}$$

$$С = \{\text{внизу присутствует синий цвет}\}$$

$$К = \{\text{внизу присутствует красный цвет}\}$$

Проверим, что эти события попарно независимы. Верно ли, что $P(Б \cap С) = P(Б)P(С)$? Очевидно, что да. Значит, попарная независимость есть. Проверим теперь совместную независимость:

$$P(Б \cap С \cap К) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(Б)P(С)P(К)$$

Это доказывает, что попарная независимость и совместная независимость неравносильны.

4 Независимые испытания Бернулли. Формулы Бернулли

Пусть у нас есть вероятностное пространство $(\Omega, \mathfrak{F}, P)$. Рассмотрим набор случайных событий $\mathfrak{A} = (A_1, \dots, A_n)$, причём $A_1, \dots, A_n$ — полная системы событий. Определим *испытание* $A_1, \dots, A_m$ как набор событий, являющийся полной системой событий.

Определение 4.1. Испытания $A_1, \dots, A_m$ будем называть *независимыми*, если для любого набора $A_{1i_1}, A_{2i_2} \dots, A_{mi_m}$ составляющие его события являются совместно независимыми.

Пример 4.2. Представим, что мы одновременно подбрасываем монетку и кубик. Каким будет вероятностное пространство? $\Omega = \{O1, O2, \dots, O6, P1, \dots, P6\}$ Зададим испытании $A_1 = \{A_{11}, A_{12}\}$, где $A_{11} = \{\text{на монете O}\}$, $A_{12} = \{\text{на монете P}\}$; $A_2 = \{A_{21}, \dots, A_{26}\}$, где $A_{2i} = \{\text{на кубике цифра } i\}$.

$$P(A_{11} \cap P(A_{23}) = \frac{1}{12}, \quad P(A_{11}) = \frac{1}{2} \quad P(A_{23}) = \frac{1}{2}$$

Испытаниями Бернулли называются набор из n независимых испытаний с двумя исходами в каждом из них, условно называемыми успехом и неудачей, и с постоянной вероятностью успеха во всех испытаниях. Будем обозначать такой набор $(У_1, У_2, Н_3, \dots, У_n$, а вероятность успеха $P(У_k) = p$.

Пусть $A = \{\text{в } n \text{ испытаниях Бернулли успех произошёл ровно } k \text{ раз}\}$. Какова вероятность A? Заметим, что $A = (\underbrace{УУ\dots У}_{k}\underbrace{НН\dots Н}_{n-k}) \cup (\underbrace{УУ\dots УУ}_{k-1}НУНН\dots Н) \cup \dots$.

$$P(УУ\dots УНН\dots Н) = P(У_1 \cap У_2 \cap \dots \cap У_k \cap Н_{k+1} \cap \dots \cap Н_n) = p^k(1-p)^{n-k}$$

Ясно, что вероятность любой цепочки, содержащей k успехов и $n-k$ неудач, равна $p^k(1-p)^{n-k}$

Получаем, что в нашем примере $P(A) = C_n^k p^k(1-p)^{n-k} = P_n(k)$ — эта формула носит имя Бернулли.

Следствие 4.3.

1. $P_n(n) = p^n$
2. $P_n(0) = (1-p)^n$
3. $P_n(\text{хотя бы один успех}) = 1 - P_n(0) = 1 - (1-p)^n$

Пример 4.4. Пусть мы подбрасываем монету 10 раз. Какова вероятность того, что все десять раз выпал орёл? По формуле Бернулли:

$$P_{10}(10) = (\frac{1}{2})^{10} = \frac{1}{1024}$$

А вероятность того, что орёл выпал ровно пять раз, равна

$$P_{10}(5) = C_{10}^5(\frac{1}{2})^5(\frac{1}{2})^5 = \frac{252}{1024}$$

Возникает вопрос, каково наиболее вероятное число выпадений орла? $P_{10}(k) = C_{10}^k(12)^{10}$, и ясно, что максимум достигается при $k = 5$

Обобщим последний вопрос примера. Пусть имеется n испытаний Бернулли. Вероятность успеха в каждом испытании равна p. Чему равно наиболее вероятное число появления успехов?

Рассмотрим неравенство:

$$P_n(k) < P_n(k+1)$$

Его можно переписать:

$$C_n^k p^k (1-p)^{n-k} < C_n^{+1-1} p^{k+1} (1-p)^{n-k-1}$$

$$\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} < \frac{n!}{(k+1)!(n-k-1)!} p^{k+1} (1-p)^{n-k-1}$$

Полученное равносильно:

$$(k+1)(1-p) < p(n-k)$$

$$k < (n+1)p - 1$$

Если мы теперь посмотрим на обратное неравенство $P_n(k) > P_n(k+1)$, то увидим, что оно равносильно $k > (n+1)p - 1$.

Рассмотрим два случая:

1. $(n+1)p \notin \mathbb{Z}$. Обозначим за k_n^* наиболее вероятное число успехов в n испытаниях. Тогда $k_n^* = [(n+1)p]$
2. $(n+1)p \in \mathbb{Z}$. В этому случае $k_n^{*_1} = (n+1)p - 1$ и $k_n^{*_2} = (n+1)p$ — наиболее вероятные числа успехов.

Пример 4.5. Пусть $p = \frac{1}{2}$, $n = 10$. Тогда $k_{10}^* = 5$. Если же $n = 11$, то $k_{11}^{*_1} = 5$ и $k_{11}^{*_2} = 6$, так как $C_{11}^5 = C_{11}^6$.

5 Предельные теоремы в схеме испытаний Бернулли

Представим, что мы подбрасываем монету 10000 раз. Ясно, что наиболее вероятное число выпадений орла равна 5000. Чему же равна вероятность такого исхода?

$P_{10000}(5000) = C_{10000}^{5000}(\frac{1}{2})^{10000}$. Мы хотели бы оценить это число.

Рассмотрим функции:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \Phi(x) = \int\limits_{-\infty}^{x} \varphi(t)dt$$

Ясно, что $\varphi(x)$ — чётная, а $\Phi(-x) = 1 - \Phi(x)$.

Пусть n — число испытаний, k — число успехов, p — вероятность успеха, $q = 1 - p$. Введём обозначение: $x_{n,k} = \frac{k-np}{\sqrt{npq}}$.

Теорема 5.1 (Локальная теорема Муавра-Лапласа). *Справедливо следующее соотношение:*

$$\frac{P_n(k)}{\frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x_{n,k}^2}{2}}} \xrightarrow[n\to\infty]{} 1 \quad \text{равномерно по всем } k: \quad |x_{n,k}| \leqslant Cn^{\frac{1}{6}-\varepsilon} \quad \forall C > 0,\ \varepsilon > 0$$

Таким образом, $P_n(k) \approx \frac{1}{\sqrt{npq}} \varphi(x_{n,k})$.

Лемма 5.2 (Формула Стирлинга).

$$n! = n^n e^{-n} \sqrt{2\pi n} \cdot (1 + o(1))\ (n \to \infty)$$

Лемма 5.3.

$$\ln(1+x) = x - \frac{x^2}{2} + \theta x^3, , \text{ где } |\theta| \leqslant 3 \quad \forall |x| < \frac{1}{2}$$

Доказательство теоремы. По формуле Бернулли:

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k q^{n-k} = \frac{n^n e^{-n}\sqrt{2\pi n}(1+o(1))}{k^k e^{-k}\sqrt{2\pi k}(1+o(1))(n-k)^{n-k}e^{-n+k}\sqrt{2\pi(n-k)}(1+o(1))}$$

Но это верно при $b \to \infty$, $k \to \infty$, $n-k \to \infty$.

$$k = np + x_{n,k}\sqrt{npq} \xrightarrow[n\to\infty]{} \infty$$

$$n - k = nq - x_{n,k}\sqrt{npq} \xrightarrow[n\to\infty]{} \infty$$

Продолжая вычисления:

$$P_n(k) = \left(\frac{k}{np}\right)^{-k-\frac{1}{2}} \left(\frac{n-k}{nq}\right)^{-n+k-\frac{1}{2}} \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}}(1+o(1))$$

Пусть $\sqrt{npq}P_n(k) = T_{n,k}$. Тогда:

$$T_{n,k} = \left(\frac{k}{np}\right)^{-k-\frac{1}{2}} \left(\frac{n-k}{nq}\right)^{-n+k-\frac{1}{2}} \frac{1}{\sqrt{2\pi}}(1+o(1))$$

$$\ln T_{n,k} = (-k - \frac{1}{2}) \ln \frac{k}{np} + (-n + k - \frac{1}{2}) \ln \frac{n-k}{nq} = \ln \frac{1}{\sqrt{2\pi}} + o(1)$$

С учётом

$$\frac{k}{np} = 1 + x\sqrt{\frac{q}{np}}$$

$$\frac{n-k}{nq} = 1 - x\sqrt{\frac{p}{nq}}$$

Применив лемму 5.3, получим:

$$\ln T_{n,k} = (-np - x\sqrt{npq} - \frac{1}{2})(x\sqrt{\frac{q}{np}} - \frac{x^2}{2}\frac{q}{np} + \theta_1 \frac{x^3 q\sqrt{q}}{np\sqrt{np}})+$$

$$+(-np + x\sqrt{npq} - \frac{1}{2})(-x\sqrt{\frac{p}{nq}} - \frac{x^2 p}{2nq} + \theta_2 x^3 \frac{p}{nq}\frac{\sqrt{p}}{\sqrt{nq}}) + \ln\frac{1}{\sqrt{2\pi}} + o(1) =$$

$$= -x\sqrt{npq} + \frac{x^2}{2}q + o(1) - x^2 q + x\sqrt{npq} + \frac{x^2}{2}p - x^2 p + \ln\frac{1}{\sqrt{2\pi}} =$$

$$= -\frac{x^2}{2} + \ln\frac{1}{\sqrt{2\pi}} + o(1)$$

Тогда само $T_{n,k}$ равно:

$$T_{n,k} = e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{o(1)}$$

Отсюда:

$$P_n(k) = \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_{n,k}^2}{2}} e^{o(1)}$$

Теорема доказана. $\square$

Попробуем ответить на вопрос: как найти $P_n(a < \text{число успехов} < b)$ при $a < b$ и больших n? Применение локальной теоремы Муавра-Лапласа может давать слишком высокие погрешности, поэтому необходимо использовать иное решение.

Теорема 5.4 (Интегральная теорема Муавра-Лапласа). *Пусть* p *— вероятность успеха,* $0 < p < 1$. *Тогда:*

$$\sup_{a<b}\left|P_n(a < \textit{число успехов} < b) - \int\limits_{\frac{a-np}{\sqrt{npq}}}^{\frac{b-np}{\sqrt{npq}}} \varphi(t)\,dt\right| \xrightarrow[n\to\infty]{} 0$$

Мы докажем эту теорему позже — как частный случай более общей теоремы.

Следствие 5.5.

$$\int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt = 1$$

Доказательство. Примем $a = -\infty$, $b = +\infty$. ☐

Если мы возьмём функцию $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt$, то теорему можно сформулировать так:

$$P_n(a < \text{число успехов} < b) \approx \Phi\left(\frac{b-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-np}{\sqrt{npq}}\right)$$

Пример 5.6. Подсчитаем вероятность того, что при подбрасывании монетки 10000 раз «орёл» выпадет 5000 раз. Для этого применим локальную теорему Муавра-Лапласа.

$$P_{10000}(5000) \approx \frac{1}{\sqrt{10000 \cdot \frac{1}{2} \cdot \frac{1}{2}}}\varphi(0) = \frac{1}{50} \cdot 0.39894$$

Если мы захотим подсчитать вероятность того, что «орёл» выпадет как минимум 4900 раз и как максимум 5100, то необходимо будет применить интегральную теорему:

$$P_{10000}(4900 < \text{число успехов} < 5100) \approx \ldots$$

Теорема 5.7 (Пуассона). *Будем рассматривать схему серий испытаний Бернулли. Допустим, первая серия состоит из одного испытания такого, что* $p_1 = P_1(\text{У})$. *Вторая серия состоит из двух испытаний и* $p_2 = P_1(\text{У})$. n*-я серия испытаний состоит из* n *испытаний и* $p_n = P_1(\text{У})$. *(Здесь* $P_1(\text{У})$ *— вероятность успеха в одном испытании для каждой серии соответственно.) Пусть также* $np_n = \lambda > 0$. *Тогда:*

$$P_n(k) \xrightarrow[n\to\infty]{} \frac{\lambda^k}{k!}e^{-\lambda}$$

Доказательство.

$$P_n(k) = C_n^k p_n^k (1-p_n)^{n-k} = \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^{n-k} =$$

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

$$= \frac{\lambda^k}{k!} \underbrace{(1-\frac{1}{n})\ldots(1-\frac{k-1}{n})}_{\to 1} \cdot \underbrace{(1-\frac{\lambda}{n})^n}_{\to e^{-\lambda}} \cdot \underbrace{(1-\frac{\lambda}{n})^{-k}}_{\to 1}$$

☐

Теорема 5.8 (Закон больших чисел Бернулли).

$$\forall \varepsilon > 0 \quad P_n\left(\left|\frac{k_n}{n} - p\right| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0$$

Здесь $\frac{k_n}{n}$ называется частотой успеха.

Доказательство. Рассмотрим вероятность:

$$P_n(\left|\frac{k_n}{n} - p\right| > \varepsilon) = 1 - P_n(\left|\frac{k_n}{n} - p\right| \leqslant \varepsilon) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\, dt - P_n(np - n\varepsilon \leqslant k_n \leqslant n\varepsilon + np) =$$

$$= \frac{1}{\sqrt{2\pi}} \int\limits_{|t|\leqslant \varepsilon\sqrt{\frac{n}{pq}}} e^{-\frac{t^2}{2}}\, dt + \frac{1}{\sqrt{2\pi}} \int\limits_{|t|> \varepsilon\sqrt{\frac{n}{pq}}} e^{-\frac{t^2}{2}}\, dt - P_n(np - n\varepsilon \leqslant k_n \leqslant n\varepsilon + np)$$

Разность первого и третьего слагаемого стремится к нулю, второе слагаемое тоже стремится к нулю. Теорема доказана. □

6 Случайная величина. Распределение случайных величин

Будем рассматривать вероятностное пространство $(\Omega, \mathfrak{F}, P)$.

Определение 6.1. Функция $\xi : \Omega \to \mathbb{R}$ такая, что $\forall B \in \mathfrak{B}\ \xi^{-1}(B) \in \mathfrak{F}$, называется *случайной величиной.* (Здесь $\mathfrak{B}$ обозначает борелевскую сигма-алгебру.)

Примеры 6.2.

1. Пусть мы подбрасываем игральный кубик. $\Omega = \{\omega_1, \dots, \omega_6\}$. Пусть $\xi_1(\omega_i) = i$, $\xi_2(\omega_i) = \begin{cases} 0, & i \neq 6 \\ 1, & i = 6 \end{cases}$. ξ_1, ξ_2 — случайные величины.
2. Пусть $\Omega = \{x \big| x \in [0, 1]\}$, $\mathfrak{F} = \{[0, \frac{1}{2}], (\frac{1}{2}, 1], [0, 1], \varnothing\}$, $\eta(x) = x$. η не будет случайной величиной, так как $\eta^{-1}([0, \frac{1}{3}]) = [0, \frac{1}{3}] \notin \mathfrak{F}$.

Рассмотрим функцию $\mathfrak{P}_\xi : \mathfrak{B} \to \mathbb{R}$ такую, что $\forall B \in \mathfrak{B}\ \mathfrak{P}_\xi(B) = P(\xi^{-1}(B))$.

Определение 6.3. $\mathfrak{P}_\xi$ называется *распределением случайной величины* ξ.

Теорема 6.4. *$\mathfrak{P}_\xi$ является вероятностью.*

Доказательство. Проверим аксиомы вероятности.

1. $\mathfrak{P}_\xi(B) = P(\xi^{-1}(B)) \geqslant 0$
2. $\mathfrak{P}_\xi(\mathbb{R}) = P(\xi^{-1}(\mathbb{R})) = P(\Omega) = 1$
3. Пусть $B_1, B_2, \ldots \in B$, $B_i \cap B_j = \varnothing$ $(i \neq j)$.

$$\mathfrak{P}_\xi(\bigcup_{i=1}^{\infty} B_i) = P(\xi^{-1}(\bigcup_{i=1}^{\infty} B_i) =$$

$$= P(\bigcup_{i=1}^{\infty} \underbrace{\xi^{-1}(B_i)}_{\text{попарно несовместны}}) = \sum_i P(\xi^{-1}(B_i)) = \sum_i \mathfrak{P}_\xi(B_i)$$

□

Теперь мы можем перейти к использованию вероятностного пространства $(\mathbb{R}, \mathfrak{B}, \mathfrak{P}_\xi)$.

Определение 6.5. Функция $F_\xi : \mathbb{R} \to \mathbb{R}$ такая, что $\forall x \in \mathbb{R}\ F_\xi(x) = \mathfrak{P}_\xi((-\infty, x)) = P(\xi^{-1}((-\infty, x))) = P(\omega : \xi(\omega) < x) = P(\xi < x)$, называется *функцией распределения случайной величины* ξ.

Рассмотрим значение $F_\xi(b) = \mathfrak{P}_\xi((-\infty, b)) = \mathfrak{P}_\xi((-\infty, a) \cup [a, b)) = \mathfrak{P}_\xi((-\infty, a)) + \mathfrak{P}_\xi([a, b)) = F_\xi(a) + \mathfrak{P}_\xi([a, b)))$. Получается, что $\mathfrak{P}_\xi([a, b)) = F_\xi(b) - F_\xi(a)$, то есть, существует взаимно-однозначное соответствие между $\mathfrak{P}_\xi$ и F_ξ.

Предложение 6.6 (Свойства функции распределения). *1.* $\forall x\ 0 \leqslant F_\xi(x) \leqslant 1$*;*

2. F_ξ *неубывает;*

3. F_ξ *непрерывна слева во всех точках.*

4. $F_\xi(x) \xrightarrow[x \to -\infty]{} 0$, $F_\xi(x) \xrightarrow[x \to \infty]{} 1$

Доказательство. Докажем свойство 3. Пусть $x_1 < x_2 < \ldots$ и $x_n \xrightarrow[n \to \infty]{} x$. Мы хотим доказать, что $F_\xi(x_n) \xrightarrow[n \to \infty]{} F_\xi(x)$. $(-\infty, x_n) \subset (-\infty, x_{n+1})$ и $\bigcup\limits_n (-\infty, x_n) = (-\infty, x)$. Отсюда, по свойству вероятности, $\underbrace{\mathfrak{P}_\xi((-\infty, x_n))}_{F_\xi(x_n)} \xrightarrow[n \to \infty]{} \underbrace{\mathfrak{P}_\xi((-\infty, x))}_{F_\xi(x)}$. Всё доказано.

Докажем свойство 4. Пусть $\forall n \geqslant 1\ [-n, n) \subset [-(n+1), n+1)$ и $\bigcup\limits_{n \geqslant 1} [-n, n) = \mathbb{R}$. Тогда:

$$\underbrace{\mathfrak{P}_\xi([-n, n))}_{F_\xi(n) - F_\xi(-n)} \xrightarrow[n \to \infty]{} \mathfrak{P}_\xi(\mathbb{R}) = 1$$

Отсюда получаем, что $\lim\limits_{n \to \infty} F_\xi(n) = 1$ и $\lim\limits_{n \to \infty} F_\xi(-n) = 0$. Свойство 4 доказано. □

Мы доказали, что функция распределения удовлетворяет указанным свойствам, однако верно и обратное.

Предложение 6.7. *Пусть* $G : \mathbb{R} \to \mathbb{R}$ *— функция, удовлетворяющая свойствам функции распределения. Тогда* G *является функцией распределения некоторой случайной величины.*

Доказательство. Упражнение. □

7 Различные типы распределений случайных величин. Случайные величины с дискретным распределением

Определение 7.1. Случайная величина ξ называется случайной величиной с дискретным распределением, если существует не более чем счётное подмножество вещественных чисел A такое, что $P(\xi \in A) = 1$. Другими словами, у этой случайной величины конечное или счётное число значений.

Занумеруем все элементы множества A: $A = \{a_1, a_2, \ldots\}$, обозначим $p_i = P(\xi = a_i)$. Правило, сопоставляющее каждому значению a_i случайной величины ξ вероятность p_i, называется *законом распределения.*

Примеры 7.2. 1. $\xi = c$, $p = 1$ — случайная величина с вырожденной в точке c распределением. График функции распределения $F_\xi(x) = P(\xi < x)$ представляет собой ступенчатую функцию, принимающую значение 0 при $x \leqslant c$ и 1 при $x > c$.

2. $(\xi, p) : (0, 1-p), (1, p)$ — случайная величина с распределением Бернулли. График — ступенчатая функция, принимающая значения $1-p$ и p в точках 0 и 1 соответственно.

3. $(\xi, p) : (a_1, p_1), (a_2, p_2), \ldots, (a_n, p_n)$. График — аналогичная ступенчатая функция.

4. $\xi : 0, 1, 2, \ldots, n$ и $p_k = P(\xi = k) = C_n^k p^k (1-p)^{n-k}$ — случайная величина с биномиальным законом распределения и параметрами n и p.

5. $\xi : 0, 1, 2, \ldots$ и $P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ — случайная величина с распределением Пуассона с параметром $\lambda > 0$.

6. $\xi : 0, 1, 2, \ldots$ и $P(\xi = k) = p^k(1-p)$ — случайная величина с геометрическим распределением с параметром p $(0 \leqslant p \leqslant 1)$.

Определение 7.3. Случайная величина ξ называется случайной величиной с абсолютно непрерывным распределением, если существует функция $p(x)$ такая, что для любого $y \in \mathbb{R}$ $F_\xi(y) = \int\limits_{-\infty}^{y} p(x)\,dx$. В этом случае $p(x) = p_\xi(x)$ называется *плотностью распределения случайной величины* ξ.

Предложение 7.4 (Свойства плотности распределения). *1.* $p(x) \geqslant 0$.

2. $\int\limits_{-\infty}^{+\infty} p(x)\,dx = 1$.

Примеры 7.5. 1. ξ — случайная величина с равномерным распределением на отрезке $[a, b]$. (Обозначается $\xi \in U_{[a,b]}$).

$$p_\xi(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b] \end{cases}$$

2. ξ — случайная величина с экспоненциальным распределением с параметром α.

$$p_\xi(x) = \begin{cases} 0, & x < 0 \\ \alpha e^{-\alpha x}, & x \geqslant 0 \end{cases}$$

3. (Важный пример!) ξ — случайная величина с нормальным распределением с параметрами $a \in \mathbb{R}$ и $\sigma^2 > 0$.

$$p_\xi = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

В случае, когда $a \neq 0$, $\sigma^2 = 1$, говорят о *стандартном нормальном распределении.* В этом случае $p_\xi = \varphi$ — функция Гаусса.

4. ξ — случайная величина с распределением Коши. $p_\xi(x) = \frac{1}{\pi(1+x^2)}$.

Заметим, что $P(a < \xi < b) = P(a \leqslant \xi < b) = P(a \leqslant \xi \leqslant b) = F_\xi(b) - F_\xi(a)$. Почему строгость неравенства не имеет значения? В силу непрерывности функции распределения $P(a < \xi < b) = F_\xi(b) - F_\xi(a) \xrightarrow[b \to a]{} 0$.

8 Случайные векторы. Распределение случайных векторов

Определение 8.1. Пусть $\xi_1, \ldots, \xi_d$ — случайные величины, определённые на одном вероятностном пространстве. Тогда $\xi = (\xi_1, \ldots, \xi_d)$ будет называться случайным d-мерным вектором, $\xi : \Omega \to \mathbb{R}^d$. Иначе говоря, $\forall B \in \mathfrak{B}^d \ \xi^{-1}(B) \in \mathfrak{F}$.

Определение 8.2. Функция $F_{(\xi_1,\ldots,\xi_d)} : \mathbb{R}^d \to \mathbb{R}$ такая, что $\forall x_1, \ldots, x_d \ F_{(\xi_1,\ldots,\xi_d)}(x_1, \ldots, x_d) = P(\xi_1 < x_1, \ldots, \xi_d < x_d)$, называется функцией распределения случайного вектора.

Рассмотрим случай $d = 2$, $F_{(\xi_1,\xi_2)}(x_1, x_2) = P(\xi_1 < x_1, \xi_2 < x_2)$.

$\xi_1, \ldots, \xi_d$ — с дискретным распределением, если существует $A \subset \mathbb{R}^d$ такое, что A не более чем счётно и $P((\xi_1, \ldots, \xi_d) \in A) = 1$.

$\xi_1, \ldots, \xi_d$ — с абсолютно непрерывным распределением, если существует $p(x_1, \ldots, x_d)$ такое, что $F_{(\xi_1,\ldots,\xi_d)}(y_1, \ldots, y_d) = \int\limits_{-\infty}^{y_1} \ldots \int\limits_{-\infty}^{y_d} p(x_1, \ldots, x_d)\, dx_1 \ldots dx_d$.

Предложение 8.3 (Свойства функции распределения случайного вектора).

1. *$0 \leqslant F_{(\xi_1,\ldots,\xi_d)}(x_1, \ldots, x_d) \leqslant 1$*
2. *Пусть $x_1 \leqslant y_1, \ldots, x_d \leqslant y_d$. Тогда $F_{(\xi_1,\ldots,\xi_d)}(x_1, \ldots, x_d) \leqslant F_{(\xi_1,\ldots,\xi_d)}(y_1, \ldots, y_d)$*
3. *$F_{(\xi_1,\ldots,\xi_d)}$ непрерывна слева по каждой координате.*
4. *$F_{(\xi_1,\ldots,\xi_d)}(x_1, \ldots, x_d) \xrightarrow[x_k \to -\infty]{} 0$.*
5. *$F_{(\xi_1,\ldots,\xi_d)}(x_1, \ldots, x_d) \xrightarrow[x_k \to +\infty]{} F_{(\xi_1,\ldots,\xi_{k-1},\xi_{k+1},\ldots\xi_d)}(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots x_d)$.*
6. *$F_{(\xi_1,\ldots,\xi_d)}(x_1, \ldots, x_d) \xrightarrow[x_1 \to +\infty, \ldots, x_d \to +\infty]{} 1$*
7. *Пусть $d = 2$. Тогда $P\big((\xi_1, \xi_2) \in [a_1, b_1) \times [a_2, b_2)\big) = F_{(\xi_1,\xi_2)}(b_1, b_2) - F_{(\xi_1,\xi_2)}(a_1, b_2) - F_{(\xi_1,\xi_2)}(b_1, a_2) + F_{(\xi_1,\xi_2)}(a_1, a_2)$. То есть, вероятность попадания точки в произвольный параллелепипед однозначно выражается через значения функции распределения в вершинах этого параллелепипеда.*

9 Независимость случайных величин

Вспомним, что события $A, B \in \mathfrak{F}$ называются независимыми, если $P(A \cap B) = P(A)P(B)$. Введём теперь понятие независимости для случайных величин.

Определение 9.1. Пусть ξ, η — случайные величины, определённые на одном вероятностном пространстве. Будем говорить, что ξ и η — независимые случайные величины, если для любых вещественных x, y $F_{(\xi,\eta)}(x, y) = F_\xi(x) F_\eta(y)$.

Определение 9.2. ξ и η — независимые случайные величины, если $\forall a_1 < b_1, a_2 < b_2$ $P\big((\xi, \eta) \in [a_1, b_1) \times [a_2, b_2)\big) = P(a_1 \leqslant \xi < b_1) \cdot P(a_2 \leqslant \eta < b_2)$.

Предложение 9.3. *Определения 9.1 и 9.2 эквивалентны.*

Доказательство.

1. $(9.1 \Rightarrow 9.2)$

$$P\big((\xi, \eta) \in [a_1, b_1) \times [a_2, b_2)\big) = P(a_1 \leqslant \xi < b_1) \cdot P(a_2 \leqslant \eta < b_2)$$

При $a_1, a_2 \to -\infty$ левая часть равенства стремится к $F_{(\xi,\eta)}(b_1, b_2)$, а правая часть — к $F_\xi(b_1) \cdot F_\eta(b_2)$.

2. (9.2 ⇒ 9.1)

$$P\big((\xi,\eta)\in[a_1,b_1)\times[a_2,b_2)\big)=F_{(\xi,\eta)}(b_1,b_2)-F_{(\xi,\eta)}(a_1,b_2)-F_{(\xi,\eta)}(b_1,b_a)+$$

$$+F_{(\xi,\eta)})(a_1,a_2)=F_\xi(b_1)F_\eta(b_2)-F_\xi(a_1)F_\eta(b_2)-F_\xi(b_1)F_\eta(a_2)+$$

$$+F_\xi(a_1)F_\eta(a_2)=\big(F_\xi(b_1)-F_\xi(a_1)\big)\big(F_\eta(b_2)-F_\eta(a_2)\big)=$$

$$P(a_1\leqslant\xi<b_1)\cdot P(a_2\leqslant\eta<b_2)$$

□

Определение 9.4. Пусть для любых $B_1,\dots,B_d\in\mathfrak{B}$ $P(\xi_1\in B_1,\dots,\xi_d\in B_d)=\prod\limits_{i=1}^{d}P(\xi_i\in B_i)$. Тогда $\xi_1,\dots,\xi_d$ называются независимыми случайными величинами.

Очевидно, что это определение также эквивалентно двум предыдущим.

Пример 9.5. Рассмотрим случайный вектор $\xi=(\xi_1,\xi_2)$ с плотностью распределения

$$p_\xi(x,y)=\begin{cases}1, & 0\leqslant x\leqslant 1,\ 0\leqslant y\leqslant 1\\ 0 & \text{иначе}\end{cases}$$

Здесь:

$$F_{\xi_1}(x)=P(\xi_1<x)=\begin{cases}0, & x\leqslant 0\\ x, & 0<x\leqslant 1\\ 1, & x>1\end{cases}$$

$$F_{\xi_2}(y)=P(\xi_2<y)=\begin{cases}0, & y\leqslant 0\\ y, & 0<y\leqslant 1\\ 1, & y>1\end{cases}$$

$$F_{(\xi_1,\xi_2)}(x,y)=\begin{cases}0, & x\leqslant 0 \text{ или } y\leqslant 0\\ xy, & 0<x\leqslant 1 \text{ и } 0<y\leqslant 1\\ x, & 0<x\leqslant 1 \text{ и } y>1\\ y, & 0<y\leqslant 1 \text{ и } x>1\\ 1, & x>1 \text{ и } y>1\end{cases}$$

Видим, что $F_{(\xi_1,\xi_2)}(x,y)=F_{\xi_1}(x)F_{\xi_2}(y)$, значит, ξ_1 и ξ_2 независимы.

Заметим, что для того, чтобы доказать, что события не являются независимыми, достаточно найти хотя бы одну точку, в которой определение нарушается.

Мы хотим найти более простой критерий независимости событий.

Рассмотрим для начала случай дискретного распределения, ограничимся двумерным случаем.

Предложение 9.6. *$\xi=(\xi_1,\xi_2)$ — случайный вектор с дискретным распределением тогда и только тогда, когда его компоненты — с дискретным распределением.*

Доказательство. Очевидно. □

Теорема 9.7. *Пусть $\xi=(\xi_1,\xi_2)$ — с дискретным распределением. Тогда независимость случайных величин ξ_1 и ξ_2 равносильна тому, что для любых вещественных a и b выполняется $P(\xi_1=a,\xi_2=b)=P(\xi_1=a)P(\xi_2=b)$.*

Доказательство. Если ξ_1, ξ_2 независимы, то $\forall a, b \in \mathbb{R}$ и $\forall \varepsilon > 0$ $P(a \leqslant \xi_1 \leqslant a+\varepsilon, b \leqslant \xi_2 < b-\varepsilon) = P(a \leqslant \xi_1 \leqslant a+\varepsilon)P(b \leqslant \xi_2 < b-\varepsilon)$ — по определению 9.2. Устремим ε к нулю.

Обратно:

$$F_{(\xi_1,\xi_2)}(y_1,y_2) = \sum_{a_i<y_1, b_j<y_2} P(\xi_1 = a_i, \xi_2 = b_j) = \sum_{a_i<y_1, b_j<y_2} P(\xi_1 = a_i)P(\xi_2 = b_j) =$$

$$= \sum_{b_j<y_2}\left(\sum_{a_i<y_1} P(\xi_1 = a_i)P(\xi_2 = b_j)\right) = \sum_{a_i<y_1} P(\xi_1 = a_i)\cdot \sum_{b_j<y_2} P(\xi_2 = b_j) = F_{\xi_1}(y_1)F_{\xi_2}(y_2)$$

□

Теперь рассмотрим случай непрерывного распределения.

Предложение 9.8. *Если $\xi = (\xi_1, \dots, \xi_d)$ — с абсолютным непрерывным распределением, то $\forall k = 1, 2, \dots, d$ ξ_k с абсолютно непрерывным распределением.*

Заметим, что обратное, вообще говоря, неверно.

Таким образом, ξ — случайный вектор с абсолютно непрерывным распределением тогда и только тогда, когда для любых $y_1, \dots, y_d$

$$F_\xi(y_1, \dots, y_d) = \int_{-\infty}^{y_1} \dots \int_{-\infty}^{y_d} p_\xi(x_1, \dots, x_d)\, dx_1 \dots dx_d$$

При $y_2, \dots, y_d \to \infty$ левая часть стремится к $F_{\xi_1}(y_1)$, а в правой будет

$$\int_{-\infty}^{y_1} \left(\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} p_\xi(x_1, \dots, x_d)\, dx_2 \dots dx_d\right) dx_1$$

Что по определению равно $p_{\xi_1}(x_1)$ — плотности распределения ξ_1.

Помимо этого, $\xi = (\xi_1, \dots, \xi_d)$ — с абсолютно непрерывным распределением $\iff$ для любых $a_1 < b_1, \dots, a_d < b_d$ выполнено

$$P(a_1 \leqslant \xi_1 < b_1, \dots, a_d \leqslant \xi_d < b_d = \int_{a_1}^{b_1} \dots \int_{a_d}^{b_d} p_\xi(x_1, \dots, x_d)\, dx_1 \dots dx_d$$

Что также равносильно

$$\forall B \in \mathfrak{B}^d \quad P(\xi \in B) = \int_B p_\xi(x_1, \dots, x_d)\, dx_1 \dots dx_d$$

Как нам проверить, будут ли случайные величины, составляющие случайный вектор с абсолютно непрерывным распределением, независимыми?

Теорема 9.9. *Пусть $\xi = (\xi_1, \xi_2)$ — с абсолютно непрерывным распределением и плотностью $p_\xi(x,y)$. Тогда независимость ξ_1, ξ_2 равносильна тому, что $p_\xi(x,y) = p_{\xi_1}(x)p_{\xi_2}(y)$.*

Доказательство. Проверим необходимость. Рассмотрим функцию распределения в произвольной точке: $F_{(\xi_1,\xi_2)}(x,y) = F_{\xi_1}(x)F_{\xi_2}(y)$. Дважды продифференцировав это равенство, получим

$$\frac{d^2F_{(\xi_1,\xi_2)}(x,y)}{dx\,dy} = p_\xi(x,y)$$

Левая часть равна $\frac{d^2 F_{\xi_1}(x) F_{\xi_2}(y)}{dx\,dy} = \frac{dF_{\xi_1}(x)}{dx} \frac{dF_{\xi_2}(y)}{dy} = p_{\xi_1}(x) p_{\xi_2}(y)$.

Обратно. Для любых x, y

$$\int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} p_\xi(s,t)\,ds\,dt = \int\limits_{-\infty}^{x} p_{\xi_1}(s)\,ds \int\limits_{-\infty}^{y} p_{\xi_2}(t)\,dt = F_{\xi_1}(x) F_{\xi_2}(y)$$

$\square$

Теорема 9.10. *Пусть ξ, η — независимые случайные величины, f, g — измеримые функции. Тогда $f(\xi)$ и $g(\eta)$ — тоже независимые случайные величины.*

Доказательство. Рассмотрим $B_1, B_2 \in \mathfrak{B}$. $P(f(\xi) \in B_1, g(\eta) \in B_2) = P(\xi \in f^{-1}(B_1), \eta g^{-1}(B_2)$ Но так как функции измеримы, то прообразы измеримых подмножеств будут также измеримыми. По определению 9.4 это равно $P(\xi \in f^{-1}(B_1)) P(\eta \in g^{-1}(B_2)) = P(f(\xi) \in B_1, g(\eta) \in B_2)$ $\square$

10 Свёртка распределения

Пусть имеется две независимые случайные величины ξ, η. Мы хотим узнать распределение $\xi + \eta$, зная функции распределения F_ξ и F_η.

Рассмотрим функцию распределения $\xi + \eta$ в произвольной точке x:

$$F_{\xi+\eta}(x) = P(\xi + \eta < x) = P((\xi, \eta) \in A_x) = \iint\limits_{A_x} dF_{(\xi,\eta)}(s,t) = \iint\limits_{A_x} d(F_\xi(s) F_\eta(t)) =$$

$$= \iint\limits_{A_x} dF_\xi(s)\,dF_\eta(t) = \int\limits_{-\infty}^{+\infty} \left(\int\limits_{-\infty}^{x-s} dF_\eta(t) \right) dF_\xi(s)$$

Таким образом, доказана следующая теорема.

Теорема 10.1 (О свёртке распределения).

$$F_{\xi+\eta}(x) = \int\limits_{-\infty}^{+\infty} F_\eta(x-s)\,dF_\xi(s) = \int\limits_{-\infty}^{+\infty} F_\xi(x-t)\,dF_\eta(t)$$

Следствие 10.2. *Если ξ, η — независимые случайные величины и ξ — с абсолютно непрерывным распределением, то $\xi + \eta$ — тоже с абсолютно непрерывным распределением и, кроме того, $p_{\xi+\eta}(x) = \int\limits_{-\infty}^{+\infty} p_\xi(x-t)\,dF_\eta(t)$. Если же обе случайные величины ξ, η с абсолютно непрерывным распределением, то $p_{\xi+\eta}(x) = \int\limits_{-\infty}^{+\infty} p_\xi(x-t) p_\eta(t)\,dt$.*

11 Числовые характеристики случайных величин. Математическое ожидание

Как обычно, будем рассматривать вероятностное пространство $(\Omega, \mathfrak{F}, P)$, случайная величина $\xi : \Omega \to \mathbb{R}$ с функцией распределения $F_\xi(x(= P(\xi < x)$ и распределением $\mathfrak{P}_\xi(x) = P(\xi^{-1}(A)) = P(\xi \in A)$.

Определение 11.1. Математическим ожиданием случайной величины ξ называется число $E\xi = \int\limits_{\Omega} \xi(\omega)\, dP(\omega)$, если таковое существует.

Интеграл можно записать следующим образом: $\int\limits_{\Omega} \xi^{+}(\omega)\, dP(\omega) + \int\limits_{\Omega} \xi^{-}(\omega)\, dP(\omega)$. Здесь:

$$\xi^{+} = \begin{cases} \xi, & \xi \geqslant 0 \\ 0, & \xi < 0 \end{cases} \qquad \xi^{-} = \begin{cases} 0, & \xi \geqslant 0 \\ \xi, & \xi < 0 \end{cases}$$

Тогда $|\xi| = \xi^{+} - \xi^{-}$. Ясно, что $E\xi$ и $E|\xi|$ существуют одновременно.

Замечание 11.2. В литературе математическое ожидание также может обозначаться как $M\xi$.

Для того, чтобы понять смысл этой величины, рассмотрим следующий пример.

Пример 11.3. Пусть ξ — случайная величина с дискретным распределением и законом распределения $(\xi, p) : (a_1, p_1), (a_2, p_2), \ldots, (a_n, p_n)$. Для такой величины:

$$E\xi = \int\limits_{\Omega} \xi(\omega)\, dP$$

Рассмотрим события $A_k = (\xi = a_k)$, причём $A_k \cap A_l = \varnothing$ при $k \neq l$ и $\bigcup\limits_{k} A_k = \Omega$. Тогда:

$$E\xi = \sum_{k} \int\limits_{A_k} \xi(\omega)\, dP = \sum_{k} a_k \underbrace{\int\limits_{A_k} dP}_{P(\xi = a_k)} = \sum_{k} a_k p_k$$

То есть, математическое ожидание является средневзвешенным значением случайной величины. В частности, если $p_1 = p_2 = \ldots = p_n = \frac{1}{n}$, то $E\xi = \frac{a_1 + \ldots + a_n}{n}$.

Предложение 11.4 (Свойства математического ожидания). *Пусть ξ, η — случайные величины. Верно:*

1. $P(\xi = c) = 1 \implies E\xi = c$
2. $E(a\xi + n\eta) = aE\xi + bE\eta$, *где* $a, b \in \mathbb{R}$
3. $P(\xi \geqslant 0) = 1 \implies E\xi \geqslant 0$
4. $P(\xi \geqslant \eta) = 1 \implies E\xi \geqslant E\eta$
5. $|E\xi| \leqslant E|\xi|$

Доказательство.

1. $E\xi = \int\limits_{\Omega} \xi(\omega)\, dP = c \int\limits_{\Omega} dP = c \cdot P(\Omega) = c$
2. Очевидно из линейности интеграла.
3. Очевидно.
4. $P(\xi \geqslant \eta) = P(\xi - \eta \geqslant 0) = 1 \implies E(\xi - \eta) = E\xi - E\eta \geqslant 0$
5. Так как $|\xi| \geqslant \xi$ и $|\xi| \geqslant -\xi$, то $E|\xi| \geqslant \xi$ и $E|\xi| \geqslant -E\xi$.

$\square$

Следующая теорема даёт правило вычисления математического ожидания случайных величин.

Теорема 11.5. *Пусть* $\xi_1, \xi_2, \ldots, \xi_n$ *— случайные величины, определённые на одном вероятностном пространстве. Обозначим через* ξ *случайные вектор* $\xi = (\xi_1, \ldots, \xi_n)$*. Рассмотрим также измеримую функцию* $f : \mathbb{R}^n \to \mathbb{R}$*. Тогда:*

$$\mathrm{E}f(\xi) = \int\limits_{\mathbb{R}^n} f(x_1, \ldots, x_n) \mathfrak{P}_\xi(\,dx_1, \ldots, \,dx_n) = \int\limits_{\mathbb{R}^n} f(x_1, \ldots, x_n)\, dF_\xi(x_1, \ldots, x_n)$$

Следствие 11.6. *Если* ξ *— случайная величина с функцией распределения* F_ξ*,* $f : \mathbb{R}^n \to \mathbb{R}$ *— измеримая функция, то*

$$\mathrm{E}f(\xi) = \int\limits_{\mathbb{R}} f(x)\, dF_\xi(x)$$

$$\mathrm{E}\xi = \int\limits_{\mathbb{R}} x\, dF_\xi(x)$$

Если известна плотность p_ξ*, то* $\mathrm{E}\xi = \int\limits_{\mathbb{R}} x p_\xi(x)\, dx$

Доказательство теоремы. Рассмотрим два случая.

1. Функция f принимает не более чем счётное число значений. Обозначим через $f_1, f_2, \ldots$ все значения функции f. Введём множества:

   $B_k = f^{-1}(f_k) \subset \mathbb{R}^n$, при этом $B_k \cap B_l = \varnothing$ $(k \neq l)$, $\bigcup\limits_k B_k = \mathbb{R}^n$;

   $A_k = \xi^{-1}(B_k) \in \mathfrak{F}$, $A_k \cap A_l = \varnothing$ $(k \neq l)$, $\bigcup\limits_k A_k = \Omega$.

   Рассмотрим математическое ожидание $\mathrm{E}f(\xi)$. По определению:

$$\mathrm{E}f(\xi) = \int\limits_{\Omega} f(\xi)\, dP = \sum_k \int\limits_{A_k} f(\xi)\, dP = \sum_k \int\limits_{A_k} f_k\, dP = \sum_k f_k \int\limits_{A_k} dP = \sum_k f_k P(A_k) =$$

$$= \sum_k f_k \mathfrak{P}_\xi(B_k) = \sum_k f_k \int\limits_{B_k} \mathfrak{P}_\xi(\,dx_1, \ldots, \,dx_n) = \sum_k \int\limits_{B_k} f_k P_\xi(\,dx_1, \ldots, \,dx_n) =$$

$$= \sum_k \int\limits_{B_k} f(x_1, \ldots, x_n) \mathfrak{P}_\xi(\,dx_1, \ldots, \,dx_n) = \int\limits_{\mathbb{R}^n} f(x_1, \ldots, x_n) \mathfrak{P}_\xi(\,dx_1, \ldots, \,dx_n)$$

2. Функция f — произвольная. Возьмём $m \geqslant 1$. Определим f^m следующим образом:

$$f^m(x_1, \ldots, x_n) = \frac{k}{m}, \text{ если } \frac{k}{m} \leqslant f(x_1, \ldots, x_n) \leqslant \frac{k+1}{m}$$

   То есть, мы будем приближать функцию f ступенчатой функцией:

$$\left| f(x_1, \ldots, x_n) - f^m(x_1, \ldots, x_n) \right| < \frac{1}{m}$$

   Отсюда:

$$\left| \mathrm{E}f(\xi) - \mathrm{E}f^m(\xi) \right| = \left| \mathrm{E}(f(\xi) - f^m(\xi)) \right| \leqslant \mathrm{E}\left| f(\xi) - f^m(\xi) \right| \leqslant \frac{1}{m}$$

$$\Big|\int\limits_{\mathbb{R}^n} f(x_1,\dots,x_n)\mathfrak{P}_\xi(\,dx_1,\dots,\,dx_n) - \int\limits_{\mathbb{R}^n} f^m(x_1,\dots,x_n)\mathfrak{P}_\xi(\,dx_1,\dots,\,dx_n)\Big| \leqslant \frac{1}{m}$$

Рассмотрим модуль:

$$\Big|Ef(\xi) - \int\limits_{\mathbb{R}^n} f(x_1,\dots,x_n)\mathfrak{P}_\xi(\,dx_1,\dots,\,dx_n)\Big|$$

Прибавим и вычтем в выражении под знаком модуля

$$\int\limits_{\mathbb{R}^n} f^m(x_1,\dots,x_n)\mathfrak{P}_\xi(\,dx_1,\dots,\,dx_n)$$

Получим:

$$\Big|Ef(\xi) - \int\limits_{\mathbb{R}^n} f(x_1,\dots,x_n)\mathfrak{P}_\xi(\,dx_1,\dots,\,dx_n)\Big| \leqslant$$

$$\leqslant \Big|Ef(\xi) - \int\limits_{\mathbb{R}^n} f^m(x_1,\dots,x_n)\mathfrak{P}_\xi(\,dx_1,\dots,\,dx_n)\Big| +$$

$$+\Big|f^m(x_1,\dots,x_n)\mathfrak{P}_\xi(\,dx_1,\dots,\,dx_n) - f(x_1,\dots,x_n)\mathfrak{P}_\xi(\,dx_1,\dots,\,dx_n)\Big| \leqslant \frac{1}{m} + \frac{1}{m} = \frac{2}{m}$$

Поскольку результат не зависит от m, то исходная разность равна нулю и теорема доказана.

□

Теорема 11.7 (Следствие для произведения независимых случайных величин). *Пусть* ξ, η *— независимые случайные величины и, кроме того, существуют* $E\xi$ *и* $E\eta$. *Тогда* $E(\xi\eta) = E\xi \cdot E\eta$.

Доказательство. Рассмотрим случайные вектор (ξ, η).

$$F_{(\xi,\eta)}(x,y) = F_\xi(x)F_\eta(y)$$

$$\mathfrak{P}_{(\xi,\eta)} = \mathfrak{P}_\xi \times \mathfrak{P}_\eta$$

Возьмём функцию $f : \mathbb{R}^2 \to \mathbb{R}$: $f(x,y) = xy$.

$$Ef(\xi,\eta) = E(\xi\eta) = \iint\limits_{\mathbb{R}^2} x_1x_2\mathfrak{P}_{(\xi,\eta)}(\,dx_1,\,dx_2) = \iint\limits_{\mathbb{R}^2} x_1x_2\mathfrak{P}_\xi(\,dx_1)\mathfrak{P}_\eta(\,dx_2) =$$

$$= \int\limits_{\mathbb{R}} x_1\mathfrak{P}_\xi(\,dx_1) \cdot \int\limits_{\mathbb{R}} x_2\mathfrak{P}_\eta(\,dx_2) = E\xi \cdot E\eta$$

□

12 Дисперсия

Определение 12.1. Дисперсией случайной величины ξ называется число $D\xi = E(\xi - E\xi)^2$, если оно существует.

Дисперсия представляет собой среднеквадратичное отклонение случайной величины от своего среднего значения.

Предложение 12.2 (Свойства дисперсии).

1. $D\xi \geqslant 0$. $D\xi = 0 \iff \exists c : P(\xi = c) = 1$
2. *Если* $a, b \in \mathbb{R}$, *то* $D(a\xi + b) = a^2 D\xi$
3. *Если* ξ, η *— случайные величины, то* $D(\xi + \eta) = D\xi + D\eta + 2E((\xi - E\xi)(\eta - E\eta))$.
4. *Если* ξ, η *— независимые случайные величины, то* $D(\xi + \eta) = D\xi + D\eta$. *Обратное, вообще говоря, неверно.*
5. $D\xi = E\xi^2 - (E\xi)^2$

Определение 12.3. Величина $E((\xi - E\xi)(\eta - E\eta))$ называется *ковариацией* случайных величин ξ и η. Обозначение: $\mathrm{cov}(\xi, \eta)$.

Доказательство предложения.

1. Неотрицательность следует из свойств математического ожидания. Далее, $D\xi = \int\limits_{\Omega} (\xi(\omega) - E\xi)^2 \, dP(\omega) = 0$. Это означает, что $P(\xi - E\xi = 0) = 1$, возьмём $c = E\xi$.
2. $D(a\xi + b) = E(a\xi + b - E(a\xi + b))^2 = E(a\xi + b - aE\xi - b)^2 = a^2 E(\xi - E\xi)^2 = a^2 D\xi$.
3. $D(\xi + \eta) = E(\xi + \eta - E\xi - E\eta)^2 = E(\xi - E\xi)^2 + E(\eta - E\eta)^2 + 2E((\xi - E\xi)(\eta - E\eta))$
4. По теореме 9.10 $\xi - E\xi$ и $\eta - E\eta$ будут независимыми, отсюда $E((\xi - E\xi)(\eta - E\eta)) = E(\xi - E\xi) \cdot E(\eta - E\eta) = 0$. Подставим полученный результат в предыдущее свойство.
5. $D\xi = E(\xi - E\xi) \ldots$

$\square$

Примеры 12.4.

1. $(\xi, p) : (c, 1)$ (вырожденное распределение). Тогда $E\xi = c$, $D\xi = 0$.
2. Если $(\xi, p) : (0, 1 - p), (1 - p, p)$, то $E\xi = p$, $D\xi = p(1 - p)$.
3. $\xi \in B(n, p)$ — случайная величина с биномиальным законом распределения, т. е. $p(\xi = k) = C_n^k p^k (q - p)^{n-k}$. Для этой величины $E\xi = \sum\limits_{k=0}^{n} k \cdot C_n^k p^k (1-p)^{n-k} = \sum\limits_{k=1}^{n} k \cdot C_n^k p^k (1-p)^{n-k} = \sum\limits_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} = np \sum\limits_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-1-(k-1))!} p^{k-1} (1-p)^{n-1-(k-1)}$. Взяв $l = k - 1$, получим $E\xi = np \sum\limits_{l=0}^{n-1} C_n^l p^l (1-p)^{n-1-l} = np$. $D\xi = np(1 - p)$.
4. Рассмотрим $\xi \in \Pi(\lambda)$ — случайную величину с пуассоновским законом распределения: $P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$. $E\xi = \sum\limits_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \sum\limits_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} = \lambda e^{-\lambda} \sum\limits_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$. Для вычисления дисперсии вычислим сначала $E\xi^2 = \sum\limits_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum\limits_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} e^{-\lambda} = \sum\limits_{k=1}^{\infty} (k-1) \frac{\lambda^k}{(k-1)!} e^{-\lambda} + \sum\limits_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} = \sum\limits_{k=2}^{\infty} \frac{\lambda^k}{(k-2)!} e^{-\lambda} + \lambda e^{-\lambda} \sum\limits_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda^2 e^{-\lambda} \sum\limits_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!}$ ...
5. Для $\xi \in U_{[a,b]}$ математическое ожидание $E\xi = \int\limits_a^b x \frac{1}{b-a} \, dx = \frac{a+b}{2}$. $E\xi^2 = \int\limits_a^b x^2 \frac{1}{b-a} \, dx = \frac{b^3 - a^3}{3(b-a)}$. Тогда $D\xi = E\xi^2 - (E\xi)^2 = \frac{(b-a)^2}{12}$.

6. $\xi \in N(a,\sigma^2)$ — случайная величина с нормальным законом распределения ($p_\xi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}$).

$$E\xi = \int\limits_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}\, dx = \int\limits_{-\infty}^{+\infty} (\sigma y - a) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2}} \sigma\, dy =$$

$$= \underbrace{\sigma \int\limits_{-\infty}^{+\infty} y \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy}_{0} + a \int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy = a$$

$$D\xi = \int\limits_{-\infty}^{+\infty} (x-a)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}\, dx = \int\limits_{-\infty}^{+\infty} \sigma^2 y^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2}} \sigma\, dy = \sigma^2 \int\limits_{-\infty}^{+\infty} y^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy =$$

$$= \sigma^2 \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} y \cdot y e^{-\frac{y^2}{2}}\, dy = \sigma^2 \frac{1}{\sqrt{2\pi}} \left[-y e^{-\frac{y^2}{2}} \Big|_{-\infty}^{+\infty} + \int\limits_{-\infty}^{+\infty} e^{-\frac{y^2}{2}}\, dy \right] = \sigma^2$$

7. ξ — случайная величина с распределением Коши ($p_\xi(x) = \frac{1}{\pi(1+x^2)}$). $\int\limits_{-\infty}^{+\infty} \frac{x}{\pi(1+x^2)}\, dx$ — этот интеграл расходится, значит, математического ожидания не существует.

13 Моменты случайных величин

Определение 13.1. Пусть k — целое неотрицательное число, ξ — случайная величина. Число $E\xi^k$, если оно существует, называется k-м моментом ξ. Число $E|\xi|^k$ — абсолютный k-й момент ξ. Число $E(\xi - E\xi)^k$ — центральный k-й момент ξ. Число $E|\xi - E\xi|^2$ — абсолютный центральный k-й момент ξ.

Заметим, что все моменты величины ξ существуют или отсутствуют одновременно.

Предложение 13.2 (Свойства моментов).

1. *Пусть $n > m$ и существует $E\xi^n$. Тогда существует и $E\xi^m$.*
2. *(Неравенство Гёльдера) Если $p, q > 0$ и $\frac{1}{p} + \frac{1}{q} = 1$, то $E|\xi\eta| \leqslant (E|\xi|^p)^{\frac{1}{p}} \cdot (E|\eta|^q)^{\frac{1}{q}}$.*
3. *(Неравенство Коши-Буняковского-Шварца) $E|\xi\eta| \leqslant \sqrt{E\xi^2} \cdot \sqrt{E\eta^2}$.*
4. *(Неравенство Ляпунова) Пусть $0 < p < q$. Тогда $(E|\xi|^p)^{\frac{1}{p}} \leqslant (E|\xi|^q)^{\frac{1}{q}}$.*

Доказательство. 1. Ясно, что $|x|^m \leqslant |x|^n + 1$ для любого x. То есть, $|\xi|^m \leqslant |\xi|^n + 1$ и $E|\xi|^m \leqslant E|\xi|^n + 1 < \infty$.

2. Очевидно из неравенства Гёльдера для интегралов.
3. Это свойство — частный случай предыдущего при $p = q = \frac{1}{2}$.
4. Данное неравенство равносильно $E|\xi|^p \leqslant (E|\xi|^q)^{\frac{p}{q}}$. С учётом $\frac{p}{q} + \frac{q-p}{q} = 1$, применим неравенство Гёльдера:

$$E(\xi|^p \cdot 1|) \leqslant (E|\xi|^q)^{\frac{p}{q}} \cdot (E(1^{\frac{q}{q-p}}))^{\frac{q-p}{q}} = (E|\xi|^q)^{\frac{p}{q}}$$

$\square$

Определение 13.3. Пусть $m, n > 0$. Смешанным моментом величин ξ, η порядка m плюс n называется число $E\xi^m\eta^n$, если оно существует. Смешанным центральным моментом ξ, η порядка m плюс n называется число $E(\xi - E\xi)^m(\eta - E\eta)^n$.

Замечание 13.4. Таким образом, $\mathrm{cov}(\xi, \eta)$ является смешанным центральным моментом величин порядка 1 плюс 1.