

Теория вероятностей и математическая статистика

С. М. Ананьевский *

11 октября 2016 г.

Содержание

1	Условные вероятности	2
2	Формула полной вероятности. Формула Байеса	2
3	Независимые события. Пример Бернштейна	3
4	Независимые испытания Бернулли. Формулы Бернулли	5
5	Предельные теоремы в схеме испытаний Бернулли	6
6	Случайная величина. Распределение случайных величин	9
7	Различные типы распределений случайных величин. Случайные величины с дискретным распределением	10
8	Случайные векторы. Распределение случайных векторов	12
9	Независимость случайных величин	12

*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 $A_1, A_2, \dots \in \mathfrak{F}$; $A_i \cap A_j = \emptyset$ (при $i \neq j$). Тогда:

$$\begin{aligned} P_B\left(\bigcup_i A_i\right) &= \frac{P\left(\bigcup_i B \cap A_i\right)}{P(B)} = \frac{\sum_i P(B \cap A_i)}{P(B)} = \\ &= \sum_i \frac{P(B \cap A_i)}{P(B)} = \sum_i P_B(A_i) \end{aligned}$$

□

Следствие 1.4. Все свойства вероятности для условной вероятности выполнены.

2 Формула полной вероятности. Формула Байеса

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

Теорема 2.2 (Формула полной вероятности). Пусть A_1, A_2, \dots — 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.

$$\begin{aligned} P(B) &= P(B \cap \Omega) = P\left(B \cap \left(\bigcup_i A_i\right)\right) = P\left(\bigcup_i (B \cap A_i)\right) = \sum_i P(B \cap A_i) = \sum_i \frac{P(B \cap A_i) \cdot P(A_i)}{P(A_i)} = \\ &= \sum_i P(B/A_i) \cdot P(A_i) \end{aligned}$$

□

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

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 C_1, 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

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

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 $\iff A, \bar{B}$ независимы $\iff \bar{A}, B$ независимы $\iff \bar{A}, \bar{B}$ независимы.

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

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

Доказательство.

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

$$P(A)(1 - P(B)) = P(A)P(B) = P(A \cap B)$$

□

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, \dots, 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, \dots, A_n независимы в совокупности, если:

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

Это равносильно:

$$\forall 2 \leq k \leq n \quad \forall i_1 < i_2 < \dots < i_k \quad P(\bigcap_{j=1}^k A_{i_j}) = \prod_{j=1}^k P(A_{i_j})$$

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

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

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

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

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

$$P(B \cap C \cap K) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(B)P(C)P(K)$$

Это доказывает, что попарная независимость и совместная независимость неравносильны.

4 Независимые испытания Бернулли. Формулы Бернулли

Пусть у нас есть вероятностное пространство (Ω, \mathcal{F}, P) . Рассмотрим набор случайных событий $\mathcal{A} = (A_1, \dots, A_n)$, 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 испытанию $A_1 = \{A_{11}, A_{12}\}$, где $A_{11} = \{\text{на монете О}\}$, $A_{12} = \{\text{на монете Р}\}$; $A_2 = \{A_{21}, \dots, A_{26}\}$, 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 $(Y_1, Y_2, H_3, \dots, Y_n)$, а вероятность успеха $P(Y_k) = p$.

Пусть $A = \{\text{в } n \text{ испытаниях Бернулли успех произошёл ровно } k \text{ раз}\}$. Какова вероятность A ? Заметим, что $A = \underbrace{(YU \dots YU)}_k \underbrace{NN \dots N}_{n-k} \cup \underbrace{(YU \dots YU)}_{k-1} \underbrace{NU \dots NU}_{n-k+1} \cup \dots$

$$P(YU \dots YU \dots NN \dots N) = P(Y_1 \cap Y_2 \cap \dots \cap Y_k \cap H_{k+1} \cap \dots \cap H_n) = p^k(1-p)^{n-k}$$

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:

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

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

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

Возникает вопрос, каково наиболее вероятное число выпадений орла? $P_{10}(k) = C_{10}^k (1/2)^{10}$, 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:

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

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

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 этом 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?

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

Рассмотрим функции:

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

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:*

$$\frac{P_n(k)}{\frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x_{n,k}^2}{2}}} \xrightarrow{n \rightarrow \infty} 1 \quad \text{равномерно по всем } k: \quad |x_{n,k}| \leq C n^{\frac{1}{\delta} - \varepsilon} \quad \forall C > 0, \varepsilon > 0$$

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

Лемма 5.2 (Формула Стирлинга).

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

Лемма 5.3.

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

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 $b \rightarrow \infty$, $k \rightarrow \infty$, $n-k \rightarrow \infty$.

$$k = np + x_{n,k} \sqrt{npq} \xrightarrow{n \rightarrow \infty} \infty$$

$$n-k = nq - x_{n,k} \sqrt{npq} \xrightarrow{n \rightarrow \infty} \infty$$

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:

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

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

С учётом

$$\begin{aligned} \frac{k}{np} &= 1 + x \sqrt{\frac{q}{np}} \\ \frac{n-k}{nq} &= 1 - x \sqrt{\frac{p}{nq}} \end{aligned}$$

Применив лемму 5.3, получим:

$$\begin{aligned} \ln T_{n,k} &= \left(-np - x\sqrt{npq} - \frac{1}{2}\right) \left(x\sqrt{\frac{q}{np}} - \frac{x^2}{2} \frac{q}{np} + \theta_1 \frac{x^3 q \sqrt{q}}{np \sqrt{np}}\right) + \\ &+ \left(-np + x\sqrt{npq} - \frac{1}{2}\right) \left(-x\sqrt{\frac{p}{nq}} - \frac{x^2}{2} \frac{p}{nq} + \theta_2 x^3 \frac{p}{nq} \frac{\sqrt{p}}{\sqrt{nq}}\right) + \ln \frac{1}{\sqrt{2\pi}} + o(1) = \\ &= -x\sqrt{npq} + \frac{x^2}{2} q + o(1) - x^2 q + x\sqrt{npq} + \frac{x^2}{2} p - x^2 p + \ln \frac{1}{\sqrt{2\pi}} = \\ &= -\frac{x^2}{2} + \ln \frac{1}{\sqrt{2\pi}} + o(1) \end{aligned}$$

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. □

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:

$$\sup_{a < b} \left| P_n(a < \text{число успехов} < b) - \int_{\frac{a-np}{\sqrt{npq}}}^{\frac{b-np}{\sqrt{npq}}} \varphi(t) dt \right| \xrightarrow{n \rightarrow \infty} 0$$

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

Следствие 5.5.

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

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

Если мы возьмём функцию $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$, 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:

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

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

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

Доказательство.

$$\begin{aligned} P_n(k) &= C_n^k p_n^k (1-p_n)^{n-k} = \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^{n-k} = \\ &= \frac{1}{k!} \frac{n(n-1)\dots(n-k+1)}{n^k} n^k p_n^k (1-p_n)^n (1-p_n)^{-k} = \\ &= \frac{\lambda^k}{k!} \underbrace{\left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right)}_{\rightarrow 1} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{\rightarrow e^{-\lambda}} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-k}}_{\rightarrow 1} \end{aligned}$$

□

Теорема 5.8 (Закон больших чисел Бернулли).

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

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

Доказательство. Рассмотрим вероятность:

$$\begin{aligned} P_n(|\frac{k_n}{n} - p| > \varepsilon) &= 1 - P_n(|\frac{k_n}{n} - p| \leq \varepsilon) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt - P_n(np - n\varepsilon \leq k_n \leq n\varepsilon + np) = \\ &= \frac{1}{\sqrt{2\pi}} \int_{|t| \leq \varepsilon \sqrt{\frac{n}{pq}}} e^{-\frac{t^2}{2}} dt + \frac{1}{\sqrt{2\pi}} \int_{|t| > \varepsilon \sqrt{\frac{n}{pq}}} e^{-\frac{t^2}{2}} dt - P_n(np - n\varepsilon \leq k_n \leq n\varepsilon + np) \end{aligned}$$

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

6 Случайная величина. Распределение случайных величин

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

Определение 6.1. Функция $\xi : \Omega \rightarrow \mathbb{R}$ такая, что $\forall B \in \mathfrak{B} \quad \xi^{-1}(B) \in \mathfrak{F}$, 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

$$\xi_1(\omega_i) = i, \quad \xi_2(\omega_i) = \begin{cases} 0, & i \neq 6 \\ 1, & i = 6 \end{cases} \quad . \quad \xi_1, \xi_2 - \text{случайные величины.}$$

2. Пусть $\Omega = \{x | x \in [0, 1]\}$, $\mathfrak{F} = \{[0, \frac{1}{2}], (\frac{1}{2}, 1], [0, 1], \emptyset\}$, $\eta(x) = x$. η не будет случайной величиной, так как $\eta^{-1}([0, \frac{1}{3}]) = [0, \frac{1}{3}] \notin \mathfrak{F}$.

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

Определение 6.3. \mathfrak{P}_ξ называется *распределением случайной величины* ξ .

Теорема 6.4. \mathfrak{P}_ξ является вероятностью.

Доказательство. Проверим аксиомы вероятности.

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

$$\begin{aligned} \mathfrak{P}_\xi\left(\bigcup_{i=1}^{\infty} B_i\right) &= P\left(\xi^{-1}\left(\bigcup_{i=1}^{\infty} B_i\right)\right) = \\ &= P\left(\bigcup_{i=1}^{\infty} \underbrace{\xi^{-1}(B_i)}_{\text{попарно несовместны}}\right) = \sum_i P(\xi^{-1}(B_i)) = \sum_i \mathfrak{P}_\xi(B_i) \end{aligned}$$

\square

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

Определение 6.5. Функция $F_\xi : \mathbb{R} \rightarrow \mathbb{R}$ 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}_ξ и F_ξ .

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

2. F_ξ неубывает;

3. F_ξ непрерывна слева во всех точках.

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

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

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

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

Отсюда получаем, что $\lim_{n \rightarrow \infty} F_\xi(n) = 1$ и $\lim_{n \rightarrow \infty} F_\xi(-n) = 0$. Свойство 4 доказано. \square

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

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

Доказательство. Упражнение. \square

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 : $A = \{a_1, a_2, \dots\}$, 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 $x \leq c$ 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. $(\xi, p) : (a_1, p_1), (a_2, p_2), \dots, (a_n, p_n)$. График — аналогичная ступенчатая функция.
4. $\xi : 0, 1, 2, \dots, n$ 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. $\xi : 0, 1, 2, \dots$ и $P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ — случайная величина с распределением Пуассона с параметром $\lambda > 0$.
6. $\xi : 0, 1, 2, \dots$ и $P(\xi = k) = p^k (1 - p)$ — случайная величина с геометрическим распределением с параметром p ($0 \leq p \leq 1$).

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

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

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

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 .

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

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 $P(a < \xi < b) = P(a \leq \xi < b) = P(a \leq \xi \leq b) = F_\xi(b) - F_\xi(a)$. Почему строгость неравенства не имеет значения? В силу непрерывности функции распределения $P(a < \xi < b) = F_\xi(b) - F_\xi(a) \xrightarrow{b \rightarrow a} 0$.

8 Случайные векторы. Распределение случайных векторов

Определение 8.1. Пусть ξ_1, \dots, ξ_d — случайные величины, определённые на одном вероятностном пространстве. Тогда $\xi = (\xi_1, \dots, \xi_d)$ будет называться случайным d -мерным вектором, $\xi: \Omega \rightarrow \mathbb{R}^d$. Иначе говоря, $\forall B \in \mathfrak{B}^d \quad \xi^{-1}(B) \in \mathfrak{F}$.

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

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

ξ_1, \dots, ξ_d — с дискретным распределением, если существует $A \subset \mathbb{R}^d$ такое, что A не более чем счётно и $P((\xi_1, \dots, \xi_d) \in A) = 1$.

ξ_1, \dots, ξ_d — с абсолютно непрерывным распределением, если существует $p(x_1, \dots, x_d)$ такое, что $F_{(\xi_1, \dots, \xi_d)}(y_1, \dots, y_d) = \int_{-\infty}^{y_1} \dots \int_{-\infty}^{y_d} p(x_1, \dots, x_d) dx_1 \dots dx_d$.

Предложение 8.3 (Свойства функции распределения случайного вектора).

1. $0 \leq F_{(\xi_1, \dots, \xi_d)}(x_1, \dots, x_d) \leq 1$
2. Пусть $x_1 \leq y_1, \dots, x_d \leq y_d$. Тогда $F_{(\xi_1, \dots, \xi_d)}(x_1, \dots, x_d) \leq F_{(\xi_1, \dots, \xi_d)}(y_1, \dots, y_d)$
3. $F_{(\xi_1, \dots, \xi_d)}$ непрерывна слева по каждой координате.
4. $F_{(\xi_1, \dots, \xi_d)}(x_1, \dots, x_d) \xrightarrow{x_k \rightarrow -\infty} 0$.
5. $F_{(\xi_1, \dots, \xi_d)}(x_1, \dots, x_d) \xrightarrow{x_k \rightarrow +\infty} F_{(\xi_1, \dots, \xi_{k-1}, \xi_{k+1}, \dots, \xi_d)}(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_d)$.
6. $F_{(\xi_1, \dots, \xi_d)}(x_1, \dots, x_d) \xrightarrow{x_1 \rightarrow +\infty, \dots, x_d \rightarrow +\infty} 1$
7. Пусть $d = 2$. Тогда $P((\xi_1, \xi_2) \in [a_1, b_1] \times [a_2, b_2]) = F_{(\xi_1, \xi_2)}(b_1, b_2) - F_{(\xi_1, \xi_2)}(a_1, b_2) - F_{(\xi_1, \xi_2)}(b_1, a_2) + F_{(\xi_1, \xi_2)}(a_1, a_2)$. 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 $x, y \quad F_{(\xi, \eta)}(x, y) = F_\xi(x)F_\eta(y)$.

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

Предложение 9.3. Определения 9.1 и 9.2 эквивалентны.

Доказательство.

1. $(9.1 \Rightarrow 9.2)$

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

При $a_1, a_2 \rightarrow -\infty$ 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 \Rightarrow 9.1)

$$\begin{aligned}
 P((\xi, \eta) \in [a_1, b_1] \times [a_2, b_2]) &= F_{(\xi, \eta)}(b_1, b_2) - F_{(\xi, \eta)}(a_1, b_2) - F_{(\xi, \eta)}(b_1, a_2) + \\
 &+ F_{(\xi, \eta)}(a_1, a_2) = F_{\xi}(b_1)F_{\eta}(b_2) - F_{\xi}(a_1)F_{\eta}(b_2) - F_{\xi}(b_1)F_{\eta}(a_2) + \\
 &+ F_{\xi}(a_1)F_{\eta}(a_2) = (F_{\xi}(b_1) - F_{\xi}(a_1))(F_{\eta}(b_2) - F_{\eta}(a_2)) = \\
 &P(a_1 \leq \xi < b_1) \cdot P(a_2 \leq \eta < b_2)
 \end{aligned}$$

□