Теория вероятностей

(Ещё не)алгебраист

10 июня 2025 г.

Предисловие

Эти записки созданы с целью аккуратно формализовать и заполнить пробелы в лекциях Елены Борисовны Яровой. В разделе 0 будут содержаться основные принятые в курсе обозначения, а также сведения и определения из разных разделов математики, которыми автор будет пользоваться. Поскольку автор считает полезным взгляд на всякий раздел математики с точки зрения теории категорий и её приложений, этот язык также будет упоминаться (тем не менее, не замещая собой прочие подходы).

Содержание

0 Предварительные сведения

0.1 Обозначения

- Ω — пространство элементарных исходов;
- ω — элементарный исход;
- $\mathfrak{F}$ — σ-алгебра событий;
- P — вероятностная мера;
- ξ, η, ζ — случайные величины;
- $\mathrm{E}\,\xi$ — математическое ожидание случайной величины ξ;
- $\mathrm{D}\,\xi$ — дисперсия случайной величины ξ;
- $\mathrm{Cov}(\xi, \eta)$ — ковариация случайных величин ξ и η;
- $\rho(\xi, \eta)$ — корреляция случайных величин ξ и η;

0.2 Предварительные сведения из действительного анализа

0.2.1 (σ-)Алгебры и меры

Пусть Ω — некоторое множество.

Система множеств S называется полуцольком, если выполнены следующие аксиомы

(1) $\varnothing \in S$;

(2) $A, B \in S : A \cap B \in S$;

(3) $A, B \in S, A \subset B\ \exists n \in \mathbb{N} \exists C_1, \ldots, C_n \in S : A = B \sqcup \bigsqcup\limits_{k=1}^{n} C_k$

Система множеств (следует понимать как синоним термина «семейство множеств») $R \subset 2^\Omega$ называется алгеброй с единицей Ω, если выполнены первые три из следующих аксиом и σ-алгеброй с единицей Ω, если выполнены все четыре аксиомы.

(1) $\Omega \in R$;

(2) $\forall\ A, B \in R : A \cup B, A \cap B \in R$;

(3) $\forall\ A \in R:\ \Omega \setminus A := \overline{A} \in R$;

(4) $\forall\ \{A_k\}_{k \in R} \subset R:\ \bigcup\limits_{k \in \mathbb{N}} A_k \in R$.

Далее, если не оговорено иное, все алгебры являются (σ-)алгебрами с единицей Ω и будут называться «(σ-)алгебрами».

Будем называть функцию $\mu\colon R \to \mathbb{R}$ (конечной) мерой на (σ-)алгебре R, если ν удовлетворяет аксиоме аддитивности

$$\forall A, B \in R, A \cap B = \varnothing\ \mu(A \sqcup B) = \mu(S) + \mu(B).$$

Если дополнительно для любой последовательности попарно непересекающихся подмножеств $\{A_k\}_{k \in \mathbb{N}}$, объединение которых есть элемент R (отметим, что это автоматически выполнено для σ-алгебры) имеет место равенство

$$\mu\left(\bigsqcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k),$$

то мера μ называется σ-аддитивной (аксиома σ-аддитивности). Можно показать, что из этой аксиомы следует, что $\mu(\varnothing) = 0$ и поэтому из неё следует аксиома аддитивности.

Лемма 0.1. *Пусть $\mathfrak{F}$ — (σ-)алгебра и $A \in \mathfrak{F}$. Тогда множество*

$$\mathfrak{F} \cap A := \{B \cap A |\, B \in \mathfrak{F}\}$$

является (σ-)алгеброй с единицей A.

Доказательство. По построению $\Omega \cap A = A$ содержится в $\mathfrak{F} \cap A$.

Пусть теперь $C_1 = B_1 \cap A, C_2 = B_2 \cap A \in \mathfrak{F} \cap A$ — два множества. Тогда $C_1 \cap C_2 = (B_1 \cap B_2) \cap A \in \mathfrak{F} \cap A$, так как $B_1 \cap B_2 \in \mathfrak{F}$. Далее, $C_1 \cup C_2 = (B_1 \cup B_2) \cap A \in \mathfrak{F} \cap A$, так как $B_1 \cup B_2 \in \mathfrak{F}$. Окончательно, $A \setminus C_1 = (A \setminus B_1) = (\Omega \setminus B_1) \cap A \in \mathfrak{F} \cap A$, поскольку $\Omega \setminus B_1 \in \mathfrak{F}$.

Предположим, что $\mathfrak{F}$ являлось σ-алгеброй. Пусть $\{C_k\}$ — счётное семейство элементов $\mathfrak{F} \cap A$ и $C_k = B_k \cap A$. Тогда

$$\bigcup_{i=1}^{\infty} C_k = \bigcup_{i=1}^{\infty} (B_k \cap A) = \left(\bigcup_{i=1}^{\infty} B_k\right) \cap A \in \mathfrak{F} \cap A,$$

принадлежность справедлива в силу того, что $\bigcup\limits_{i=1}^{\infty} B_k \in \mathfrak{F}$. $\square$

0.2.2 Лебеговское продолжение меры

0.2.3 Мера Лебега-Стилтьеса

0.2.4 Измеримое отображение

0.2.5 Интеграл Лебега

0.2.6 Прямой образ меры (pushforward measure)

0.3 Предварительные сведения из анализа Фурье

0.4 Предварительные сведения из линейной алгебры

0.4.1 Билинейные функции и квадратичные формы

Пусть $\Bbbk$ — некоторое поле (в нашем случае будут рассматриваться только поля вещественных чисел $\mathbb{R}$) и V — векторное пространство над $\Bbbk$.

Отображение $B\colon V \times V \to \Bbbk$ называется билинейной функцией, если выполнены следующие аксиомы

(1) $\forall v, u, w \in V\ B(u + v, w) = B(u, w) + B(v, w)$;

(2) $\forall v, u \in V, \lambda \in \Bbbk\ B(\lambda u, v) = \lambda B(u, v)$;

(3) $\forall v, u, w \in V\ B(u, v + w) = B(u, w) + B(u, v)$;

(4) $\forall v, u \in V, \lambda \in \Bbbk\ B(u, \lambda v) = \lambda B(u, v)$.

Билинейная функция называется симметрической, если дополнительно для любых $u, v \in V$ выполнено $B(u, v) = B(v, u)$.

Пример. Пусть $V = \Bbbk$ и $B(a, b) = a \cdot b$, где $\cdot$ — умножение в поле $\Bbbk$. Тогда B — симметрическая билинейная функция.

Пример. Пусть в векторном пространстве V фиксирован базис $e_1, \ldots, e_n$. Тогда если $B(x, y) = \sum\limits_{i=1}^{n} x_i y_i$, где $x = \sum\limits_{i=1}^{n} x_i e_i$ и $y = \sum\limits_{i=1}^{n} y_i e_i$, то B — также билинейная симметрическая форма.

Квадратичной формой называется отображение $Q\colon V \to \Bbbk$ такое, что для некоторой билинейной формы и любой вектора $v \in V$ имеет место равенство $Q(v) = B(v, v)$. Если B — билинейная функция, то квадратичная форма Q, заданная формулой $Q(v) = B(v, v)$ называется квадратичной формой соответствующей билинейной функции B. Пусть $\Bbbk = \mathbb{R}$, Q — квадратичная форма и для любого ненулевого вектора $v \in V$ выполнено неравенство $Q(v) > 0$. Тогда форма Q называется положительно определённой. Если для любого $v \in V$ выполнено неравенство $Q(v) \geqslant 0$, то форма Q называется неотрицательно определённой.

Симметрическую билинейную форму с положительно определённой соответствующей квадратичной формой называют скалярным произведением. Вместо $B(u, v)$ часто пишут (u, v) или $\langle u, v \rangle$.

Примеры. Квадратичные формы, соответствующие билинейным функциям из примеров выше являются положительно определёнными.

Теорема 0.2 (Коши, Буняковский, Шварц). *Пусть V — векторное пространство над полем $\mathbb{R}$ и B — скалярное произведение на V. Тогда дл любых двух векторов $u, v \in V$ выполнено равенство*

$$B(u,v)^2 \leqslant B(u,u)B(v,v),$$

причём равенство достигается тогда и только тогда, когда u и v коллинеарны.

Доказательство. Рассмотрим вектор $u + tv$, где $t \in \mathbb{R}$ и значение квадратичной формы на нём. По билинейности, симметричности и положительной определённости имеем

$$B(u+tv, u+tv) = B(u,u)+tB(u,v)+tB(v,u)+t^2B(v,v) = B(u,u)+2tB(u,v)+t^2B(v,v) \geqslant 0,$$

причём последнее равенство достигается тогда и только тогда, когда $u + tv = 0$.

Многочлен второй степени принимает только неотрицательные (положительные) значения тогда и только тогда, когда его дискриминант меньше или равен 0 (меньше 0). Итого

$$D = 4B(u,v)^2 - 4B(u,u)B(v,v) \leqslant 0 \Leftrightarrow B(u,v)^2 \leqslant B(u,u)B(v,v)$$

и $D = 0 \Leftrightarrow B(u,v)^2 = B(u,u)B(v,v)$. Последнее равносильно тому, что многочлен имеет корень t и $u + tv = 0$, то есть u и v пропорциональны. □

Заметьте, что доказательство этого неравенства в случае поля комплексных чисел требует добавления дополнительной «поправки» λ.

0.4.2 Полуторалинейные функции

В этом подразделе будем рассматривать только векторные пространства над полем комплексных чисел.

Отображение $S\colon V \times V \to \mathbb{k}$ называется полуторалинейной функцией (по второму аргументу), если выполнены следующие аксиомы

(1) $\forall v, u, w \in V$ $S(u + v, w) = S(u, w) + S(v, w)$;

(2) $\forall v, u \in V, \lambda \in \mathbb{k}$ $S(\lambda u, v) = \lambda S(u, v)$;

(3) $\forall v, u, w \in V$ $S(u, v + w) = S(u, w) + S(u, v)$;

(4) $\forall v, u \in V, \lambda \in \mathbb{k}$ $S(u, \lambda v) = \overline{\lambda} S(u, v)$, где надчёркивание означает комплексное сопряжение.

Полуторалинейная функция называется эрмитовой, если для любых векторов u и v дополнительно выполнено равенство $S(u,v) = \overline{S(v,u)}$.

Эрмитова функция называется скалярным произведением, если для любого ненулевого вектора v выполнено неравенство $S(v,v) > 0$.

Теорема 0.3 (Коши, Буняковский, Шварц)**.** *Пусть V — векторное пространство над полем $\mathbb{C}$ и S — скалярное произведение на V. Тогда для любых двух векторов $u, v \in V$ выполнено равенство*

$$S(u,v)\overline{S(u,v)} \leqslant S(u,u)S(v,v),$$

причём равенство достигается тогда и только тогда, когда u и v коллинеарны.

Доказательство. Если $S(u,v) = 0$, то неравенство выполнено. При таком условии u и v пропорциональны тогда и только тогда, когда один из этих векторов равен 0. Последнее в свою очередь равносильно тому, что правая часть неравенства обращается в нуль. Далее будем считать, что $S(u,v) \neq 0$.

Рассмотрим вектор $u + t\lambda v$, где $t \in \mathbb{R}$ и $\lambda = S(u,v)$. Поскольку S — скалярное произведение и из условий наложенных на λ, то

$$S(u + t\lambda v, u + t\lambda v) = S(u,u) + t\overline{\lambda}S(u,v) + t\lambda S(v,u) + t^2\lambda\overline{\lambda}S(v,v) =$$

$$= S(u,u) + 2tS(u,v)S(v,u) + t^2 S(u,v)S(v,u)S(v,v) \leqslant 0$$

причём последнее равенство достигается тогда и только тогда, когда $u + t\lambda v = 0$.

Многочлен второй степени принимает только неотрицательные (положительные) значения тогда и только тогда, когда его дискриминант меньше или равен 0 (меньше 0). Итого

$$D = 4S(u,v)^2S(v,u)^2 - 4S(u,u)S(v,v)S(u,v)S(v,u) \leqslant 0 \Leftrightarrow S(u,v)S(v,u) \leqslant S(u,u)S(v,v)$$

и $D = 0 \Leftrightarrow S(u,v)^2 = S(u,u)S(v,v)$. Последнее равносильно тому, что многочлен имеет корень t_0 и $u + t_0 S(u,v)v = 0$, то есть u и v пропорциональны. □

0.5 Теория категорий и взгляд на измеримые пространства с её точки зрения

0.5.1 Категория измеримых пространств

0.5.2 Прямой образ σ-алгебры

0.5.3 Обратный образ σ-алгебры

0.5.4 Функтор борелевской σ-алгебры

1 Вероятностное пространство, случайные события

Пусть Ω — некоторое множество, $\mathfrak{F}$ — σ-алгебра с единицей Ω и P — σ-аддитивная мера на $\mathfrak{F}$, удовлетворяющая свойству $\mathrm{P}(\Omega) = 1$. Тогда тройка $(\Omega, \mathfrak{F}, \mathrm{P})$ называется вероятностным пространством. Множество Ω называется пространством элементарных событий (исходов), элементы σ-алгебры $\mathfrak{F}$ называются событиями.

Вероятностное пространство называется дискретным, если множество Ω не более, чем счётно.

Для кратности, если множество $\{\omega\}$ является событием, вместо $\mathrm{P}(\omega)$ будем писать $\mathrm{P}(\omega)$.

Примеры. Пусть $\Omega = \{1, 2, 3, 4, 5, 6\}$ — числа, возникающие при броске игральной кости. Будем считать, что все элементарные исходы равновероятны, то есть $\mathrm{P}(1) = \mathrm{P}(2) = \mathrm{P}(3) = \mathrm{P}(4) = \mathrm{P}(5) = \mathrm{P}(6) = \frac{1}{6}$. Тогда вероятность события $A = \{2, 4, 6\}$ — «>выпало чётное число> равна $\mathrm{P}(A) = \mathrm{P}(2)+\mathrm{P}(4)+\mathrm{P}(6) = \frac{1}{6}+\frac{1}{6}+\frac{1}{6} = \frac{1}{2}$.

Рассмотренный пример мотивирует нас ввести параллельные определения для дискретного пространства. Дискретным вероятностным пространством мы будем называть пару (Ω, P), где $\Omega = \{\omega_k\}_{k\in\mathbb{N}}$ — не более чем счётное множество (также называемое пространством элементарных исходов), а $\mathrm{P}\colon \Omega \to \mathbb{R}$ — неотрицательная функция, удовлетворяющая свойству $\sum\limits_{k\in\mathbb{N}} \mathrm{P}(\omega_k) = 1$. Говорят, что в этом случае на Ω заданы вероятности элементарных событий и что функция P задаёт на Ω распределение вероятностей. Событиями называются подмножества Ω. Вероятностью события $A \subset \Omega$ называется величина

$$\mathrm{P}(A) = \sum_{\omega\in A} P(\omega),$$

которую мы также будем обозначать буквой P. Последнее данное определение корректно, поскольку ряд в правой части сходится абсолютно.

Предложение 1.1. *Пусть* (Ω, P) *— дискретное вероятностное пространство в смысле последнего определения. Пусть* $\mathrm{P}\colon 2^\Omega \to \mathbb{R}$ *— функция, сопоставляющая событию его вероятность. Тогда тройка* $(\Omega, 2^\Omega, \mathrm{P})$ *является вероятностным пространством в смысле исходного определения.*

Доказательство. Множество 2^Ω является σ-алгеброй, поэтому достаточно проверить, что функция P удовлетворяет аксиомам вероятностной меры.

Из определения P имеем

$$\mathrm{P}(\Omega) = \sum_{i=1}^{+\infty} \mathrm{P}(\omega_i) = 1.$$

Пусть $A, B \subset \Omega$ и $A \cap B = \varnothing$. Положим $A = \{\omega_i\}_{i\in I_A}$, $B = \{\omega_i\}_{i\in I_B}$ и $A \sqcup B = \{\omega_i\}_{i\in I_{A\sqcup B}}$. Поскольку A и B не пересекаются, то $I_A \sqcup I_B = I_{A\sqcup B}$. Тогда, так как ряды в формуле ниже сходятся абсолютно, имеем

$$\mathrm{P}(A \sqcup B) = \sum_{i\in I_{A\sqcup B}} \omega_i = \sum_{i\in I_A} \omega_i + \sum_{i\in I_B} \omega_i = \mathrm{P}(A) + \mathrm{P}(B).$$

Пусть теперь $\{A_k\}_{k\in\mathbb{N}}$ — счётное семейство непересекающихся подмножеств множества Ω. Положим $A_k = \{\omega_i\}_{i\in I_k}$, $A = \bigsqcup\limits_{k\in I} A_k$. Снова, поскольку A_k попарно не пересекаются, то $I = \bigsqcup\limits_{k\in\mathbb{N}} I_k$. Поскольку все ряды ниже сходятся абсолютно, то выполнены равенства

$$\mathrm{P}(A) = \sum_{i\in I}\mathrm{P}(\omega_i) = \sum_{k\in\mathbb{N}}\sum_{i\in I_k}\mathrm{P}(\omega_i) = \sum_{k\in\mathbb{N}}\mathrm{P}(A_k).$$

□

Пусть $A, B \in \mathfrak{F}$ — события. Введём основные операции над событиями и приведём их классические наименования и обозначения в теории вероятностей.

Событие $\Omega\setminus A$ называется дополнением к событию A и обозначается $\overline{A}$ («событие A не произошло»).

Событие $A\cup B$ называется суммой событий A и B и обозначается $A+B$ («произошло событие A или B»). В курсе лекций это обозначение использовалось для случаев, когда $A\cap B = \varnothing$.

Событие $A\cap B$ называется произведением событий A и B и обозначается AB («произошло и событие A и событие B»).

События Ω и $\varnothing$ называются достоверным и невозможным, соответственно.

Если $AB = \varnothing$, то события A и B называются несовместными. («события A и B не происходят одновременно»).

Предложение 1.2 (Начальные свойства вероятностной меры)**.** *Пусть $A, B, A_k \in \mathfrak{F}$ — события. Тогда имеет место следующее:*

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

(2) *если* $A\subset B$, *то* $\mathrm{P}(B\setminus A) = \mathrm{P}(B) - \mathrm{P}(A)$;

(3) *если* $A\subset B$, *то* $\mathrm{P}(A) \leqslant \mathrm{P}(B)$;

(4) $\mathrm{P}(A\cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(AB)$;

(5) $\mathrm{P}(A\cup B) \leqslant \mathrm{P}(A) + \mathrm{P}(B)$;

(6) $\mathrm{P}(\bigcup\limits_{k=1}^{n} A_k) = \sum\limits_{k=1}^{n}(-1)^{k-1}\sum\limits_{i_1<i_2<\ldots<i_k}\mathrm{P}(A_{i_1}\ldots A_{i_k})$;

(7) $\mathrm{P}\left(\bigcup\limits_{k=1}^{+\infty} A_k\right) \leqslant \sum\limits_{k=1}^{+\infty}\mathrm{P}(A_k)$ *(это свойство называется субаддитивностью).*

Доказательство. Равенство (1) следует из цепочки

$$1 = \mathrm{P}(\Omega) = \mathrm{P}(A\sqcup\overline{A}) = \mathrm{P}(A) + \mathrm{P}(\overline{A}).$$

Равенство (2) — из цепочки

$$\mathrm{P}(B) = \mathrm{P}(A \cup (B \setminus A)) = \mathrm{P}(A) + \mathrm{P}(B \setminus A).$$

Неравенство (3) следует из этого равенства и неотрицательности вероятности.

Равенство (4) — из цепочки

$$\mathrm{P}(A \cup B) = \mathrm{P}((A \setminus B) \sqcup (A \cap B) \sqcup (B \setminus A)) =$$

$$= \mathrm{P}(A \setminus B) + \mathrm{P}(A \cap B) + \mathrm{P}(B \setminus A) + \mathrm{P}(A \cap B) - \mathrm{P}(A \cap B) =$$
$$= \mathrm{P}((A \setminus B) \sqcup (A \cap B)) + \mathrm{P}((B \setminus A) \sqcup (A \cap B)) - \mathrm{P}(A \cap B) =$$
$$= \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B).$$

Неравенство (5) немедленно следует из равенства (4).

Докажем (6) по индукции.

База $n = 2$ была доказана в пункте 3.

Докажем шаг. Положим $B = \bigcup\limits_{k=1}^{n-1} A_k$. По базе индукции

$$\mathrm{P}(B \cup A_n) = \mathrm{P}(B) + \mathrm{P}(A_n) - \mathrm{P}(BA_n).$$

Далее, положим $B_k = A_k A_n$. Тогда $BA_n = \bigcup\limits_{k=1}^{n-1} B_k$. По индукционному предположению вероятность $\mathrm{P}(B \cup A_n)$ равна

$$\sum_{k=1}^{n-1} (-1)^{k-1} \sum_{i_1 < i_2 < \ldots < i_k} \mathrm{P}(A_{i_1} \ldots A_{i_k}) + \mathrm{P}(A_n) - \left(\sum_{k=1}^{n-1} (-1)^{k-1} \sum_{i_1 < i_2 < \ldots < i_k} \mathrm{P}(A_{i_1} \ldots A_{i_k} A_n) \right) =$$

$$= \sum_{k=1}^{n} (-1)^{k-1} \sum_{i_1 < i_2 < \ldots < i_k} \mathrm{P}(A_{i_1} \ldots A_{i_k}).$$

Докажем неравенство (7). Положим $B_k = A_k \setminus \bigcup\limits_{i=1}^{k-1} A_i$. Тогда $\bigcup\limits_{k=1}^{+\infty} B_k = \bigcup\limits_{k=1}^{+\infty} A_k$, причём B_k попарно не пересекаются и $\mathrm{P}(B_k) \leqslant \mathrm{P}(A_k)$ по (2). Тогда по σ-аддитивности имеем

$$\bigcup_{k=1}^{+\infty} A_k = \bigcup_{k=1}^{+\infty} B_k = \sum_{k=1}^{+\infty} \mathrm{P}(B_k) \leqslant \sum_{k=1}^{+\infty} \mathrm{P}(A_k).$$

$\square$

2 Условные вероятности, формула Байеса, независимость событий

2.1 Условная вероятность

В задачах бывает полезно рассмотреть вероятность того, что произойдёт некоторое событие B при условии, что произойдёт событие A. Пусть $\mathrm{P}(A) > 0$. Тогда вероятность $\mathrm{P}(B \mid A) = \frac{\mathrm{P}(AB)}{\mathrm{P}(A)}$ называется условной вероятностью события B при условии того, что событие A произойдёт с вероятностью $\mathrm{P}(A) > 0$. Вероятность $\mathrm{P}(B)$ также иногда называется априорной вероятностью события B.

Предложение 2.1. *Пусть* $(\Omega, \mathfrak{F}, \mathrm{P})$ *— вероятностное пространство. Пусть* $A \in \mathfrak{F}$ *— событие, удовлетворяющее условию* $\mathrm{P}(A) > 0$*. Тогда тройка*

$$(\Omega, \mathfrak{F}, \mathrm{P}|_A),$$

где $\mathrm{P}|_A(B) := \mathrm{P}(B \mid A) = \frac{\mathrm{P}(AB)}{\mathrm{P}(A)}$*, является вероятностным пространством.*

Доказательство. Достаточно проверить аксиомы вероятностной меры (аксиомы σ-аддитивной меры и равенство $\mathrm{P}|_A(\Omega) = 1$).

Так как обе величины $\mathrm{P}(AB)$ и $\mathrm{P}(A)$ неотрицательны (а последняя и вовсе положительна), то $\mathrm{P}(B \mid A) \leqslant 0$.

Справедливость упомянутого равенства выводится из определения условной вероятности:

$$\mathrm{P}|_A(\Omega) = \frac{\mathrm{P}(A\Omega)}{\mathrm{P}(A)} = \frac{\mathrm{P}(A)}{\mathrm{P}(A)} = 1.$$

Пусть $\{B_k\}_{k\in\mathbb{N}}$ — счётная последовательность попарно не пересекающихся элементов алгебры $\mathfrak{F}$. Тогда элементы последовательности $\{B_k \cap A\}_{k\in\mathbb{N}}$ также попарно не пересекаются. Тогда

$$\mathrm{P}|_A\left(\bigsqcup_{k=1}^{+\infty} B_k\right) = \frac{1}{\mathrm{P}(A)}\mathrm{P}\left(A \cap \bigsqcup_{k=1}^{+\infty} B_k\right) = \frac{1}{\mathrm{P}(A)}\mathrm{P}\left(\bigsqcup_{k=1}^{+\infty} AB_k\right) = \sum_{k=1}^{+\infty}\frac{\mathrm{P}(AB_k)}{\mathrm{P}(A)} = \sum_{k=1}^{+\infty}\mathrm{P}|_A(B_k).$$

□

Следствие 2.1. *Пусть* $A \in \mathfrak{F}$ *— событие, вероятность которого больше* 0*,* $B_1, B_2 \in \mathfrak{F}$*. Тогда справедливы следующие свойства*

(1) *если* $B_1 \supset A$*, то* $\mathrm{P}(B_1 \mid A) = 1$*;*

(2) $\mathrm{P}(B_1 \cup B_2 \mid A) = \mathrm{P}(B_1 \mid A) + \mathrm{P}(B_2 \mid A) - \mathrm{P}(B_1B_2 \mid A)$*;*

(3) *если* B_1 *и* B_2 *несовместны, то* $\mathrm{P}(B_1 + B_2 \mid A) = \mathrm{P}(B_1 \mid A) + \mathrm{P}(B_2 \mid A)$*.*

2.2 Формула полной вероятности и формула Байеса

Теперь мы покажем, как связаны условные вероятности с вероятностями произведений событий, как можно вычислять вероятность события, зная его условные вероятности для несовместных событий (формула полной вероятности) и как можно вычислить условную вероятность «с переставленными причиной и следствием» (формула Байеса).

Лемма 2.2. *Пусть* $A, B \in \mathfrak{F}$ *— события и* $\mathrm{P}(A), \mathrm{P}(B) > 0$*. Тогда имеют место равенства*

$$\mathrm{P}(AB) = \mathrm{P}(A \mid B)\,\mathrm{P}(B) = \mathrm{P}(B \mid A)\,\mathrm{P}(A),$$

$$\mathrm{P}(A \mid B) = \frac{\mathrm{P}(B|A)\,\mathrm{P}(A)}{\mathrm{P}(B)}.$$

Доказательство. Первое равенство немедленно следует из определения условной вероятности, второе – немедленно из первого и предположения, что $\mathrm{P}(B) > 0$. □

Второе равенство, доказанное в лемме иногда (особенно в школьных программах), называют формулой Байеса. Ниже, пользуясь этим простым свойством, мы докажем более общую формулу и в дальнейшем будем называть формулой Байеса её.

Теорема 2.3 (Формула произведения вероятностей)**.** *Пусть* $A_1, \ldots, A_n \in \mathfrak{F}$ *— события. Если вероятности событий* $A_2A_3\ldots A_n, \ldots, A_{n-1}A_n, A_n$ *не равны нулю, то имеет место формула*

$$\mathrm{P}(A_1A_2\ldots A_n) = \mathrm{P}(A_1 \mid A_2\ldots A_n)\,\mathrm{P}(A_2 \mid A_3\ldots A_n)\ldots \mathrm{P}(A_{n-1} \mid A_n)\,\mathrm{P}(A_n).$$

Если вероятности событий $A_1A_2\ldots A_n, A_1A_2\ldots A_{n-1}, \ldots, A_1$ *не равны нулю, то имеет место формула*

$$\mathrm{P}(A_1A_2\ldots A_n) = \mathrm{P}(A_n \mid A_1\ldots A_{n-1})\,\mathrm{P}(A_{n-1} \mid A_1\ldots A_{n-2})\ldots \mathrm{P}(A_2 \mid A_1)\,\mathrm{P}(A_1).$$

Доказательство. Докажем первое утверждение индукцией по n, второе получается из первого перестановкой индексов в обратном порядке.

База: $n = 2$ есть определение условной вероятности.

Докажем шаг индукции. Пусть для $n-1$ утверждение выполнено. Положим $B = A_2A_3\ldots A_n$. Тогда по базе индукции (здесь мы пользуемся тем, что $\mathrm{P}(B) > 0$) и затем по индукционному предположению (а здесь всеми остальными условиями) имеем

$$\mathrm{P}(A_1B) = \mathrm{P}(A_1 \mid B)\,\mathrm{P}(B) = \mathrm{P}(A_1 \mid A_2A_3\ldots A_n)\,\mathrm{P}(A_2 \mid A_3\ldots A_n)\ldots \mathrm{P}(A_{n-1} \mid A_n)\,\mathrm{P}(A_n).$$

□

Набор событий $A_1, \ldots, A_n \in \mathfrak{F}$ называется разбиением пространства Ω (или просто «разбиение Ω»), если $\mathrm{P}(A_i) > 0$ для каждого i, A_i попарно несовместны ($A_iA_j = \varnothing$ при $i \neq j$) и $A_1 + A_2 + \ldots + A_n = \Omega$.

Теорема 2.4 (Формула полной вероятности)**.** *Пусть* $A_1, \ldots, A_n \in \mathfrak{F}$ *— разбиение* Ω. *Тогда для всякого события* B *имеет место равенство формула полной вероятности*

$$\mathrm{P}(B) = \sum_{i=k}^{n} \mathrm{P}(B \mid A_k)\,\mathrm{P}(A_k).$$

Доказательство. Так как события A_k попарно несовместны, то события $A_k B$ также попарно несовместны. Имеем цепочку равенств

$$\mathrm{P}(B) = \mathrm{P}(B\Omega) = \mathrm{P}(B(A_1 + \ldots + A_n)) = \mathrm{P}(BA_1 + \ldots + BA_n) \overset{\text{несовместность}}{=}$$

$$\overset{\text{несовместность}}{=} \sum_{k=1}^{n} \mathrm{P}(BA_k) = \sum_{k=1}^{n} \mathrm{P}(B \mid A_k)\,\mathrm{P}(A_k).$$

□

Формула полной вероятности остаётся справедливой, если отказаться от требования $A_1 + A_2 + \ldots + A_n = \Omega$ и заменить его на условие $B \subset A_1 + \ldots + A_n$ (сохраняя требования попарной несовместности событий A_i и $\mathrm{P}(A_i) > 0$).

Теорема 2.5 (Формула Байеса)**.** *Пусть события* $A_1, \ldots, A_n \in \mathfrak{F}$ *образуют разбиение* Ω, *пусть* $B \in \mathfrak{F}$ *— ещё одно событие и* $\mathrm{P}(B) > 0$. *Тогда справедлива формула Байеса*

$$\mathrm{P}(A_i \mid B) = \tfrac{\mathrm{P}(B|A_i)\,\mathrm{P}(A_i)}{\sum\limits_{k=1}^{n} \mathrm{P}(B|A_k)\,\mathrm{P}(A_k)}.$$

Доказательство. По лемме 2.2 («простейшая формула Байеса») имеем равенство

$$\mathrm{P}(A_i \mid B) = \tfrac{\mathrm{P}(B|A_i)\,\mathrm{P}(B)}{\mathrm{P}(B)}.$$

По формуле полной вероятности имеем

$$\mathrm{P}(B) = \sum_{i=k}^{n} \mathrm{P}(B \mid A_k)\,\mathrm{P}(A_k),$$

откуда следует искомая формула. □

2.3 Независимость событий

Интуиция говорит нам, что события A и B «независимы», когда от того с какой вероятностью произойдёт событие A не зависит вероятность того, что произойдёт событие B и наоборот. Математически это выражается формулами $\mathrm{P}(B \mid A) = \mathrm{P}(B)$ и $\mathrm{P}(A \mid B) = \mathrm{P}(A)$. Чтобы не ограничиваться случаями, когда вероятности событий больше 0, мы определим независимость следствием формул выше. События A и B называются независимыми, если справедливо равенство $\mathrm{P}(AB) = \mathrm{P}(A)\,\mathrm{P}(B)$.

Предложение 2.2 (Свойства независимости)**.** *Имеют место следующие утверждения:*

(1) *если* $\mathrm{P}(B) > 0$*, то независимость* A *и* B *равносильна равенству* $\mathrm{P}(A \mid B) = \mathrm{P}(A)$*;*

(2) *если* A *и* B *независимы, то* $\overline{A}$ *и* B *независимы;*

(3) *если события* B_1 *и* B_2 *несовместны,* A *и* B_1 *независимы, а также* A *и* B_2 *независимы, то* A *и* $B_1 + B_2$ *независимы.*

Доказательство. Проверим (1). Если $\mathrm{P}(B) > 0$, то по независимости имеем

$$\mathrm{P}(A \mid B) = \tfrac{\mathrm{P}(AB)}{\mathrm{P}(B)} = \tfrac{\mathrm{P}(A)\,\mathrm{P}(B)}{\mathrm{P}(B)} = \mathrm{P}(A).$$

Обратно, если $\mathrm{P}(A \mid B) = \mathrm{P}(A)$, то $\frac{\mathrm{P}(AB)}{\mathrm{P}(B)} = \mathrm{P}(A)$, откуда $\mathrm{P}(AB) = \mathrm{P}(A)\,\mathrm{P}(B)$.

Для доказательства (2) выпишем цепочку равенств

$$\mathrm{P}(\overline{A}B) = \mathrm{P}((\Omega \setminus A)B) = \mathrm{P}(B \setminus AB) = \mathrm{P}(B) - \mathrm{P}(AB) \overset{\text{независимость}}{=}$$

$$\overset{\text{независимость}}{=} \mathrm{P}(B) - \mathrm{P}(A)\,\mathrm{P}(B) = \mathrm{P}(B)(1 - \mathrm{P}(A)) = \mathrm{P}(B)\,\mathrm{P}(\overline{A}).$$

Наконец, для доказательства (3) заметим, что события B_1A и B_2A несовместны. Тогда

$$\mathrm{P}((B_1 + B_2)A) = \mathrm{P}(B_1A + B_2A) = \mathrm{P}(B_1A) + \mathrm{P}(B_2A) = \mathrm{P}(B_1)\,\mathrm{P}(A) + \mathrm{P}(B_2)\,\mathrm{P}(A) =$$

$$= (\mathrm{P}(B_1) + \mathrm{P}(B_2))\,\mathrm{P}(A) = \mathrm{P}(B_1 + B_2)\,\mathrm{P}(A).$$

□

Теперь определим независимость для набора событий. Пусть $B_1, \ldots, B_n \in \mathfrak{F}$ — события. Будем говорить, что они попарно независимы, если для всяких двух индексов $i \neq j$ выполнено равенство $\mathrm{P}(B_iB_j) = \mathrm{P}(B_i)\,\mathrm{P}(B_j)$ (то есть B_i и B_j независимы). Будем называть эти события независимыми, если для всякого набора индексов $i_1 < \ldots < i_k$ (здесь $2 \leqslant k \leqslant n$) имеет место равенство

$$\mathrm{P}\left(\bigcap_{s=1}^{k} B_{i_s}\right) = \prod_{s=1}^{k} \mathrm{P}(B_{i_s}).$$

Предложение 2.3. *Если события* $B_1, \ldots, B_n$ *независимы, то они попарно независимы.*

Пример. Вообще говоря из попарной независимости не следует независимость, что демонстрируется следующим примером. Рассмотрим тетраэдр, три грани которого покрашены в красный, зелёный и синий цвета, соответственно, а последняя разбита на три треугольника, покрашенных в те же цвета. Пусть вероятности выпадения граней равны $\frac{1}{4}$. покажем, что события «выпала грань с цветом A», где A — цвет попарно независимы, но не являются таковыми в совокупности. Формально ситуация выглядит следующим образом $\Omega = \{\omega_R, \omega_G, \omega_B, \omega_{RGB}\}$ — элементарное событие — выпала грань с данной раскраской. По условию $\mathrm{P}(\omega_R) = \mathrm{P}(\omega_G) = \mathrm{P}(\omega_B) = \mathrm{P}(\omega_{RGB}) = \frac{1}{4}$. Обозначим через $R = \{\omega_R, \omega_{RGB}\}$ ($G = \{\omega_G, \omega_{RGB}\}, B = \{\omega_B, \omega_{RGB}\}$) события «выпала грань с красным (зелёным, синим) цветом», соответственно. Тогда $\mathrm{P}(R) = \mathrm{P}(G) = \mathrm{P}(B) = \frac{1}{2}$, $\mathrm{P}(RG) = \mathrm{P}(GB) = \mathrm{P}(BR) = \frac{1}{4} = \mathrm{P}(R)\,\mathrm{P}(G) = \mathrm{P}(G)\,\mathrm{P}(B) = \mathrm{P}(B)\,\mathrm{P}(R)$, но $\mathrm{P}(RGB) = \mathrm{P}(\omega_{RGB}) = \frac{1}{4} \neq \mathrm{P}(R)\,\mathrm{P}(G)\,\mathrm{P}(B) = \frac{1}{8}$.

Пример. Пользуясь примером выше, можно показать, что условие несовместности в пункте (3) предложения 2.2 нельзя опустить. Положим $A = R$ и $B_1 = G$ и $B_2 = B$. Тогда $\mathrm{P}((B_1 \cup B_2)A) = \mathrm{P}(\omega_{RGB}) = \frac{1}{4}$, но $\mathrm{P}(A) = \frac{1}{2}$, $\mathrm{P}(B_1 + B_2) = \frac{3}{4}$ и $\frac{1}{2} \cdot \frac{3}{4} \neq \frac{1}{4}$. Таким образом, события $B_1 + B_2$ и A не являются независимыми.

Пример. Покажем, что из условия независимости нельзя удалить ни одно из равенств. Более того, мы докажем, что для всякого натурального n и семейства наборов индексов $S_J = \{(i_{1,j}, \dots, i_{k_j,j})\}_{j \in J}$ можно построить пример вероятностного пространства $(\Omega, \mathfrak{F}, \mathrm{P})$ и событий $A_1, \dots, A_n \in \mathfrak{F}$ для которых множество наборов, на которых выполнены равенства

$$\mathrm{P}\left(\bigcap_{s=1}^{k} B_{i_s}\right) = \prod_{s=1}^{k} \mathrm{P}(B_{i_s})$$

в точности совпадает с J.

Построим пример для дискретного вероятностного пространства. Положим $\Omega = \{(\varepsilon_1, \dots, \varepsilon_n) |\, \varepsilon_i \in \{0,1\}\}$ — множество всех кортежей из нулей и единиц длины n, $\mathrm{P}((\varepsilon_1, \dots, \varepsilon_n)) = p_{(\varepsilon_1, \dots, \varepsilon_n)}$ — будущее распределение вероятностей. Также положим $A_k = \{(\varepsilon_1, \dots, \varepsilon_n) |\, \varepsilon_i \in \{0,1\}, \varepsilon_k = 1\}$. Рассмотрим отображение $\varphi\colon \mathbb{R}^{2^n} \to \mathbb{R}^{2^n}$, заданное в некоторых фиксированных базисах этих пространств по правилу

$$\varphi\colon \begin{pmatrix} \dots \\ p_{(\varepsilon_1, \dots, \varepsilon_n)} \\ \dots \end{pmatrix} \mapsto \begin{pmatrix} \dots \\ \mathrm{P}(A_{i_1} \dots A_{i_k}) \\ \dots \end{pmatrix},$$

где для $k = 0$ предполагается, что в матрице стоит $\mathrm{P}(\Omega)$. Поскольку $\mathrm{P}(A_{i_1} \dots A_{i_k}) = \mathrm{P}(\{(\varepsilon_1, \dots, \varepsilon_n) |\, \varepsilon_i \in \{0,1\}, \varepsilon_{i_s} = 1, 1 \leqslant s \leqslant k\}) = \sum\limits_{\varepsilon_{i_s} = 1, 1 \leqslant s \leqslant k} p_{(\varepsilon_1, \dots, \varepsilon_n)}$, то φ — линейное отображение. Можно показать, что φ сюръективно (проверьте с помощью элементарных преобразований, что его матрица имеет ранг 2^n) и, следовательно, биективно. Таким образом, достаточно подобрать значения вероятностей все возможных произведений A_i так, чтобы вероятности $p_{(\varepsilon_1, \dots, \varepsilon_n)}$ были неотрицательны, в сумме давали 1 ($\mathrm{P}(\Omega) = 1$) и при этом выполнялись в точности все желаемые равенства на вероятности произведений событий A_i. Положим $\mathrm{P}(A_i) = \frac{1}{2^{2n}}$, $\mathrm{P}(\Omega) = 1$. Если $(i_1, \dots, i_k) \in S_J$, то положим $\mathrm{P}(A_{i_1} \dots A_{i_k}) = \frac{1}{2^{2kn}}$. Иначе положим $\mathrm{P}(A_{i_1} \dots A_{i_k}) = \frac{1}{2^{2kn+1}}$. Проверим, что имеют место неравенство $\frac{1}{2^{2kn+2}} \leqslant p_{(\varepsilon_1, \dots, \varepsilon_n)} \leqslant \frac{1}{2^{2kn}}$, для кортежей с $k > 0$ числом единиц. Для кортежа $(1, \dots, 1)$ неравенство выполнено по построению. Докажем неравенства для оставшихся кортежей с данным условием индукцией по числу нулей в кортеже. Пусть в текущем кортеже $(\varepsilon_1, \ldots, \varepsilon_n)$ присутствует $n - k \geqslant 1$ нулей. Прибавим к $p_{(\varepsilon_1,\ldots,\varepsilon_n)}$ все остальные значения вероятностей элементарных исходов — кортежей, в которых некоторые нули из данного кортежа заменены на единицы. Тогда $p_{(\varepsilon_1,\ldots,\varepsilon_n)} \leqslant \frac{1}{2^{2kn}}$, так как по предположению индукции все остальные слагаемые положительны, а сумма не превосходит $\frac{1}{2^{2kn}}$. С другой стороны, $p_{(\varepsilon_1,\ldots,\varepsilon_n)} \geqslant \frac{1}{2^{2kn+1}} - \frac{2^k-1}{2^{2(k+1)n}} = \frac{1}{2^{2kn+1}} - \frac{1}{2^{2kn+2+(2n-k-2)}}$. Так как $n - k - 1 \leqslant 0$ и $n - 1 \leqslant 0$, то последнее слагаемое по модулю не превосходит $\frac{1}{2^{2kn+2}}$. Тогда $p_{(\varepsilon_1,\ldots,\varepsilon_n)}$ не меньше $\frac{1}{2kn+2}$. Остаётся убедиться в том, что $p_{(0,\ldots,0)} = 1 - \sum\limits_{(\varepsilon_1,\ldots,\varepsilon_n)\neq(0,\ldots,0)} p_{(\varepsilon_1,\ldots,\varepsilon_n)} \geqslant 1 - \frac{2^n-1}{2^{2n}} > 0$.

3 Случайные величины, их распределения, функции распределения и плотности

НУЖНО: вписать все определения

НУЖНО: ввести функцию распределения

НУЖНО: определить распределение как прямой образ вероятностной меры

НУЖНО: доказать, что прямой образ вероятностной меры и мера Лебега-Стилтьеса, порождённая функцией распределения совпадают

НУЖНО: ввести понятие абсолютно непрерывной случайной величины и её плотности

4 Классические примеры распределений

НУЖНО: вписать описания для всех классических распределений

Дискретные распределения.

4.1 Распределение константы

4.2 Распределение Бернулли

4.3 Дискретное равномерное распределение

4.4 Биномиальное распределение

4.5 Распределение Пуассона

4.6 Геометрическое распределение

4.7 Гипергеометрическое распределение

4.8 Отрицательное биномиальное распределение

Абсолютно непрерывные случайные величины

4.9 Равномерное распределение

4.10 Экспоненциальное (показательное) распределение

4.11 Нормальное распределение (распределение Гаусса)

4.12 Распределение Коши

5 Численные характеристики случайных величин

НУЖНО: записать определения и свойства, описать ковариацию как скалярное произведение

НУЖНО: доказать формулы для вычисленя матожидания через интегралы Лебега, Лебега-Стилтьеса и интеграл Римана для абсолютно непрерывной случайно величины

6 Сходимости случайных величин

НУЖНО: записать определения всех сходимостей и вывод одних сходимостей из других

7 Производящие функции

НУЖНО: записать определение

8 Характеристические функции

Теорема 8.1 (Бохнер, Хинчин).

9 Предельные теоремы

НУЖНО: дописать ниже доказательства теорем

9.1 Неравенства

9.2 Закон больших чисел

Теорема 9.1 (Закон больших чисел в форме Бернулли).

Теорема 9.2 (Закон больших чисел в форме Чебышёва).

Теорема 9.3 (Усиленный закон больших чисел).

Теорема 9.4 (Закон больших чисел в форме Хинчина). *content*

9.3 Теорема Муавра-Лапласа

Теорема 9.5 (Теорема Пуассона).

Теорема 9.6 (Формула Стирлинга).

Теорема 9.7 (Муавр, Лапласа).

9.4 Закон нуля или единицы

Лемма 9.8 (Борель, Кантелли).

Лемма 9.9 (Борель, Кантелли).

Теорема 9.10 (Закон нуля или единицы Колмогорова).

9.5 Закон повторного логарифма

Теорема 9.11 (Закон повторного логарифма).

9.6 Закон арксинуса

Теорема 9.12 (Закон арксинуса).

9.7 Правило трёх сигм

Теорема 9.13 (Правило трёх сигм).

9.8 Центральная предельная теорема

Теорема 9.14 (Центральная предельная теорема).

Теорема 9.15 (Оценка Берри-Эссена).

10 Совместные распределения случайных величин

d

11 Свёртки случайных величин

d

12 Указатель терминов

d

13 Указатель теорем

d

Список литературы

[1] Колмогоров А.Н., Фомин С.В. *Элементы теории функций и функционального анализа*, Физматлит, 2004, 572с.

[2] Дьяченко М. И., Ульянов П. Л. *Мера и интеграл*, Факториал, 1998, 160с.

[3] Боровков А. А. *Теория вероятностей*, Физматлит, 1986, 432с.