

Теория вероятностей

(Ещё не)алгебраист

15 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	4
0.1	Обозначения	4
0.2	Предварительные сведения из действительного анализа	4
0.2.1	Системы множеств и структуры на них	4
0.2.2	Минимальное кольцо и минимальная алгебра	6
0.2.3	Мера на полукольце и её продолжение на минимальное кольцо	7
0.2.4	Лебеговское продолжение меры	8
0.2.5	Единственность продолжения меры на минимальную σ -алгебру	9
0.2.6	Непрерывность и полнота меры	11
0.2.7	Мера Лебега-Стилтьеса	12
0.2.8	Измеримые функции	13
0.2.9	Интеграл Лебега	16
0.2.10	Прямой образ меры (pushforward measure)	18
0.3	Теория категорий и взгляд на измеримые пространства с её точки зрения	19
0.3.1	Категория измеримых пространств	19
0.3.2	Прямой образ σ -алгебры	20
0.3.3	Обратный образ σ -алгебры	21

0.3.4	Связь между минимальной σ -алгеброй, прямым и обратным образами σ -алгебры	21
0.4	Произведение в категории измеримых пространств	23
0.4.1	Функтор борелевской σ -алгебры	24
0.5	Предварительные сведения из анализа Фурье	25
0.6	Предварительные сведения из линейной алгебры	25
0.6.1	Билинейные функции и квадратичные формы	25
0.6.2	Полуторалинейные функции	26
0.7	Элементарная комбинаторика	27
0.7.1	Элементарные факты для подсчёта числа комбинаций	27
0.7.2	Классические комбинаторные величины	28
0.7.3	Свойства комбинаторных величин	29
1	Вероятностное пространство, случайные события	31
2	Условные вероятности, формула Байеса, независимость событий	34
2.1	Условная вероятность	34
2.2	Формула полной вероятности и формула Байеса	35
2.3	Независимость событий	37
2.4	Произведение вероятностных пространств	40
3	Случайные величины, их распределения, функции распределения и плотности	41
4	Классические примеры распределений	46
4.1	Дискретные распределения.	46
4.1.1	Распределение константы	46
4.1.2	Распределение Бернулли	47
4.1.3	Дискретное равномерное распределение	47
4.1.4	Биномиальное распределение	47
4.1.5	Распределение Пуассона	48
4.1.6	Геометрическое распределение	48
4.1.7	Гипергеометрическое распределение	49
4.1.8	Отрицательное биномиальное распределение	49
4.2	Абсолютно непрерывные случайные величины	50
4.2.1	Равномерное распределение	50
4.2.2	Экспоненциальное (показательное) распределение	51
4.2.3	Нормальное распределение (распределение Гаусса)	52
4.2.4	Распределение Коши	52
4.3	Сингулярные распределения	53
5	Совместные распределения случайных величин	54

6	Независимость случайных величин	56
6.1	Независимость конечного набора случайных величин	56
6.2	Последовательность независимых случайных величин	59
7	Численные характеристики случайных величин	60
7.1	Математическое ожидание, моменты и абсолютные моменты	60
7.2	Моменты и абсолютные моменты случайной величины	64
7.3	Дисперсия, ковариация и корреляция	64
7.4	Связь с независимостью случайных величин	65
7.5	Мода, медиана	65
7.6	Квантили, асимметрия и эксцесс	65
7.7	Вычисление матожиданий и дисперсий некоторых случайных величин	66
8	Сходимости случайных величин	66
8.1	Сходимость почти наверное	66
8.2	Сходимость по вероятности	66
8.3	Пространство \mathcal{L}_p и сходимость в нём	66
8.4	Сходимость по распределению	66
8.5	Связь сходимостей	66
9	Производящие функции	66
10	Характеристические функции	66
11	Предельные теоремы	67
11.1	Неравенства	67
11.2	Закон больших чисел	67
11.3	Теорема Муавра-Лапласа	67
11.4	Закон нуля или единицы	67
11.5	Закон повторного логарифма	67
11.6	Закон арксинуса	67
11.7	Правило трёх сигм	68
11.8	Центральная предельная теорема	68
12	Свёртки случайных величин	68
13	Указатель терминов	68
14	Указатель теорем	68

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;
- $E\xi$ — математическое ожидание случайной величины ξ ;
- $D\xi$ — дисперсия случайной величины ξ ;
- $\text{Cov}(\xi, \eta)$ — 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 Системы множеств и структуры на них

Система множеств (следует понимать как синоним термина «семейство множеств») S называется **полукольцом**, если она удовлетворяет следующим аксиомам:

- (1) $\emptyset \in S$;
- (2) $\forall A, B \in S : A \cap B \in S$;
- (3) $\forall A, B \in S, A \subset B \exists n \in \mathbb{N} \exists C_1, \dots, C_n \in S : A = B \sqcup \bigsqcup_{k=1}^n C_k$.

Множество $\Omega \in U$ называется **единицей системы множеств U** , если всякий элемент $A \in U$ является подмножеством Ω .

Система множеств R называется **кольцом**, если она удовлетворяет следующим аксиомам:

- (1) $\forall A, B \in R : A \cap B \in R$;
- (2) $\forall A, B \in R : A \Delta B \in R$.

Следующее утверждение проверяется непосредственно, исходя из теоретико-множественных тождеств, но его доказательство приведено, например, в книге [2].

Предложение 0.1. Пусть R — кольцо. Тогда R является полукольцом. Кроме того, для любых элементов $A, B \in R$ в R также содержатся их объединение $A \cup B$ и разность $A \setminus B$.

Кольцо называется **σ -кольцом**, если для любого счётного набора его элементов $\{A_k\}_{k \in \mathbb{N}} \subset R$ их объединение содержится в R ($\bigcup_{k \in \mathbb{N}} A_k \in R$) и **δ -кольцом**, если для любого счётного набора его элементов $\{A_k\}_{k \in \mathbb{N}} \subset R$ их пересечение содержится в R .

Кольцо с единицей Ω называется **алгеброй (подмножеств множества Ω)**.

В книгах по теории вероятностей понятие алгебры часто вводится с использованием другого равносильного набора аксиом, что выражает следующее

Предложение 0.2 (Определение алгебры в традиции теории вероятностей). Система множеств R является алгеброй подмножеств множества Ω тогда и только тогда, когда R удовлетворяет следующим аксиомам

- (1) $\Omega \in R$;
- (2) $\forall A, B \in R : A \cup B, A \cap B \in R$;
- (3) $\forall A \in R : \Omega \setminus A := \bar{A} \in R$.

Мы снова опускаем доказательство, сводящееся к тождествам теории множеств.

Если алгебра является σ -кольцом или δ -кольцом, то её называют **σ -алгеброй** или **δ -алгеброй**, соответственно.

Предложение 0.3. Имеет место следующее:

- (1) Всякое σ -кольцо является δ -кольцом, обратное вообще говоря не верно.
- (2) Всякая σ -алгебра является δ -алгеброй и наоборот.

Лемма 0.4. Пусть R — (σ) -кольцо и $A \in R$. Тогда множество

$$R \cap A := \{B \cap A \mid B \in R\}$$

является (σ) -алгеброй подмножеств A . Также $R \cap A \subset R$.

Доказательство. По построению $\Omega \cap A = A$ содержится в $R \cap A$ и всякий элемент $R \cap A$ есть подмножество A . Так как кольцо замкнуто относительно пересечений, то $R \cap A \subset R$.

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

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

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

принадлежность справедлива в силу того, что $\bigcup_{i=1}^{\infty} B_k \in R$. □

Можно показать, что кольцо множеств является кольцом в алгебраическом смысле этого слова с операциями сложения Δ и умножения \cap , а алгебра множеств является булевой алгеброй (в частности, \mathbb{F}_2 -алгеброй).

0.2.2 Минимальное кольцо и минимальная алгебра

Следующее утверждение сводится к проверке аксиом кольца или алгебры, но его доказательство также можно прочесть в книге [2].

Предложение 0.5. Пусть $\{R_\alpha\}_{\alpha \in A}$ — семейство (σ, δ) -колец множеств. Тогда система $R = \bigcap_{\alpha \in A} R_\alpha$ является (σ, δ) -кольцом. Кроме того, если все кольца R_α являются (σ) -алгебрами подмножеств множества Ω (то есть у них есть общая единица), то R также является (σ) -алгеброй подмножеств множества Ω .

Теорема 0.6 (О минимальной (σ) -алгебре). Пусть U — система множеств. Тогда существует как минимум одно (σ) -кольцо, содержащее U . Пересечение всех таких (σ) -колец $R(U)$ ($R_\sigma(U)$) само является (σ) -кольцом. Всякое (σ) -кольцо, содержащее U , содержит и $R(U)$. Если $\Omega \in U$ — единица U , то $R(U)$ ($R_\sigma(U)$) является (σ) -алгеброй подмножеств множества Ω .

Доказательство. В качестве (σ) -кольца, содержащего U можно взять булеан 2^Σ , где множество Σ определено как объединение $\bigcup_{A \in U} A$.

Пересечение всех таких (σ) -колец существует, поскольку имеется хотя бы одно кольцо и по предложению 0.5 это пересечение само является (σ) -кольцом.

Пусть (σ) -кольцо R' содержит множество U . Тогда по построению $R(U)$ ($R_\sigma(U)$) содержится в пересечении $2^\Sigma \cap R'$, откуда $R(U) \subset R'$ ($R_\sigma(U) \subset R'$).

Если Ω — единица U , то по построению $\Sigma = \Omega$. Для всякого (σ) -кольца R' , содержащего U имеем $\Omega \in R$ и по лемме 0.4 система множеств $R' \cap \Omega \subset R'$ является (σ) -алгеброй подмножеств Ω . Так как Ω являлось единицей U , то U содержится в $R' \cap \Omega$. Следовательно, достаточно рассматривать пересечение только (σ) -алгебр подмножеств множества Ω , содержащих U . По предложению 0.5 их пересечение является (σ) -алгеброй подмножеств Ω . □

Опираясь на предложение 0.6 дадим определение. Для системы множеств U пересечение всех колец, содержащих U называется **минимальным кольцом, порождённым U** и обозначается $R(U)$.

Предложение 0.7. Пусть S — полукольцо и $R(S)$ — минимальное кольцо, порождённое S . Тогда $R(S)$ допускает следующие описания

- $R(S) = \{A_1 \cup \dots \cup A_n \mid n \in \mathbb{N}, \{A_i\}_{i=1}^n \subset S\};$
- $R(S) = \{A_1 \sqcup \dots \sqcup A_n \mid n \in \mathbb{N}, \{A_i\}_{i=1}^n \subset S\}.$

Доказательство. **НУЖНО:** дописать доказательство или сослаться □

0.2.3 Мера на полукольце и её продолжение на минимальное кольцо

Пусть S — некоторое полукольцо. Будем называть неотрицательную функцию $m: S \rightarrow \mathbb{R}$ **мерой** на полукольце S , если m удовлетворяет аксиоме аддитивности

$$\forall A, B \in S, A \cap B = \emptyset, A \cup B \in S : m(A \cup B) = m(A) + m(B).$$

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

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

то мера m называется **σ -аддитивной** (аксиома σ -аддитивности). Подразумевается, что ряд в правой части сходится для любой последовательности попарно не пересекающихся $\{A_k\}$, объединение которых лежит в S . Можно показать, что из этой аксиомы следует, что $m(\emptyset) = 0$ и поэтому из неё следует аксиома аддитивности.

Пользуясь предложением 0.7 введём функцию $\nu: R(S) \rightarrow \mathbb{R}$ по правилу $\nu(A_1 \sqcup \dots \sqcup A_n) = \sum_{i=1}^n m(A_i)$, где $A_i \in S$. Следующее предложение позволяет назвать ν **продолжением меры m с полукольца S на его минимальное кольцо.**

Предложение 0.8. Справедливо следующее

- (1) функция ν определена корректно, то есть значение ν не зависит от выбора представления $A_1 \sqcup \dots \sqcup A_n$;
- (2) функция ν является мерой на кольце $R(S)$;
- (3) ограничение функции ν на полукольцо S совпадает с m ;
- (4) если мера m была σ -аддитивной, то функция ν также является σ -аддитивной.

Доказательство. См. доказательство в [2, Глава 1, Теорема 2.2]. □

Лемма 0.9. Пусть S — полукольцо и $m: S \rightarrow \mathbb{R}$ — мера на S . Тогда m удовлетворяет следующим свойствам

(1) если для $A, B \in S$ выполнено $A \subset B$, то $m(A) \leq m(B)$;

(2) если $A, A_1, \dots, A_n \in S$ и $A \subset \bigcup_{i=1}^n A_i$, то

$$m(A) \leq \sum_{i=1}^n m(A_i);$$

(3) если $A_1, \dots, A_n \subset S$ — попарно не пересекающиеся множества и $\bigcup_{i=1}^n A_i \subset A \in S$, то

$$\sum_{i=1}^n m(A_i) \leq m(A);$$

(4) если $\{A_i\}_{i=1}^{+\infty} \subset S$ — попарно не пересекающиеся множества и $\bigcup_{i=1}^{+\infty} A_i \subset A \in S$, то

$$\sum_{i=1}^{+\infty} m(A_i) \leq m(A).$$

Доказательство. **НУЖНО:** дописать доказательство или сослаться □

Последнее свойство, доказанное в лемме, называют **полуаддитивностью**.

0.2.4 Лебеговское продолжение меры

Далее будем рассматривать полукольцо S с единицей Ω и σ -аддитивной мерой m . Пусть $\nu: R(S) \rightarrow \mathbb{R}$ — продолжение этой меры на минимальное кольцо.

Введём функцию **внешней меры** $\mu^*: 2^\Omega \rightarrow \mathbb{R}$, заданную по правилу

$$\mu^*(A) := \inf_{A \subset \bigcup_{i=1}^{+\infty} B_i, B_i \in S} \sum_{i=1}^{+\infty} m(B_i).$$

Предложение 0.10. Для всякого $A \subset \Omega$ в определении внешней меры можно заменить дизъюнктные объединения на произвольные:

$$\mu^*(A) = \inf_{A \subset \bigcup_{i=1}^{+\infty} B_i, B_i \in S} \sum_{i=1}^{+\infty} m(B_i).$$

Доказательство. **НУЖНО:** дописать доказательство или сослаться □

Множество $A \subset \Omega$ называется **измеримым**, если для любого $\varepsilon > 0$ найдётся множество $B \in R(S)$ такое, что $\mu^*(A \Delta B) < \varepsilon$. Если A измеримо, то его **мерой** называется значение $\mu(A) := \mu^*(A)$. Обозначим через \mathcal{M} системы всех измеримых подмножеств единицы Ω .

Лемма 0.11. Пусть $\{A_i\}_{i=1}^{+\infty} \subset \mathcal{M}$ — последовательность множеств, $A \in \mathcal{M}$ и $A \subset \bigcup_{i=1}^n A_i$. Тогда

$$\mu^*(A) \leq \sum_{i=1}^n \mu^*(A_i);$$

Доказательство. НУЖНО: дописать доказательство или сослаться □

Теорема 0.12. Система измеримых множеств \mathcal{M} является алгеброй.

Доказательство. НУЖНО: дописать доказательство или сослаться □

Теорема 0.13. Функция μ на алгебре множеств \mathcal{M} является мерой.

Доказательство. НУЖНО: дописать доказательство или сослаться □

Теорема 0.14. Алгебра измеримых множеств \mathcal{M} является σ -алгеброй.

Доказательство. НУЖНО: дописать доказательство или сослаться □

Теорема 0.15. Мера μ на σ -алгебре измеримых множеств \mathcal{M} является σ -аддитивной.

Доказательство. НУЖНО: дописать доказательство или сослаться □

Ограничение внешней меры μ^* на σ -алгебру измеримых подмножеств \mathcal{M} мы будем называть **лебеговским продолжением меры μ** .

0.2.5 Единственность продолжения меры на минимальную σ -алгебру

Прежде, чем приступить к доказательству теорема Каратеодори мы докажем лемму, характеризующую минимальную σ -алгебру, порождённую данной алгеброй.

Лемма 0.16. Пусть \mathcal{A} — некоторая алгебра. Тогда существует наименьшая по включению система множеств $\text{Mon}(\mathcal{A})$, удовлетворяющая свойствам

- (1) $\mathcal{A} \subset \text{Mon}(\mathcal{A})$;
- (2) если имеется последовательность множеств $\{A_i\} \subset \text{Mon}(\mathcal{A})$ и либо $A_1 \subset A_2 \subset \dots$, либо $A_1 \supset A_2 \supset \dots$, то либо $\bigcup_{i=1}^{+\infty} A_i \in \text{Mon}(\mathcal{A})$, либо $\bigcap_{i=1}^{+\infty} A_i \in \text{Mon}(\mathcal{A})$, соответственно (мы будем называть это свойство **монотонностью**).

Более того, система множеств $\text{Mon}(\mathcal{A})$ совпадает с минимальной σ -алгеброй $R_\sigma(\mathcal{A})$, порождённой \mathcal{A} .

Доказательство. Из определения σ -алгебры следует, что $R_\sigma(\mathcal{A})$ удовлетворяет свойствам из условия. Кроме того, пересечение произвольного семейства систем множеств, удовлетворяющих данным условиям снова будет удовлетворять им. Поэтому взяв пересечение всех таких систем множеств, мы получим наименьшую по включению систему $Mon(\mathcal{A})$. Так как всегда можно ограничиться системами множеств с той же единицей, что была в \mathcal{A} , то $Mon(\mathcal{A})$ обладает единицей.

Из сказанного выше следует, что $Mon(\mathcal{A})$ содержится в σ -алгебре $R_\sigma(\mathcal{A})$. Мы докажем, что система $Mon(\mathcal{A})$ сама является σ -алгеброй. Отсюда по минимальности (теорему 0.6) будет следовать обратное включение.

Для множества $B \in Mon(\mathcal{A})$ обозначим через

$$L(B) = \{A \in Mon(\mathcal{A}) \mid A \setminus B, B \setminus A, A \cup B \in Mon(\mathcal{A})\}$$

систему множеств из $Mon(\mathcal{A})$, разности и объединение которых с B снова лежат в $Mon(\mathcal{A})$. По построению $A \in L(B)$ равносильно тому, что $B \in L(A)$.

Рассмотрим случай, когда система множеств $L(B)$ непуста и в ней содержится вложенную последовательность $\{A_i\} \subset L(B)$ то есть либо $A_1 \subset A_2 \subset \dots$, либо $A_1 \supset A_2 \supset \dots$. Тогда из того, что $Mon(\mathcal{A})$ удовлетворяет свойству монотонности, то $\bigcup_{i=1}^{+\infty} A_i \in L(B)$, либо $\bigcap_{i=1}^{+\infty} A_i \in L(B)$, соответственно. Поэтому система множеств $L(B)$ тоже удовлетворяет свойству монотонности.

Пусть теперь $A \in \mathcal{A}$. Тогда $\mathcal{A} \subset L(A)$, то есть $L(A)$ удовлетворяет обоим свойствам из условия. Отсюда по минимальности $Mon(\mathcal{A}) \subset L(A)$. Тогда для любого $B \in Mon(\mathcal{A})$ имеем $\mathcal{A} \subset L(B)$. Поэтому, аналогично, $Mon(\mathcal{A}) \subset L(B)$, то есть $Mon(\mathcal{A}) = L(B)$ для любого $B \in Mon(\mathcal{A})$. Следовательно, система множеств $Mon(\mathcal{A})$ является алгеброй.

Из монотонности следует, что $Mon(\mathcal{A})$ является σ -алгеброй. □

Мы приведём доказательство теоремы Каратеодори для случая конечной меры, он отметим, что она остаётся верной и для σ -конечной меры.

Теорема 0.17 (Каратеодори). *Пусть \mathcal{A} — алгебра подмножеств множества Ω и $\nu: \mathcal{A} \rightarrow \mathbb{R}$ — конечная мера на \mathcal{A} . Тогда существует единственная σ -аддитивная мера μ на минимальной σ -алгебре $R_\sigma(\mathcal{A})$, ограничение которой на \mathcal{A} совпадает с ν .*

Доказательство. Для доказательства существования рассмотрим лебеговское продолжение меры ν на σ -алгебру измеримых подмножеств \mathcal{M} . По минимальности (теореме 0.6) имеем включение $R_\sigma(\mathcal{A}) \subset \mathcal{M}$. Тогда ограничение лебеговского продолжения на $R_\sigma(\mathcal{A})$ будет являться σ -аддитивной мерой.

Докажем единственность. Предположим, что μ_1 и μ_2 — два продолжения меры ν . Положим $\mathcal{A}_\pm = \{B \in R_\sigma(\mathcal{A}) \mid \mu_1(B) = \mu_2(B)\} \subset R_\sigma(\mathcal{A})$ — система множеств, на которых меры μ_1 и μ_2 совпадают.

Проверим, что \mathcal{A}_\pm удовлетворяет условиям леммы 0.16. По предположению о совпадении ограничений μ_1 и μ_2 на алгебре \mathcal{A} с мерой t имеем включение $\mathcal{A} \subset \mathcal{A}_\pm$. Остаётся проверить только второе условие.

Пусть имеется вложенная последовательность $\{A_i\} \subset \mathcal{A}_=$, то есть либо $A_1 \subset A_2 \subset \dots$, либо $A_1 \supset A_2 \supset \dots$. По непрерывности мер в обоих случаях имеем равенства

$$\begin{aligned}\mu_1 \left(\bigcup_{i=1}^{+\infty} A_i \right) &= \lim_{i \rightarrow +\infty} \mu_1(A_i) = \lim_{i \rightarrow +\infty} \mu_2(A_i) = \mu_2 \left(\bigcup_{i=1}^{+\infty} A_i \right), \\ \mu_1 \left(\bigcap_{i=1}^{+\infty} A_i \right) &= \lim_{i \rightarrow +\infty} \mu_1(A_i) = \lim_{i \rightarrow +\infty} \mu_2(A_i) = \mu_2 \left(\bigcap_{i=1}^{+\infty} A_i \right).\end{aligned}$$

Тогда $\mathcal{A}_=$ удовлетворяет условию монотонности и по лемме 0.16 содержит (поскольку мы не проверили минимальность лемма даёт только такой результате) σ -алгебру $R_\sigma(\mathcal{A})$.

Из всего сказанного следует равенство $\mathcal{A}_= = R_\sigma(\mathcal{A})$. \square

0.2.6 Непрерывность и полнота меры

Мера μ на кольце R называется **непрерывной**, если для любой последовательности вложенных подмножеств $\{A_i\}_{i=1}^{+\infty} \subset R$, $A_1 \supset A_2 \supset A_3 \supset \dots$, пересечения которых $A = \bigcap_{i=1}^{\infty} A_i$ лежит в R , имеет место равенство

$$\lim_{i \rightarrow +\infty} \mu(A_i) = \mu(A) = \mu \left(\bigcap_{i=1}^{\infty} A_i \right).$$

Предложение 0.18. Конечная мера μ на кольце R непрерывна тогда и только тогда, когда она σ -аддитивна.

Доказательство. **НУЖНО:** дописать доказательство \square

Следствие 0.19. Пусть задана мера μ на кольце R и последовательность вложенных подмножеств $\{A_i\}_{i=1}^{+\infty} \subset R$, $A_1 \supset A_2 \supset A_3 \supset \dots$, пересечение которых $A = \bigcap_{i=1}^{\infty} A_i$ лежит в R и, кроме того, мера $\mu(A_1) < +\infty$. Тогда имеет место равенство

$$\lim_{i \rightarrow +\infty} \mu(A_i) = \mu(A) = \mu \left(\bigcap_{i=1}^{\infty} A_i \right).$$

Доказательство. **НУЖНО:** дописать доказательство \square

Предложение 0.20. Пусть задана конечная σ -аддитивная мера μ на кольце R . Тогда для любой последовательности вложенных подмножеств $\{A_i\}_{i=1}^{+\infty} \subset R$, $A_1 \subset A_2 \subset A_3 \subset \dots$, объединение которых $A = \bigcup_{i=1}^{\infty} A_i$ лежит в R , имеет место равенство

$$\lim_{i \rightarrow +\infty} \mu(A_i) = \mu(A) = \mu \left(\bigcup_{i=1}^{\infty} A_i \right).$$

Доказательство. **НУЖНО:** дописать доказательство □

Мера μ на кольце R называется **полной**, если для любого множества $A \in R$ из равенства нулю меры $\mu(A) = 0$ вытекает, что всякое подмножество $B \subset A$ лежит в R . Из свойств меры в этом случае $\mu(B) = 0$.

0.2.7 Мера Лебега-Стилтьеса

Пусть $f: \mathbb{R} \rightarrow \mathbb{R}$ — неубывающая ограниченная функция. Можно рассматривать и случай, когда функция не ограничена и для неё развить теорию, используя σ -конечную меру, но мы не будем сталкиваться с такими случаями в дальнейшем и поэтому остановимся на рассмотрении ограниченной функции f . Рассмотрим полукольцо S полуинтервалов вида $(a, b]$ или $(a; +\infty), (-\infty; b]$ и введём на нём функцию m_f по правилу $m_f((a, b]) = f(b) - f(a)$, $m_f((a; +\infty)) = \lim_{x \rightarrow +\infty} f(x) - f(a)$, $m_f((-\infty; b]) = f(b) - \lim_{x \rightarrow -\infty} f(x)$ и $\mu_f(\mathbb{R}) = \lim_{x \rightarrow +\infty} f(x) - \lim_{x \rightarrow -\infty} f(x)$.

Предложение 0.21. *Функция m_f является мерой на полукольце S . Функция f непрерывна справа тогда и только тогда, когда мера m_f является σ -аддитивной.*

Доказательство. Докажем, что m_f удовлетворяет аксиоме аддитивности. Пусть $(a, b] = \bigsqcup_{i=1}^n (a_i, b_i]$. Можно считать, что $a = a_1 < b_1 = a_2 < b_2 < \dots < b_n = b$. Тогда

$$m_f((a, b]) = f(b) - f(a) = \sum_{i=1}^n (f(b_i) - f(a_i)) = \sum_{i=1}^n m_f((a_i, b_i]).$$

Доказательство для бесконечных интервалов аналогично, нужно лишь заменить числа a или b на соответствующие пределы.

Предположим, что f непрерывна справа. Пусть $(a, b] = \bigsqcup_{i=1}^{+\infty} (a_i, b_i]$. По **полуаддитивности**

$$\sum_{i=1}^{+\infty} m_f((a_i, b_i]) \leq m_f((a, b]).$$

Фиксируем $\varepsilon > 0$. По непрерывности справа найдём такое $a' \in (a, b]$, что $F(a') - F(a) < \frac{\varepsilon}{2}$ и такие $b'_i > b_i$, чтобы $F(b'_i) - F(b_i) < \frac{\varepsilon}{2^{i+1}}$. Тогда имеем

$$[a', b] \subset (a, b] = \bigsqcup_{i=1}^{+\infty} (a_i, b_i] \subset \bigsqcup_{i=1}^{+\infty} (a_i, b'_i].$$

Поскольку отрезок $[a', b]$ компактен, то из его покрытия $\{(a_i, b'_i)\}$ можно выбрать конечное подпокрытие $\{(a_{i_k}, b'_{i_k})\}_{k=1}^n$. Тогда

$$m_f((a, b]) \leq m_f((a', b]) + \frac{\varepsilon}{2} \leq \sum_{k=1}^n m_f((a_{i_k}, b'_{i_k})) + \frac{\varepsilon}{2} \leq \sum_{i=1}^{+\infty} m_f((a_i, b'_i]) + \varepsilon.$$

Отсюда следует равенство $\sum_{i=1}^{+\infty} m_f((a_i, b_i]) = m_f((a, b])$.

Случай с бесконечными концами сводится к случаю конечного полуинтервала путём выбора точки c такой, что $F(c) - F(-\infty) < \varepsilon$ или $F(+\infty) - F(c) < \varepsilon$ и рассмотрения пересечения полуинтервалов разбиения с $(-\infty; c]$ или $(c; +\infty)$. \square

Мерой Лебега-Стилтьеса (порождённой функцией f) мы назовём лебеговское продолжение меры m_f и обозначим её через μ_f .

0.2.8 Измеримые функции

Пусть \mathfrak{F} — σ -алгебра подмножеств Ω . Назовём отображение $f: \Omega \rightarrow \mathbb{R}$ **измеримой функцией**, если для любого **борелевского множества** $B \subset \mathbb{R}$ его прообраз является элементом \mathfrak{F} (то есть f является **измеримым отображением** $f: (\Omega, \mathfrak{F}) \rightarrow (\mathbb{R}, \mathcal{B})$). Можно показать (см. лемму 0.39), что данное требование равносильно тому, что прообраз любого интервала или тому, что прообраз любого бесконечного полуинтервала $(-\infty, b]$ измерим. Будем говорить, что функция f , определённая на подмножестве $A \subset \Omega$, **измерима на A** , если также для любого борелевского множества B его прообраз $f^{-1}(B) \subset A$ лежит в \mathfrak{F} (это равносильно тому, что ограничение $f|_A$ является измеримым отображением $f|_A: (A, \mathfrak{F} \cap A) \rightarrow (\mathbb{R}, \mathcal{B})$).

Далее будем считать, что на σ -алгебре \mathfrak{F} задана σ -аддитивная мера μ и элементы \mathfrak{F} мы будем называть измеримыми множествами.

Предложение 0.22. Пусть функция f измерима и g — непрерывная на $\text{Im } f$ функция. Тогда композиция $g \circ f$ измерима.

Доказательство. Пусть U — открытое множество. Тогда $g^{-1}(U)$ открыто и, следовательно, $f^{-1}(g^{-1}(U))$ измеримо. \square

Предложение 0.23. Пусть f, g — измеримые функции. Тогда множество $A_{f \leq g} = \{x \in \Omega \mid f(x) \leq g(x)\}$ измеримо. Функции $a + f, af, |f|, f^2, f + g$ и fg , где a — константа, измеримы. Если функция g не принимает значения 0, то функции $\frac{1}{g}$ и $\frac{f}{g}$ измеримы.

Доказательство. Как было замечено в определении измеримой функции, для проверки измеримости функции h достаточно доказать, что для любой константы t множество $A_{h \leq t}$ измеримо.

Первое утверждение следует из представления

$$A_{f \leq g} = \bigcup_{q \in \mathbb{Q}} \{x \in \Omega \mid f(x) < q < g(x)\} = \bigcup_{q \in \mathbb{Q}} (\{x \in \Omega \mid f(x) < q\} \cap \{x \in \Omega \mid q < g(x)\}),$$

где все множества справа измеримы, так как измеримы функции f и g .

Измеримость функций $a + f, af, |f|, f^2, \frac{1}{g}$ следует из предложения 0.22.

По доказанному выше функция $a - g$ измерима, Тогда также по доказанному выше множество $A_{f \leq a-g}$ измеримо и, поэтому измерима функция $f + g$.

Измеримость функции fg следует из представления $fg = \frac{1}{4}(f+g)^2 - \frac{1}{4}(f-g)^2$ и доказанного выше. Из этого следует измеримость функции $\frac{f}{g}$. \square

Простой функцией мы будем называть измеримую функцию f такую, что множество значений f конечно и мера множества, на котором f принимает ненулевые значения конечна.

Обобщённой простой функцией мы будем называть произвольную измеримую функцию, принимающую не более чем счётное число значений. Всякая простая функция является обобщённой простой функцией.

Индикатором множества A мы будем называть функцию $\mathbb{1}_A$, заданную по правилу

$$\mathbb{1}_A(x) = \begin{cases} 1, & x \in A; \\ 0, & x \notin A. \end{cases}$$

Предложение 0.24. *Верно следующее.*

- (1) *Индикатор $\mathbb{1}_A$ измеримого множества A является простой функцией.*
- (2) *Всякая (обобщённая) простая функция единственным образом может быть представлена в виде конечной линейной комбинации (ряда) с попарно различными коэффициентами индикаторов попарно непересекающихся измеримых множеств, объединение которых равно Ω .*
- (3) *Всякая (обобщённая) простая функция единственным образом может быть представлена в виде линейной комбинации (ряда) с попарно различными ненулевыми коэффициентами индикаторов попарно непересекающихся измеримых множеств.*
- (4) *Всякая простая функция является обобщённой простой и её представления из пунктов выше как для простой и как для обобщённой простой функции совпадают.*
- (5) *Ряд из индикаторов с коэффициентами (не обязательно различными) попарно непересекающихся измеримых множеств является обобщённой простой функцией.*

Доказательство. Индикатор множества принимает всего два значения — 0 и 1. Так как A измеримо, то прообраз любого борелевского множества относительно индикатора есть одно из четырёх измеримых множеств: $\emptyset, A, \Omega \setminus A, \Omega$. Следовательно, $\mathbb{1}_A$ — простая функция.

Пусть $\text{Im } f = \{c_i\}_{i=1}^{n(+\infty)}$ — образ функции f . Множества $A_i = f^{-1}(\{c_i\})$ измеримы, так как f измерима, попарно не пересекаются и покрывают всё Ω . Тогда

$$f = \sum_{i=1}^{n(+\infty)} c_i \mathbb{1}_{A_i}.$$

Пусть имеется другое представление f в конечной суммы (ряда), удовлетворяющее условию пункта (2). Пусть множество B , индикатор которого входит в это представление не совпадает ни с одним из множеств A_i . Если B строго содержится в некотором A_i , то найдётся её одно множество C из второго разбиения, пересекающееся с A_i . Тогда функция f принимает на B и на C равные значения, что противоречит условию. Иначе B пересекается как минимум с двумя множествами A_i и функция f принимает на B не менее двух различных значений, что снова противоречит условию. Следовательно, данное представление единственно.

Чтобы получить представление для пункта (3) удалим из построенной суммы (ряда) слагаемое, соответствующее $c_i = 0$. Доказательство единственности аналогично.

Построенное выше представление для простой функции содержит только конечное число ненулевых слагаемых, поэтому выполнен пункт (4).

Функция, построенная в пункте (5) принимает счётное число различных значений, причём каждое конкретное значение — на измеримом множестве. Тогда прообраз всякого борелевского множества является объединением (не более, чем счётным) этих множеств (и, возможно, дополнения до их объединения) и поэтому измерим. \square

Предложение 0.25. *Сумму двух (обобщённых) простых функций является (обобщённой) простой функцией. Функция, пропорциональная (обобщённой) простой функции, сама является таковой. Модуль (обобщённой) простой функции, является таковой функцией.*

Доказательство. Сумма двух измеримых функций принимающих конечное (не более, чем счётное) число значений снова является измеримой функцией по предложению 0.23 и снова принимает конечное (не более, чем счётное) число значений.

Функция, пропорциональная измеримой, измерима и модуль измеримой функции измерим по предложению 0.23. Модуль функции, принимающей не более, чем счётное число значений и эта функция умноженная на константу также принимают не более, чем счётное число значений. \square

Предложение 0.26. *Для всякой измеримой функции f существует последовательность обобщённых простых функций $\{f_n\}$, равномерно сходящаяся к f .*

Доказательство. Поскольку f измерима, то для всяких натурального n и целого k множество $A_{n,k} = f^{-1}([\frac{k}{2^n}, \frac{k+1}{2^n}))$ измеримо и для фиксированного n и различных k эти множества образуют разбиение Ω . Положим

$$f_n(x) = \sum_{k \in \mathbb{Z}} \frac{k}{2^n} \mathbb{1}_{A_{n,k}}.$$

Тогда f_n удовлетворяет определению обобщённой простой функции. Кроме того, для всякого n и $x \in \Omega$ выполнены неравенства $0 \leq f(x) - f_n(x) \leq \frac{1}{2^n}$. Поэтому последовательность $\{f_n\}$ сходится к f равномерно. \square

0.2.9 Интеграл Лебега

Пусть $f = \sum_{i=1}^n c_i \mathbb{1}_{A_i}$, $c_i \neq 0$ — простая функция в представлении из предложения 0.24. Интегралом простой функции f на пространстве Ω по мере μ мы назовём сумму $\sum_{i=1}^n c_i \cdot \mu(A_i)$ и обозначим через $\int_{\Omega} f d\mu$.

Обобщённую простую функцию $f = \sum_{i=1}^{+\infty} c_i \mathbb{1}_{A_i}$ (в представлении из предложения 0.24) назовём интегрируемой по Лебегу на пространстве Ω по мере μ , если для $c_i \neq 0$ меры $\mu(A_i)$ конечны и ряд $\sum_{i=1}^{+\infty} c_i \cdot \mu(A_i)$ сходится абсолютно (предполагается, что $0 \cdot \mu(A) = 0$, даже если мера $\mu(A)$ бесконечна). В этом случае обозначим через $\int_{\Omega} f d\mu$ сумму ряда и назовём её интегралом обобщённой простой функции f на пространстве Ω по мере μ .

Из определения простой функции следует, что она является интегрируемой на пространстве Ω по мере μ обобщённой простой функцией.

Лемма 0.27. Пусть f, g — обобщённые простые функции. Тогда

- (1) если f, g — интегрируемые на Ω по мере μ , то функция $f + g$ интегрируема и имеет место формула

$$\int_{\Omega} f d\mu + \int_{\Omega} g d\mu = \int_{\Omega} (f + g) d\mu;$$

- (2) интегрируемость функции f равносильна интегрируемости функции $|f|$ и в случае интегрируемости справедливо неравенство

$$\left| \int_{\Omega} f d\mu \right| \leq \int_{\Omega} |f| d\mu;$$

- (3) если $|f| \leq g$ и g интегрируема, то $|f|$ интегрируема и выполнено неравенство

$$\int_{\Omega} |f| d\mu \leq \int_{\Omega} g d\mu.$$

Доказательство. Первые два свойства следуют из свойств сходимости абсолютно сходящихся рядов. Третье — из признака абсолютной сходимости. \square

Лемма 0.28. Пусть мера μ конечна на Ω . Тогда выполнены следующие утверждения.

- (1) Всякая ограниченная обобщённая простая функция интегрируема на Ω по μ .

(2) Интеграл $\int_{\Omega} 1_A d\mu$ от индикатора 1_A равен мере A .

Доказательство. Пусть константа C ограничивает функцию f . Так как мера Ω конечна, то функция $C \cdot 1_{\Omega}$ интегрируема, откуда по лемме 0.27 следует, что f интегрируема. Второе утверждение выполнено из построения интеграла обобщённой простой функции \square

Теорема 0.29. Пусть мера μ конечна на Ω . Пусть f — измеримая функция и существует последовательность интегрируемых обобщённых простых функций $\{f_n\}$, равномерно сходящаяся к f . Тогда

- (1) существует предел $\int_{\Omega} f_n d\mu =: I$;
- (2) для любой другой последовательности интегрируемых обобщённых простых функций $\{g_n\}$, равномерно сходящихся к f предел $\int_{\Omega} g_n d\mu$ также равен I ;
- (3) если f — обобщённая простая функция, то f интегрируема и $\int_{\Omega} f d\mu = I$.

Доказательство. Фиксируем $\varepsilon > 0$. Так как функции f_n равномерно сходятся к f , то начиная с некоторого N для $n, m > N$ и для любого $x \in \Omega$ выполнено неравенство $|f_n(x) - f_m(x)| < \varepsilon$. Тогда по леммам 0.28 и 0.28 имеем

$$\left| \int_{\Omega} f_n d\mu - \int_{\Omega} f_m d\mu \right| \leq \int_{\Omega} |f_n - f_m| d\mu \leq \varepsilon \mu(X).$$

По критерию Коши существует предел $\int_{\Omega} f_n d\mu =: I$.

Заметим, что если $f = 0$, то для $n > N(\varepsilon)$ выполнено равенство

$$\left| \int_{\Omega} f_n d\mu \right| < \varepsilon \mu(X).$$

Поэтому в данном случае $I = 0$. Поскольку, если последовательности f_n и g_n сходятся равномерно к f , то последовательность $f_n - g_n$ сходится равномерно к нулевой функции, то предел I будет одинаковым для обеих последовательностей.

Пусть f сама является обобщённой простой функцией. Найдётся индекс n такой, что для любого $x \in \Omega$ выполнено неравенство $|f(x) - f_n(x)| < 1$. По леммам 0.28 и 0.27 функция $f - f_n$ интегрируема. Так как f_n интегрируема, то снова по лемме 0.27 функция $f = (f - f_n) + f_n$ интегрируема. Теперь возьмём в качестве последовательности f_n постоянную последовательность из функции f . Тогда предел этой последовательности равен одновременно $\int_{\Omega} f d\mu$ и I . \square

Измеримая функция f на Ω с $\mu(\Omega) < +\infty$ называется **интегрируемой**, если существует последовательность интегрируемых обобщённых простых функций $\{f_n\}$, равномерно сходящаяся к f . **Интегралом (Лебега) f на Ω по мере μ** называется предел $\int_{\Omega} f_n d\mu$. В дальнейшем мы будем обозначать его через $\int_{\Omega} f d\mu$.

Следствие 0.30. Пусть мера μ конечна на Ω . Пусть измеримая функция f интегрируема на Ω по μ и существует последовательность обобщённых простых функций $\{f_n\}$, равномерно сходящаяся к f . Тогда начиная с некоторого номера N все функции в этой последовательности интегрируемы.

Доказательство. Рассмотрим некоторую последовательность интегрируемых обобщённых простых функций $\{f_n\}$, равномерно сходящуюся к f . Тогда последовательность простых функций $\{f_n - g_n\}$ равномерно сходится к тождественно нулевой функции. Следовательно, начиная с некоторого индекса N для всех $n > N$ и $x \in \Omega$ выполнено неравенство $|f_n(x) - g_n(x)| < 1$. Тогда по леммам 0.24 и 0.27 функция $g_n - f_n$ интегрируема. Тогда интегрируема и функция $g_n = (g_n - f_n) + f_n$. \square

0.2.10 Прямой образ меры (pushforward measure)

Пусть (X, \mathfrak{A}_X) и (Y, \mathfrak{A}_Y) — **измеримые пространства**, μ — мера на σ -алгебре \mathfrak{A}_X и $f: (X, \mathfrak{A}_X) \rightarrow (Y, \mathfrak{A}_Y)$ — **измеримое отображение**. Тогда **прямым образом меры μ при отображении f** называется функция $f_*\mu$, заданная по правилу $f_*\mu(A) = \mu(f^{-1}(A))$ для любого $A \in \mathfrak{A}_Y$.

Предложение 0.31. Прямой образ (σ -аддитивной) меры $f_*\mu$ является (σ -аддитивной) мерой на алгебре \mathfrak{A}_Y .

Доказательство. Пусть $\{A_k\}$ — последовательность попарно не пересекающихся множеств из \mathfrak{A}_Y . Тогда множества из последовательности $\{f^{-1}(A_k)\}$ также попарно не пересекаются. Из определения прямого образа меры $f_*\mu$ имеем цепочку равенств

$$f_*\mu\left(\bigsqcup_{k=1}^{+\infty} A_k\right) = \mu\left(f^{-1}\left(\bigsqcup_{k=1}^{+\infty} A_k\right)\right) = \mu\left(\bigsqcup_{k=1}^{+\infty} f^{-1}(A_k)\right) = \sum_{k=1}^{+\infty} \mu(f^{-1}(A_k)) = \sum_{k=1}^{+\infty} f_*\mu(A_k).$$

Доказательство аддитивности проводится аналогично. \square

Теорема 0.32. Пусть мера μ конечна на X . Рассмотрим измеримую функцию $g: (Y, \mathfrak{A}_Y) \rightarrow (\mathbb{R}, \mathcal{B})$. Тогда g интегрируема на Y по прямому образу меры $f_*\mu$ тогда и только тогда, когда композиция $g \circ f$ интегрируема на X по мере μ . В случае интегрируемости равны интегралы

$$\int_X (g \circ f) d\mu = \int_Y g d f_*\mu.$$

Доказательство. Из определения обобщённых простых функций g и $g \circ f$ одновременно являются или не являются простыми обобщёнными функциями.

Для обобщённой простой функции истинность утверждения следует из того, что интегралы $\int_X (g \circ f) d\mu$ и $\int_Y g df_* \mu$ равны сумме одного и того же ряда (поскольку $f_* \mu(g^{-1}(\{a\})) = \mu(f^{-1}(g^{-1}(\{a\}))) = \mu((g \circ f)^{-1}(\{a\}))$ для $a \in \mathbb{R}$).

Если существует последовательность интегрируемых обобщённых простых функций $\{g_n\}$ на Y , равномерно сходящаяся к g , то последовательность $\{g_n \circ f\}$ будет являться последовательностью интегрируемых обобщённых простых функций на X , равномерно сходящейся к $g \circ f$. Следовательно, из интегрируемости g следует интегрируемость $g \circ f$.

Обратно, пусть $g \circ f$ интегрируема. По предложению 0.26 существует последовательность обобщённых простых функций $\{g_n\}$ на Y , равномерно сходящаяся к g . Тогда последовательность обобщённых простых функций $\{g_n \circ f\}$ равномерно сходится к $g \circ f$. По следствию 0.30 начиная с некоторого номера N функции $g_n \circ f$ являются интегрируемыми. По доказанному выше, это означает, что для всех $n > N$ функции g_n интегрируемы, поэтому g интегрируема.

Пусть теперь g и $g \circ f$ интегрируемы. Фиксируем последовательность интегрируемых обобщённых простых функций $\{g_n\}$ на Y , равномерно сходящуюся к g . Имеем цепочку равенств

$$\int_X (g \circ f) d\mu = \lim_{n \rightarrow +\infty} \int_X (g_n \circ f) d\mu = \lim_{n \rightarrow +\infty} \int_Y g_n df_* \mu = \lim_{n \rightarrow +\infty} \int_Y g df_* \mu.$$

□

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

Теория категорий в её лучшем проявлении выражает собой формализацию понятия «математическая конструкция» через понятия объектов, морфизмов, функторов, естественных преобразований, а также формализует интуицию в виде универсальных свойств, сопряжённости функторов, эквивалентности категорий и так далее.

В этом подразделе будет предполагаться, что вы знакомы с определением категории (например, основанном на теории множеств) и знакомы с понятиями объекта, морфизма между объектами и функтором из одной категории в другую. Остальные определения по возможности будут приведены здесь.

0.3.1 Категория измеримых пространств

Пусть \mathfrak{A}_X — σ -алгебра подмножеств множества X . Пара (X, \mathfrak{A}_X) называется **измеримым пространством**.

Пусть (X, \mathfrak{A}_X) и (Y, \mathfrak{A}_Y) — пара измеримых пространств. **Морфизмом измеримых пространств (измеримым отображением)** $f: (X, \mathfrak{A}_X) \rightarrow (Y, \mathfrak{A}_Y)$ называется отображение множеств $f: X \rightarrow Y$ такое, что для любого $U \in \mathfrak{A}_Y$ выполнено $f^{-1}(U) \in \mathfrak{A}_X$, где f^{-1} — это полный прообраз.

В категории измеримых пространств \mathfrak{Meas} объектами являются измеримые пространства, а морфизмами — морфизмы измеримых пространств.

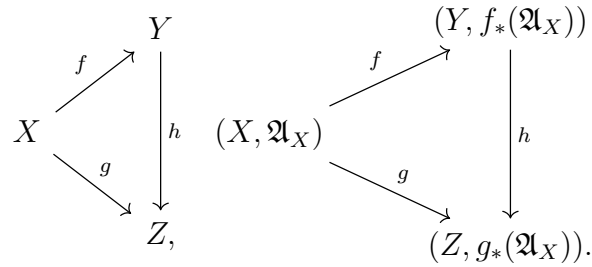
0.3.2 Прямой образ σ -алгебры

Пусть (X, \mathfrak{A}_X) — измеримое пространство и $f: X \rightarrow Y$ — отображение множеств. Существует способ естественным образом построить σ -алгебру подмножеств Y .

Положим $f_*(\mathfrak{A}_X) = \{B \subset Y \mid f^{-1}(B) \in \mathfrak{A}_X\}$ — все подмножества Y , полный прообраз которых лежит в \mathfrak{A}_X . Построенная система множеств называется **прямым образом** σ -алгебры \mathfrak{A}_X при отображении f .

Предложение 0.33. Конструкция f_* обладает следующими свойствами.

- (1) Система множеств $f_*(\mathfrak{A}_X)$ является σ -алгеброй подмножеств множества Y .
- (2) Отображение f является морфизмом измеримых пространств $f: (X, \mathfrak{A}_X) \rightarrow (Y, f_*(\mathfrak{A}_X))$.
- (3) Если \mathfrak{A}_Y — σ -алгебра подмножеств Y , то отображение f является морфизмом измеримых пространств $f: (X, \mathfrak{A}_X) \rightarrow (Y, \mathfrak{A}_Y)$ тогда и только тогда, когда $\mathfrak{A}_Y \subset f_*(\mathfrak{A}_X)$.
- (4) Пусть имеются отображения множеств $f: X \rightarrow Y, g: X \rightarrow Z$ и $h: Y \rightarrow Z$ такие, что $h \circ f = g$, то h является морфизмом измеримых пространств $h: (Y, f_*(\mathfrak{A}_X)) \rightarrow (Z, g_*(\mathfrak{A}_X))$.



Доказательство. Первый пункт проверяется непосредственно и следует из теоретико-множественных тождеств. Второй пункт — из построения $f_*(\mathfrak{A}_X)$, третий — из определения морфизма измеримых пространств.

Докажем четвёртый пункт. Пусть $U \in g_*(\mathfrak{A}_X)$. Тогда $g^{-1}(U) = f^{-1}(h^{-1}(U)) \in \mathfrak{A}_X$. По определению прямого образа $f_*(\mathfrak{A}_X)$ множество $h^{-1}(U)$ содержится в нём. Следовательно, h является морфизмом измеримых пространств. \square

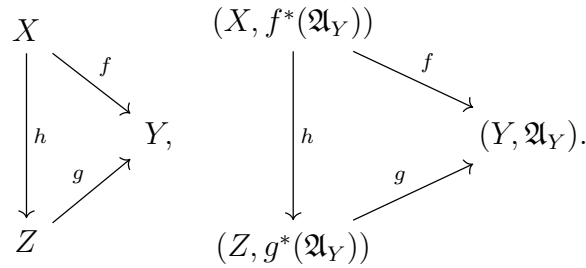
0.3.3 Обратный образ σ -алгебры

Пусть (Y, \mathfrak{A}_Y) — измеримое пространство и $f: X \rightarrow Y$ — отображение множеств. Существует способ естественным образом построить σ -алгебру подмножеств X .

Положим $f^*(\mathfrak{A}_Y) = \{f^{-1}(B) \mid B \in \mathfrak{A}_Y\}$ — полные прообразы всех элементов σ -алгебры \mathfrak{A}_Y . Построенная система множеств называется **обратным образом** σ -алгебры \mathfrak{A}_Y при отображении f .

Предложение 0.34. Конструкция f^* обладает следующими свойствами.

- (1) Система множеств $f^*(\mathfrak{A}_Y)$ является σ -алгеброй подмножеств множества X .
- (2) Отображение f является морфизмом измеримых пространств $f: (X, f^*(\mathfrak{A}_Y)) \rightarrow (Y, \mathfrak{A}_Y)$.
- (3) Если \mathfrak{A}_X — σ -алгебра подмножеств X , то отображение f является морфизмом измеримых пространств $f: (X, \mathfrak{A}_X) \rightarrow (Y, \mathfrak{A}_Y)$ тогда и только тогда, когда $f^*(\mathfrak{A}_Y) \subset \mathfrak{A}_X$.
- (4) Пусть имеются отображения множеств $f: X \rightarrow Y, g: Z \rightarrow Y$ и $h: X \rightarrow Z$ такие, что $g \circ h = f$, то h является морфизмом измеримых пространств $h: (X, f^*(\mathfrak{A}_Y)) \rightarrow (Z, g^*(\mathfrak{A}_Y))$.



Доказательство. Первый пункт проверяется непосредственно и следует из теоретико-множественных тождеств. Второй пункт — из построения $f^*(\mathfrak{A}_Y)$, третий — из определения морфизма измеримых пространств.

Докажем четвёртый пункт. Пусть $U \in g^*(\mathfrak{A}_Y)$. Тогда для некоторого $V \in \mathfrak{A}_Y$ выполнено $g^{-1}(V) = U$. Далее, $h^{-1}(U) = h^{-1}(g^{-1}(V)) = f^{-1}(V)$. По определению обратного образа $f^*(\mathfrak{A}_Y)$ множество $h^{-1}(U)$ содержится в нём. Следовательно, h является морфизмом измеримых пространств. \square

0.3.4 Связь между минимальной σ -алгеброй, прямым и обратным образами σ -алгебры

Лемма 0.35. Пусть X, Y — множества, $f: X \rightarrow Y$ — отображение множеств. Если \mathfrak{A}_X — σ -алгебра подмножеств X , то $f^*(f_*(\mathfrak{A}_X)) \subset \mathfrak{A}_X$. Если \mathfrak{A}_Y — σ -алгебра подмножеств Y , то $\mathfrak{A}_Y \subset f_*(f^*(\mathfrak{A}_Y))$.

Доказательство. По предложению 0.33 отображение f является морфизмом измеримых пространств $f: (X, \mathfrak{A}_X) \rightarrow (Y, f_*(\mathfrak{A}_X))$. Тогда по предложению 0.34 имеем включение $f^*(f_*(\mathfrak{A}_X)) \subset \mathfrak{A}_X$.

Теперь По предложению 0.34 отображение f является морфизмом измеримых пространств $f: (X, f^*(\mathfrak{A}_Y)) \rightarrow (Y, \mathfrak{A}_Y)$. Тогда по предложению 0.33 имеем включение $\mathfrak{A}_Y \subset f_*(f^*(\mathfrak{A}_Y))$. \square

Рассмотрим категорию $\sigma\text{-}\mathfrak{Alg}_X$ всех σ -алгебр с единицей X , в которой объектами выступают σ -алгебры подмножеств X , а единственный существующий морфизм из σ -алгебры \mathfrak{A} идёт в σ -алгебру \mathfrak{A}' , если $\mathfrak{A} \subset \mathfrak{A}'$. Построенные нами для отображения множеств $f: X \rightarrow Y$ конструкции f_* и f^* являются функторами между категориями $\sigma\text{-}\mathfrak{Alg}_X$ и $\sigma\text{-}\mathfrak{Alg}_Y$ и, более того, что эти функторы сопряжены.

Теорема 0.36. *Пусть $f: X \rightarrow Y$ — отображение множеств. Тогда $f_*: \sigma\text{-}\mathfrak{Alg}_X \rightarrow \sigma\text{-}\mathfrak{Alg}_Y$ и $f^*: \sigma\text{-}\mathfrak{Alg}_Y \rightarrow \sigma\text{-}\mathfrak{Alg}_X$ — функторы. Кроме того, функтор f^* является левым сопряжённым к функтору f_* (и обратно, функтор f_* является правым сопряжённым к f^*). Другими словами, для любого объекта $\mathfrak{A}_X \in \sigma\text{-}\mathfrak{Alg}_X$ и любого объекта $\mathfrak{A}_Y \in \sigma\text{-}\mathfrak{Alg}_Y$ и имеется естественный изоморфизм множеств $\text{Hom}(f^*(\mathfrak{A}_Y), \mathfrak{A}_X) \cong \text{Hom}(\mathfrak{A}_Y, f_*(\mathfrak{A}_X))$.*

Доказательство. Из построения f_* и f^* сохраняют отношения включения и равенства (и, следовательно, композицию включений), поэтому они функториальны.

Теперь по предложению 0.34 имеет место включение $f^*(\mathfrak{A}_Y) \subset \mathfrak{A}_X$ тогда и только тогда, когда $f: (X, \mathfrak{A}_X) \rightarrow (Y, \mathfrak{A}_Y)$ является измеримым отображением. По предложению 0.33 последнее равносильно включению $\mathfrak{A}_X \subset f_*(\mathfrak{A}_Y)$. Множества $\text{Hom}(f^*(\mathfrak{A}_Y), \mathfrak{A}_X)$ и $\text{Hom}(\mathfrak{A}_Y, f_*(\mathfrak{A}_X))$ одновременно либо пусты, либо состоят из одного элемента. Тогда пусть данный изоморфизм будет пустым в первом случае и сопоставляет единственный элемент единственному элементу во втором случае.

Если имеются включения $\mathfrak{A}_{X1} \subset \mathfrak{A}_{X2}$ и $\mathfrak{A}_{Y1} \subset \mathfrak{A}_{Y2}$, то возникает коммутативная диаграмма, индуцированная этими включениями, выражающая естественность изоморфизма

$$\begin{array}{ccccc}
 & & \text{Hom}(f^*(\mathfrak{A}_{Y2}), \mathfrak{A}_{X2}) & \xrightarrow{\quad} & \text{Hom}(\mathfrak{A}_{Y2}, f_*(\mathfrak{A}_{X2})) \\
 & \nearrow & \vdots & & \nearrow \\
 \text{Hom}(f^*(\mathfrak{A}_{Y2}), \mathfrak{A}_{X1}) & \xrightarrow{\quad} & \text{Hom}(\mathfrak{A}_{Y2}, f_*(\mathfrak{A}_{X1})) & & \\
 \downarrow & & \downarrow & & \downarrow \\
 & & \text{Hom}(f^*(\mathfrak{A}_{Y1}), \mathfrak{A}_{X2}) & \cdots \rightarrow & \text{Hom}(\mathfrak{A}_{Y1}, f_*(\mathfrak{A}_{X2})) \\
 & \nearrow & \vdots & & \nearrow \\
 \text{Hom}(f^*(\mathfrak{A}_{Y1}), \mathfrak{A}_{X1}) & \xrightarrow{\quad} & \text{Hom}(\mathfrak{A}_{Y1}, f_*(\mathfrak{A}_{X1})) & &
 \end{array}$$

\square

Лемма 0.37. *(О коммутировании конструкций минимальной σ -алгебры и обратного образа) Пусть X, Y — множества, T — система подмножеств Y , $Y \in T$.*

Пусть $f: X \rightarrow Y$ — отображение множеств. Обозначим через $f^{-1}(T)$ систему множеств $\{f^{-1}(U) \mid U \in T\}$. Тогда $R_\sigma(f^{-1}(T)) = f^*(R_\sigma(T))$.

Доказательство. Так как $T \subset R_\sigma(T)$, то по построению f^* имеем включение $f^{-1}(T) \subset f^*(R_\sigma(T))$. По предложению 0.34 система множеств $f^*(R_\sigma(T))$ является σ -алгеброй. Тогда по минимальности (теореме 0.6) имеем включение $R_\sigma(f^{-1}(T)) \subset f^*(R_\sigma(T))$.

Далее, система множеств $f_*(R_\sigma(f^{-1}(T)))$ является σ -алгеброй по предложению 0.33. По построению f_* имеем включение $T \subset f_*(R_\sigma(f^{-1}(T)))$. Снова по минимальности (теореме 0.6) имеем включение $R_\sigma(T) \subset f_*(R_\sigma(f^{-1}(T)))$.

По лемме 0.35 имеем включение $f^*(f_*(R_\sigma(f^{-1}(T)))) \subset R_\sigma(f^{-1}(T))$. Собирая вместе все включения и пользуясь тем, что $f^*(A) \subset f^*(B)$ для $A \subset B$ получаем

$$f^*(R_\sigma(T)) \subset f^*(f_*(R_\sigma(f^{-1}(T)))) \subset R_\sigma(f^{-1}(T)) \subset f^*(R_\sigma(T)),$$

откуда следует требуемое. \square

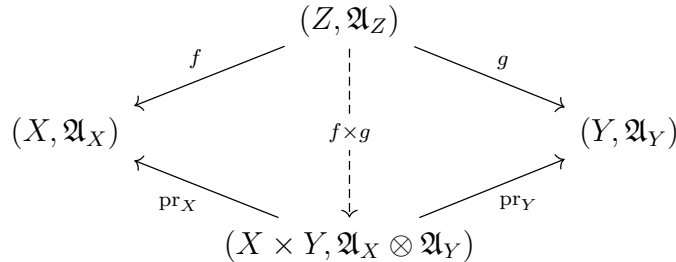
0.4 Произведение в категории измеримых пространств

Пусть (X, \mathfrak{A}_X) и (Y, \mathfrak{A}_Y) — измеримые пространства. Построим по ним новые измеримые пространства следующим образом.

Произведением измеримых пространств (X, \mathfrak{A}_X) и (Y, \mathfrak{A}_Y) мы назовём измеримое пространство $(X \times Y, \mathfrak{A}_X \otimes \mathfrak{A}_Y)$, где $\mathfrak{A}_X \otimes \mathfrak{A}_Y$ есть минимальная σ -алгебра, порождённая полукольцом всех возможных декартовых произведений $A_X \times A_Y$, где $A_X \in \mathfrak{A}_X$ и $A_Y \in \mathfrak{A}_Y$.

Теорема 0.38. Пусть (X, \mathfrak{A}_X) и (Y, \mathfrak{A}_Y) — измеримые пространства, $(X \times Y, \mathfrak{A}_X \otimes \mathfrak{A}_Y)$ — их произведение. Тогда

- (1) отображения проекций $\text{pr}_X: X \times Y \rightarrow X$ и $\text{pr}_Y: X \times Y \rightarrow Y$ являются морфизмами измеримых пространств.
- (2) пространство $(X \times Y, \mathfrak{A}_X \otimes \mathfrak{A}_Y)$ удовлетворяет универсальному свойству произведения, то есть для всякого измеримого пространства (Z, \mathfrak{A}_Z) и всяких двух морфизмов $f: (Z, \mathfrak{A}_Z) \rightarrow (X, \mathfrak{A}_X)$, $g: (Z, \mathfrak{A}_Z) \rightarrow (Y, \mathfrak{A}_Y)$ существует единственный морфизм $f \times g: (Z, \mathfrak{A}_Z) \rightarrow (X \times Y, \mathfrak{A}_X \otimes \mathfrak{A}_Y)$, удовлетворяющий условию $\text{pr}_X \circ (f \times g) = f$ и $\text{pr}_Y \circ (f \times g) = g$.



Доказательство. Пусть $A \in \mathfrak{A}_X$. Тогда $\text{pr}_X^{-1}(A) = A \times Y \in \mathfrak{A}_X \otimes \mathfrak{A}_Y$, поэтому отображение проекции на X является морфизмом. Аналогично доказывается для второй проекции.

Пусть теперь f, g — морфизмы из условия. Положим $(f \times g)(z) = (f(z), g(z))$ — единственный возможный вариант определить это отображение так, чтобы выполнялись условия на проекции. Проверим, что $f \times g$ является морфизмом. Пусть $A = A_X \times A_Y$, где $A_X \in \mathfrak{A}_X$ и $A_Y \in \mathfrak{A}_Y$. Тогда $(f \times g)^{-1}(A) = \{z \in Z \mid f(z) \in A_X, g(z) \in A_Y\} = f^{-1}(A_X) \cap g^{-1}(A_Y) \in \mathfrak{A}_Z$. Положим $S = \{A = A_X \times A_Y \mid A_X \in \mathfrak{A}_X, A_Y \in \mathfrak{A}_Y\}$. По определению $\mathfrak{A}_X \otimes \mathfrak{A}_Y = R_\sigma(S)$. По лемме 0.37 имеем $R_\sigma((f \times g)^{-1}(S)) = (f \times g)^*(R_\sigma(S)) = (f \times g)^*(\mathfrak{A}_X \otimes \mathfrak{A}_Y)$. Так как $(f \times g)^{-1}(S) \subset \mathfrak{A}_Z$, то по минимальности (по теореме 0.6) имеем включение $R_\sigma((f \times g)^{-1}(S)) \subset \mathfrak{A}_Z$. Тогда $(f \times g)^*(\mathfrak{A}_X \otimes \mathfrak{A}_Y) \subset \mathfrak{A}_Z$ и по свойствам обратного образа (см. предложение 0.34) отображение $f \times g$ является морфизмом. \square

0.4.1 Функтор борелевской σ -алгебры

Пусть (X, τ) — топологическое пространство. Минимальная σ -алгебра, порождённая системой открытых множеств τ называется **борелевской σ -алгеброй**, а её элементы называются **борелевскими множествами**. Мы будем обозначать её через $\mathcal{B}(\tau)$, а соответствующее измеримое пространство через $\text{Bor}((X, \tau)) = (X, \mathcal{B}(\tau))$. Мы докажем, что конструкция $\text{Bor}: \mathfrak{Top} \rightarrow \mathfrak{Meas}$, сопоставляющая топологическому пространству (X, τ) измеримое пространство $(X, \mathcal{B}(\tau))$, а непрерывному отображению f его же как отображение множеств, функториальна.

Лемма 0.39. Пусть $f: X \rightarrow Y$ — отображение множеств, θ — топология Y , \mathfrak{A}_X — σ -алгебра подмножеств X . Пусть также S — база топологии θ такая, что всякое открытое подмножество представляется в виде не более, чем счётного объединения элементов базы, и T — предбаза топологии θ такая, что всякое открытое множество представляется в виде не более, чем счётного объединения конечных пересечений элементов T . Тогда следующие утверждения равносильны

- (1) прообраз всякого элемента борелевской σ -алгебры $\mathcal{B}(\theta)$ лежит в \mathfrak{A}_X ;
- (2) прообраз всякого открытого множества лежит в \mathfrak{A}_X ;
- (3) прообраз всякого элемента базы S лежит в \mathfrak{A}_X ;
- (4) прообраз всякого элемента предбазы T лежит в \mathfrak{A}_X .

Доказательство. Все пункты являются частным случаем пункта (1), а пункты (3) и (4) — пункта (2). Так как база топологии (с данным дополнительным условием) является частным случаем предбазы топологии (с дополнительным условием), то достаточно вывести из пункта (4) пункт (1).

Рассмотрим систему множеств $T_X = f^{-1}(T) \cup \{X\} := \{f^{-1}(U) \mid U \in T\} \cup \{X\}$. По условию $T_X \subset \mathfrak{A}_X$. Пусть $R_\sigma(T_X)$ — минимальная σ -алгебра, содержащая

T_X . Так как \mathfrak{A}_X является σ -алгеброй, то $R_\sigma(T_X) \subset \mathfrak{A}_X$. Из построения f_* имеем $f_*(R_\sigma(T_X)) \subset f_*(\mathfrak{A}_X)$. Так же из построения f_* имеем включение $T \subset f_*(R_\sigma(T_X))$. Из условия наложенного на T следует, что минимальная σ -алгебра, порождённая T совпадает с $\mathcal{B}(\theta)$. Тогда по минимальности (теореме 0.6) имеем включение $\mathcal{B}(\theta) \subset f_*(R_\sigma(T_X))$. Следовательно, $\mathcal{B}(\theta) \subset f_*(\mathfrak{A}_X)$ и по предложению 0.33 f является морфизмом измеримых пространств $f: (X, \mathfrak{A}_X) \rightarrow (Y, \mathcal{B}(\theta))$, что и утверждается в пункте (1). \square

Теорема 0.40. Пусть $(X, \tau), (Y, \theta) — топологические пространства, $f: (X, \tau) \rightarrow (Y, \theta) — непрерывное отображение. Тогда отображение f является морфизмом измеримых пространств $f: (X, \mathcal{B}(\tau)) \rightarrow (Y, \mathcal{B}(\theta))$.$$

Доказательство. Следует из эквивалентности пунктов (1) и (2) леммы 0.39 для случая, когда $\mathfrak{A}_X = \mathcal{B}(\tau)$. \square

0.5 Предварительные сведения из анализа Фурье

0.6 Предварительные сведения из линейной алгебры

0.6.1 Билинейные функции и квадратичные формы

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

Отображение $B: V \times V \rightarrow \mathbb{K}$ 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) $\forall v, u \in V, \lambda \in \mathbb{K} \ B(\lambda u, v) = \lambda B(u, v);$
- (3) $\forall v, u, w \in V \ B(u, v + w) = B(u, w) + B(u, v);$
- (4) $\forall v, u \in V, \lambda \in \mathbb{K} \ B(u, \lambda v) = \lambda B(u, v).$

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

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

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

Квадратичной формой называется отображение $Q: V \rightarrow \mathbb{K}$ такое, что для некоторой билинейной формы и любого 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 $\mathbb{K} = \mathbb{R}$, 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 $Q(v) \geq 0$, 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 0.41 (Коши, Буняковский, Шварц). Пусть V — векторное пространство над полем \mathbb{R} и B — скалярное произведение на V . Тогда для любых двух векторов $u, v \in V$ выполнено равенство

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

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

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

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

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

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. \square

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

0.6.2 Полуторалинейные функции

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

Отображение $S: V \times V \rightarrow \mathbb{C}$ 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) $\forall v, u \in V, \lambda \in \mathbb{C} \ S(\lambda u, v) = \lambda S(u, v);$

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

$$(4) \quad \forall v, u \in V, \lambda \in \mathbb{K} \quad S(u, \lambda v) = \bar{\lambda} S(u, v), \text{ где надчёркивание означает комплексное сопряжение.}$$

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 0.42 (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

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

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

$$\begin{aligned} S(u + t\lambda v, u + t\lambda v) &= S(u, u) + t\bar{\lambda} S(u, v) + t\lambda S(v, u) + t^2 \lambda \bar{\lambda} S(v, v) = \\ &= S(u, u) + 2t S(u, v) S(v, u) + t^2 S(u, v) S(v, u) S(v, v) \leq 0 \end{aligned}$$

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

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

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. \square

0.7 Элементарная комбинаторика

0.7.1 Элементарные факты для подсчёта числа комбинаций

Центральной задачей комбинаторной математики можно считать задачу расположения объектов в соответствии со специальными правилами и нахождения числа способов, которыми это может быть сделано. Для нахождения числа способов полезны следующие результаты.

Лемма 0.43. *Мощность объединения двух конечных множеств A и B равна*

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Доказательство проводится изображением диаграммы Элера-Венна и подсчётом элементов в её частях. \square

Лемма 0.44 (формула включений-исключений).

$$\begin{aligned} \left| \bigcup_{i=1}^n A_i \right| = & \sum_i |A_i| - \sum_{i < j} |A_i \cap A_j| + \sum_{i < j < k} |A_i \cap A_j \cap A_k| - \dots \\ & \dots + (-1)^{n-1} |A_1 \cap A_2 \cap \dots \cap A_n|. \end{aligned}$$

Доказательство проводится по индукции с помощью предыдущей леммы, которая используется как база и шаг индукции одновременно. \square

Лемма 0.45. *Число пар, которые можно составить, когда первый элемент выбирается из множества A , а второй независимо от первого берётся из множества B равно $|A| \cdot |B|$.*

Доказательство. Следует из формулы $|A \times B| = |A| \cdot |B|$. \square

Лемма 0.46. *Число кортежей (a_1, \dots, a_n) в которых независимо выбираются первый элемент a_1 из множества A_1 , второй a_2 — из множества A_2 , \dots , a_n — из множества, равно $|A_1| \cdot |A_2| \cdot \dots \cdot |A_n|$.*

Доказательство. По индукции из предыдущей леммы. \square

0.7.2 Классические комбинаторные величины

Перестановка — это расположение n различных объектов по различным местам. Число перестановок множества в n элементов обозначаем P_n . Например, число способов разложить девять подписанных писем в девять подписанных конвертов — это число перестановок P_9 .

Перестановка с повторениями — это расположение n объектов, некоторые из которых могут быть тождественны друг другу, по различным местам. Число перестановок множества в n элементов, где n_1 элементов 1-го типа, \dots , n_k элементов k -того типа (причём $n_1 + \dots + n_k = n$) обозначаем $P(n_1, \dots, n_k)$. Например, число способов составить из букв слова «математика» всевозможные (абстрактные) слова — это число перестановок с повторениями $P(2, 3, 2, 1, 1, 1)$.

Упорядоченная выборка (кортеж) в t элементов из множества в n элементов называется **размещением**, неупорядоченная выборка (подмножество) из n элементов по t называется **сочетанием**. Выборка может быть без повторений, если элементы повторяться не могут, и может быть с повторениями, если элементы в

выборке повторяются. Число размещений без повторений обозначается A_n^m , с повторениями — \tilde{A}_n^m . Число сочетаний без повторений обозначается C_n^m или $\binom{n}{k}$, с повторениями — \tilde{C}_n^m или $\left(\binom{n}{k}\right)$.

Примеры:

- (1) Число способов расставить шесть книг из имеющихся восьми на книжной полке — это число размещений без повторения A_8^6 .
- (2) Число способов выбрать команду в пять человек из девяти данных игроков — это число сочетаний без повторения C_9^5 .
- (3) Число способов купить семь ручек из продающихся ручек пяти фирм — это число сочетаний с повторениями \tilde{C}_5^7 .

Теорема 0.47 (о комбинаторных числах). Числа перестановок, размещений и сочетаний находятся для неотрицательных целых n, n_i, t по следующим формулам (для комбинаций без повторений дополнительно требуется $t \leq n$):

Комбинация	без повторений	с повторениями
перестановка	$P_n = n!$	$P(n_1, \dots, n_k) = \frac{(n_1 + \dots + n_k)!}{n_1! \dots n_k!}$
размещение	$A_n^m = n^{[m]} = \frac{n!}{(n-m)!}$	$\tilde{A}_n^m = n^m$
сочетание	$C_n^m = \frac{n!}{m!(n-m)!}$	$\tilde{C}_n^m = C_{n+m-1}^m = \frac{(n+m-1)!}{m!(n-1)!}$

Доказательство.

Для A_n^k применяем принцип умножения: $n \times (n-1) \times \dots \times (n-k+1) = \frac{n!}{(n-k)!}$.

Для C_n^k замечаем, что выбрать k и упорядочить их можно A_n^k способами, причём каждый набор упорядочен $k!$ раз, поэтому $C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!}$.

Для \tilde{C}_n^m рассуждаем следующие образом. Необходимо выбрать из набора из n различных типов предметов ровно m предметов, причём каждый тип можно выбрать несколько раз (т.е. с повторениями), и порядок выбора не важен. Представим каждую из выбранных m «штук» как звёздочку (*), а разделители между типами — как палочки (|). Всего типов n , значит между этими типами нужно разместить $n-1$ разделитель. Расположение звёзд и палочек вдоль одной оси однозначно кодирует, сколько звёзд (предметов) взято каждого типа. Всего символов («звёзд» и «палочек») = $m + (n-1)$. Из них нужно выбрать m позиций для звёзд (или, эквивалентно, $n-1$ позиций для палочек). Таким образом число всех конфигураций равно числу способов выбрать m позиций из $(m+n-1)$, т.е. $\tilde{C}_n^m = C_{n+m-1}^m$. \square

0.7.3 Свойства комбинаторных величин

Теорема 0.48 (о свойствах числа сочетаний).

$$C_n^m = P(m, n-m) \quad (m \leq n);$$

$$C_n^m = C_n^{n-m} \quad (m \leq n);$$

$$C_n^m = \frac{n}{m} \cdot C_{n-1}^{m-1} \quad (m \leq n);$$

$$C_n^m + C_n^{m+1} = C_{n+1}^{m+1} \quad (m < n).$$

Теорема 0.49 (о бинOME Ньютона).

$$(x + y)^n = \sum_{k=0}^n C_n^k x^{n-k} y^k =$$

$$= C_n^0 x^n + C_n^1 x^{n-1} y + \dots + C_n^k x^{n-k} y^k + \dots + C_n^n y^n;$$

В связи с этой теоремой числа сочетаний $C_n^m = \binom{n}{m}$ называют биномиальными коэффициентами.

Бином Ньютона может быть обобщён до полинома Ньютона — возведения в степень суммы произвольного числа слагаемых:

Теорема 0.50 (о мультиномине Ньютона).

$$(x_1 + x_2 + \dots + x_k)^n = \sum_{\substack{n_j \geq 0 \\ n_1 + n_2 + \dots + n_k = n}} P(n_1, \dots, n_k) x_1^{n_1} \dots x_k^{n_k}.$$

В связи с формулой из теоремы число $P(n_1, \dots, n_k)$ также называется мультиномиальным коэффициентом и обозначается $\binom{n}{n_1, n_2, \dots, n_k}$.

При $k = 2$ мультиномиальный коэффициент совпадает с биномиальным $\binom{n_1+n_2}{n_1, n_2} = \binom{n_1+n_2}{n_1} = \binom{n_1+n_2}{n_2}$.

Урновая схема. Пусть имеется n урн, пронумерованных от 1 до n . Разместим в урны произвольным образом m шаров, где $m \leq n$. Число способов разместить шары зависит от двух условий: (1) различимы ли шары или нет; (2) имеет ли место принцип исключения, который не позволяет положить второй шар в урну, уже содержащую один шар. Это число способов в зависимости от комбинации условий находится как одно из известных комбинаторных чисел:

Комбинация	с принципом исключений	без принципа исключений
шары различимы	$A_n^m = n^{[m]} = \frac{n!}{(n-m)!}$	$\tilde{A}_n^m = n^m$
шары неразличимы	$C_n^m = \frac{n!}{m!(n-m)!}$	$\tilde{C}_n^m = C_{n+m-1}^m = \frac{(n+m-1)!}{m!(n-1)!}$

Число решений (x_1, \dots, x_n) уравнения

$$x_1 + \dots + x_n = k$$

в неотрицательных целых x_i равно числу \tilde{C}_n^k сочетаний с повторениями из n элементов по k .

Задача о беспорядках. Требуется найти число перестановок (p_1, p_2, \dots, p_n) множества $\{1, 2, \dots, n\}$ таких что $p_i \neq i$ для всех i . Такие перестановки называются **беспорядками**

Теорема 0.51 (о числе беспорядков). Число беспорядков на множестве из n элементов равно $n! \left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \dots + (-1)^n \frac{1}{n!}\right)$.

Доказательство следует из формулы включений-исключений. \square

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 $P(\Omega) = 1$. Тогда тройка $(\Omega, \mathfrak{F}, P)$ называется **вероятностным пространством**. Множество Ω называется **пространством элементарных событий (исходов)**, элементы σ -алгебры \mathfrak{F} называются **событиями**, а мера P **вероятностной мерой**.

Иногда мы будем называть вероятностное пространство «экспериментом» или «испытанием» или говорить, что «эксперименту» или «испытанию» соответствует вероятностное пространство, выражая таким образом физический смысл этого понятия: проходит эксперимент (испытание), у которого есть различные элементарные исходы. Эти исходы могут в результате этого эксперимента в разных комбинациях возникнуть с разной вероятностью.

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

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

Пример. Пусть $\Omega = [0, 1]$ и $\mathfrak{F} = M$ — σ -алгебра измеримых относительно меры Лебега подмножеств отрезка $[0, 1]$, $P = \mu$ — классическая мера Лебега. Тогда $P((\frac{1}{2}, \frac{3}{4})) = \frac{1}{4}$, $P(\frac{2}{9}) = 0$ и $P((0, \frac{1}{2}) \cup (\frac{2}{3}, 1)) = \frac{5}{6}$.

Пример. Пусть теперь $\Omega = [0, 2]$ и $\mathfrak{F} = M$ — σ -алгебра измеримых относительно меры Лебега подмножеств отрезка $[0, 2]$, $P = \frac{1}{2}\mu$ — мера, пропорциональная мере Лебега (мы выбрали именно такую меру, чтобы удовлетворить условию $P(\Omega) = 1$). Тогда $P((0, 1)) = \frac{1}{2}$, $P(\frac{4}{5}) = 0$ и $P((0, \frac{1}{2}) \cup (1, \frac{4}{3})) = \frac{5}{12}$.

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

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 $P: \Omega \rightarrow \mathbb{R}$ — неотрицательная функция, удовлетворяющая свойству $\sum_{k \in \mathbb{N}} P(\omega_k) = 1$. 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

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

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 $P: 2^\Omega \rightarrow \mathbb{R}$ — функция, сопоставляющая событию его вероятность. Тогда тройка $(\Omega, 2^\Omega, P)$ 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

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

Пусть $A, B \subset \Omega$ и $A \cap B = \emptyset$. 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

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

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

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

□

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 \bar{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 $A \cap B = \emptyset$.

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

События Ω и \emptyset называются **достоверным** и **невозможным**, соответственно.

Если $AB = \emptyset$, 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) $P(\bar{A}) = 1 - P(A)$;
- (2) если $A \subset B$, то $P(B \setminus A) = P(B) - P(A)$;
- (3) если $A \subset B$, то $P(A) \leq P(B)$;
- (4) $P(A \cup B) = P(A) + P(B) - P(AB)$;
- (5) $P(A \cup B) \leq P(A) + P(B)$;
- (6) $P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n \left(\sum_{i_1 < i_2 < \dots < i_k} (-1)^{k-1} P(A_{i_1} \dots A_{i_k}) \right)$;
- (7) $P\left(\bigcup_{k=1}^{+\infty} A_k\right) \leq \sum_{k=1}^{+\infty} P(A_k)$ (это свойство называется **субаддитивностью**).

Доказательство. Равенство (1) следует из цепочки

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

Равенство (2) — из цепочки

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

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

Равенство (4) — из цепочки

$$\begin{aligned} P(A \cup B) &= P((A \setminus B) \sqcup (A \cap B) \sqcup (B \setminus A)) = \\ &= P(A \setminus B) + P(A \cap B) + P(B \setminus A) + P(A \cap B) - P(A \cap B) = \\ &= P((A \setminus B) \sqcup (A \cap B)) + P((B \setminus A) \sqcup (A \cap B)) - P(A \cap B) = \\ &= P(A) + P(B) - P(A \cap B). \end{aligned}$$

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

Докажем (6) по индукции.

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

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

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

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

$$\begin{aligned} \sum_{k=1}^{n-1} \left(\sum_{i_1 < i_2 < \dots < i_k} (-1)^{k-1} P(A_{i_1} \dots A_{i_k}) \right) + P(A_n) - \sum_{k=1}^{n-1} \left(\sum_{i_1 < i_2 < \dots < i_k} (-1)^{k-1} P(A_{i_1} \dots A_{i_k} A_n) \right) = \\ = \sum_{k=1}^n \left(\sum_{i_1 < i_2 < \dots < i_k} (-1)^{k-1} P(A_{i_1} \dots A_{i_k}) \right). \end{aligned}$$

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

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

□

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 $P(A) > 0$. Тогда вероятность $P(B | A) = \frac{P(AB)}{P(A)}$ называется **условной вероятностью события B при условии того, что событие A произойдёт с вероятностью $P(A) > 0$** . Вероятность $P(B)$ также иногда называется **априорной вероятностью события B** .

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

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

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

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

Так как обе величины $P(AB)$ и $P(A)$ неотрицательны (а последняя и вовсе положительна), то $P(B | A) \geq 0$.

Справедливость упомянутого равенства выводится из определения условной вероятности:

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

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

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

□

Следствие 2.2. 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 $P(B_1 | A) = 1$;
- (2) $P(B_1 \cup B_2 | A) = P(B_1 | A) + P(B_2 | A) - P(B_1 B_2 | A)$;
- (3) если B_1 и B_2 несовместны, то $P(B_1 + B_2 | A) = P(B_1 | A) + P(B_2 | A)$.

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

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

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

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

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 2.4 (Формула произведения вероятностей). Пусть $A_1, \dots, A_n \in \mathfrak{F}$ — события. Если вероятности событий $A_2 A_3 \dots A_n, \dots, A_{n-1} A_n, A_n$ не равны нулю, то имеет место формула

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

Если вероятности событий $A_1 A_2 \dots A_n, A_1 A_2 \dots A_{n-1}, \dots, A_1$ не равны нулю, то имеет место формула

$$P(A_1 A_2 \dots A_n) = P(A_n | A_1 \dots A_{n-1}) P(A_{n-1} | A_1 \dots A_{n-2}) \dots P(A_2 | A_1) P(A_1).$$

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 $B = A_2 A_3 \dots A_n$. Тогда по базе индукции (здесь мы пользуемся тем, что $P(B) > 0$) и затем по индукционному предположению (а здесь всеми остальными условиями) имеем

$$P(A_1 B) = P(A_1 | B) P(B) = P(A_1 | A_2 A_3 \dots A_n) P(A_2 | A_3 \dots A_n) \dots P(A_{n-1} | A_n) P(A_n).$$

□

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

Теорема 2.5 (Формула полной вероятности). Пусть $A_1, \dots, A_n \in \mathfrak{F}$ — разбиение Ω . Тогда для всякого события B имеет место равенство формула полной вероятности

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

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

$$\begin{aligned} P(B) &= P(B\Omega) = P(B(A_1 + \dots + A_n)) = P(BA_1 + \dots + BA_n) \stackrel{\text{несовместность}}{=} \\ &\stackrel{\text{несовместность}}{=} \sum_{k=1}^n P(BA_k) = \sum_{k=1}^n P(B | A_k) P(A_k). \end{aligned}$$

□

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

Теорема 2.6 (Формула Байеса). Пусть события $A_1, \dots, A_n \in \mathfrak{F}$ образуют разбиение Ω , пусть $B \in \mathfrak{F}$ — ещё одно событие и $P(B) > 0$. Тогда справедлива формула Байеса

$$P(A_i | B) = \frac{P(B|A_i) P(A_i)}{\sum_{k=1}^n P(B|A_k) P(A_k)}.$$

Доказательство. По лемме 2.3 («простейшая формула Байеса») имеем равенство

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

По формуле полной вероятности имеем

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

откуда следует искомая формула. \square

Пример. Приведём стандартный пример на применение простейшего вида формулы Байеса. Пусть в популяции заболевание встречается с вероятностью $P(B) = 0.1$. ПЦР-тест на выявление заболевания устроен так, что:

- При наличии заболевания он даёт положительный результат с вероятностью $P(+ | B) = 0.9$,
- При отсутствии заболевания он даёт ложноположительный результат с вероятностью $P(+ | \bar{B}) = 0.2$.

Найдём вероятность того, что человек действительно болен, если результат теста оказался положительным. По формуле Байеса:

$$P(B | +) = \frac{P(+ | B) \cdot P(B)}{P(+ | B) \cdot P(B) + P(+ | \bar{B}) \cdot P(\bar{B})}.$$

Подставим известные значения:

$$P(B | +) = \frac{0.99 \cdot 0.01}{0.99 \cdot 0.01 + 0.05 \cdot 0.99} = \frac{0.0099}{0.0099 + 0.0495} = \frac{0.0099}{0.0594} \approx 0.1667.$$

Формально мы рассматривали дискретное вероятностное пространство $\Omega = \{+B, -B, +\bar{B}, -\bar{B}\}$ с четырьмя элементарными исходами, выражающими все возможные комбинации результата тестирования и реального состояния тестируемого человека. Здесь событие B «болен» являлось объединением $\{+B, -B\}$, событие «пцр-тест дал положительный результат» — объединением $\{+B, +\bar{B}\}$ и так далее.

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 $P(B | A) = P(B)$ и $P(A | B) = P(A)$. Чтобы не ограничиваться случаями, когда вероятности событий больше 0, мы определим независимость следствием формул выше. События A и B называются **независимыми**, если справедливо равенство $P(AB) = P(A)P(B)$.

Предложение 2.7 (Свойства независимости). *Имеют место следующие утверждения:*

- (1) если $P(B) > 0$, то независимость A и B равносильна равенству $P(A | B) = P(A)$;
- (2) если A и B независимы, то \bar{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 $P(B) > 0$, то по независимости имеем

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

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

Для доказательства (2) выпишем цепочку равенств

$$\begin{aligned} P(\bar{A}B) &= P((\Omega \setminus A)B) = P(B \setminus AB) = P(B) - P(AB) \stackrel{\text{независимость}}{=} \\ &\stackrel{\text{независимость}}{=} P(B) - P(A)P(B) = P(B)(1 - P(A)) = P(B)P(\bar{A}). \end{aligned}$$

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

$$\begin{aligned} P((B_1 + B_2)A) &= P(B_1A + B_2A) = P(B_1A) + P(B_2A) = P(B_1)P(A) + P(B_2)P(A) = \\ &= (P(B_1) + P(B_2))P(A) = P(B_1 + B_2)P(A). \end{aligned}$$

□

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

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

Предложение 2.8. Если события B_1, \dots, 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 записывается так: $\Omega = \{\omega_R, \omega_G, \omega_B, \omega_{RGB}\}$ — элементарное событие — выпала грань с данной раскраской. По условию $P(\omega_R) = P(\omega_G) = P(\omega_B) = P(\omega_{RGB}) = \frac{1}{4}$. 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 $P(R) = P(G) = P(B) = \frac{1}{2}$, $P(RG) = P(GB) = P(BR) = \frac{1}{4} = P(R)P(G) = P(G)P(B) = P(B)P(R)$, но $P(RGB) = P(\omega_{RGB}) = \frac{1}{4} \neq P(R)P(G)P(B) = \frac{1}{8}$.

Пример. Пользуясь примером выше, можно показать, что условие несовместности в пункте (3) предложения 2.7 нельзя опустить. Положим $A = R$ и $B_1 = G$ и $B_2 = B$. Тогда $P((B_1 \cup B_2)A) = P(\omega_{RGB}) = \frac{1}{4}$, но $P(A) = \frac{1}{2}$, $P(B_1 + B_2) = \frac{3}{4}$ 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 $(\Omega, \mathfrak{F}, P)$ и событий $A_1, \dots, A_n \in \mathfrak{F}$ для которых множество наборов, на которых выполнены равенства

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

в точности совпадает с J .

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

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

где для $k = 0$ предполагается, что в матрице стоит $P(\Omega)$. Поскольку $P(A_{i_1} \dots A_{i_k}) = P(\{(\varepsilon_1, \dots, \varepsilon_n) \mid \varepsilon_i \in \{0, 1\}, \varepsilon_{i_s} = 1, 1 \leq s \leq k\}) = \sum_{\varepsilon_{i_s}=1, 1 \leq s \leq k} p_{(\varepsilon_1, \dots, \varepsilon_n)}$, то φ — линейное

отображение. Можно показать, что φ сюръективно (проверьте с помощью элементарных преобразований, что его матрица имеет ранг 2^n) и, следовательно, биективно. Достаточно подобрать значения вероятностей все возможных произведений A_i так, чтобы вероятности $p_{(\varepsilon_1, \dots, \varepsilon_n)}$ были неотрицательны, в сумме давали 1 ($P(\Omega) = 1$) и при этом выполнялись в точности все желаемые равенства на вероятности произведений событий A_i . Положим $P(A_i) = \frac{1}{2^{2^n}}$, $P(\Omega) = 1$. Если $(i_1, \dots, i_k) \in S_J$, то положим $P(A_{i_1} \dots A_{i_k}) = \frac{1}{2^{2^{kn}}}$. Иначе положим $P(A_{i_1} \dots A_{i_k}) = \frac{1}{2^{2^{kn+1}}}$. Проверим, что имеют место неравенство $\frac{1}{2^{2^{kn+2}}} \leq p_{(\varepsilon_1, \dots, \varepsilon_n)} \leq \frac{1}{2^{2^{kn}}}$, 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 $(\varepsilon_1, \dots, \varepsilon_n)$ при-

существует $n - k \geq 1$ нулей. Прибавим к $p_{(\varepsilon_1, \dots, \varepsilon_n)}$ все остальные значения вероятностей элементарных исходов — кортежей, в которых некоторые нули из данного кортежа заменены на единицы. Тогда $p_{(\varepsilon_1, \dots, \varepsilon_n)} \leq \frac{1}{2^{2kn}}$, так как по предположению индукции все остальные слагаемые положительны, а сумма не превосходит $\frac{1}{2^{2kn}}$. С другой стороны, $p_{(\varepsilon_1, \dots, \varepsilon_n)} \geq \frac{1}{2^{2kn+1}} - \frac{2^k - 1}{2^{2(k+1)n}} = \frac{1}{2^{2kn+1}} - \frac{1}{2^{2kn+2+(2n-k-2)}}$. Так как $n - k - 1 \leq 0$ и $n - 1 \leq 0$, то последнее слагаемое по модулю не превосходит $\frac{1}{2^{2kn+2}}$. Тогда $p_{(\varepsilon_1, \dots, \varepsilon_n)}$ не меньше $\frac{1}{2^{2kn+2}}$. Остаётся убедиться в том, что $p_{(0, \dots, 0)} = 1 - \sum_{(\varepsilon_1, \dots, \varepsilon_n) \neq (0, \dots, 0)} p_{(\varepsilon_1, \dots, \varepsilon_n)} \geq 1 - \frac{2^n - 1}{2^{2n}} > 0$.

Предложение 2.9. Пусть событие A независимо с самим собой тогда и только тогда, когда $P(A) = 0$ или $P(A) = 1$.

Доказательство. По определению независимости $P(AA) = P(A)P(A)$, откуда $P(A)^2 - P(A) = 0$ и либо $P(A) = 0$, либо $P(A) = 1$. Обратно, если одно из этих равенств выполнено, то $P(A)P(A) = P(A) = P(AA)$. \square

2.4 Произведение вероятностных пространств

Если считать, что каждый исход получен в результате отдельного испытания, то мы обнаружим, что любое событие, относящееся к фиксированному испытанию будет независимым от любого события, относящегося к другим испытаниям. В таких случаях говорят о последовательности независимых испытаний.

Формализуем данную ситуацию. Рассмотрим два вероятностных пространства $(\Omega_1, \mathfrak{F}_1, P_1)$ и $(\Omega_2, \mathfrak{F}_2, P_2)$. Их **произведением** мы будем называть вероятностное пространство $(\Omega, \mathfrak{F}, P)$, где $\Omega = \Omega_1 \times \Omega_2$ — декартово произведение, $\mathfrak{F} = \mathfrak{F}_1 \otimes \mathfrak{F}_2$ — σ -алгебра, порождённая всеми возможными произведениями $A_1 \times A_2$, где $A_1 \in \mathfrak{F}_1, A_2 \in \mathfrak{F}_2$ — события в исходных вероятностных пространствах (заметьте, что сами по себе таким произведения образуют полукольцо, но вообще говоря не образуют даже кольца), а мера P определяется как продолжение меры m с полукольца произведений, заданной по правилу $m(A_1 \times A_2) := P_1(A_1) \cdot P_2(A_2)$.

Пусть теперь $(\Omega, \mathfrak{F}, \tilde{P})$ — вероятностное пространство с теми же Ω и \mathfrak{F} , то возможно другой вероятностной мерой. Это можно представлять себе так: происходят два испытания и результаты одного могут «повлиять» на результаты другого. Неформальное «повлиять» выражается понятием независимости испытаний. Мы будем говорить, что испытания, которым соответствуют вероятностные пространства $(\Omega_1, \mathfrak{F}_1, P_1)$ и $(\Omega_2, \mathfrak{F}_2, P_2)$, а составному эксперименту, состоящему из двух этих испытаний, **независимы**, если для любых двух событий $A_1 \in \mathfrak{F}_1$ и $A_2 \in \mathfrak{F}_2$ выполнено равенство

$$\tilde{P}(A_1 \times A_2) = P_1(A_1)P_2(A_2) = \tilde{P}(A_1 \times \Omega_2)\tilde{P}(\Omega_2 \times A_2).$$

3 Случайные величины, их распределения, функции распределения и плотности

Пусть $(\Omega, \mathfrak{F}, P)$ — вероятностное пространство. Функция $\xi: \Omega \rightarrow \mathbb{R}$ называется **случайной величиной**, если прообраз любого борелевского множества $B \in \mathcal{B}$ при отображении f лежит в \mathfrak{F} . Соотнося это из определениями из действительного анализа, мы можем сказать, что ξ является **измеримой функцией**. Также, вспомним, что мы получим эквивалентное определение, если вместо всех борелевских множеств будем рассматривать все интервалы, все отрезки, все интервалы одного из двух видов $(-\infty, b)$ и $(a, +\infty)$ или все полуинтервалы одного из двух видов $(-\infty, b]$ и $[a, +\infty)$.

Случайные величины удовлетворяют всем тем стандартным свойствам, которыми удовлетворяют измеримые функции:

Предложение 3.1. Пусть ξ — случайная величина и g — непрерывная на $\text{Im } \xi$ функция. Тогда композиция $g(\xi)$ является случайной величиной.

Доказательство. Переформулировка предложения 0.22. □

Предложение 3.2. Пусть ξ, η — случайные величины. Тогда множество $A_{\xi \leq \eta} = \{\omega \in \Omega \mid \xi(\omega) \leq \eta(\omega)\}$ измеримо. Функции $a + \xi, a\xi, |\xi|, \xi^2, \xi + \eta$ и $\xi\eta$, где a — константа, являются случайными величинами. Если случайная величина η не принимает значения 0, то функции $\frac{1}{\eta}$ и $\frac{\xi}{\eta}$ также являются случайными величинами.

Доказательство. Переформулировка предложения 0.23. □

Случайная величина называется **дискретной**, если она принимает не более, чем счётное число различных значений. Случайные величины на дискретном вероятностном пространстве всегда дискретны. Кроме того, всякая функция на дискретном вероятностном пространстве (в смысле **второго определения**) является случайной величиной.

Пример. Пусть $\Omega = \{ГГ, ГР, РГ, РР\}$ — все возможные результаты выпадения на двух монетках с гербом (Г) на одной стороне и решкой (Р) на другой. Будем считать, что вероятности всех элементарных событий равны $\frac{1}{4}$. Пусть $\xi: \Omega \rightarrow \{0, 1\}$ сопоставляет элементам ГГ, ГР и РГ единицу, а РР — ноль. Тогда ξ — дискретная случайная величина. Мы можем интерпретировать 0 как «неудачу» в эксперименте «бросить две монетки и получить хотя бы один герб» и 1 — «успех». В нашем случае ξ принимает значение 1 с вероятностью $P(\xi^{-1}(\{1\})) = P(\{ГГ, ГР, РГ\}) = \frac{3}{4}$ и значение 0 с вероятностью $P(\xi^{-1}(\{0\})) = P(\{РР\}) = \frac{1}{4}$.

Вместо $P(\xi^{-1}(\{a\}))$ мы будем пользоваться записью $P(\xi = a)$, отражающей смысл этого выражения — «вероятность случайной величины ξ принять значение a ». Также вместо перегруженных скобками выражений $P(\xi^{-1}((-\infty; b)))$, $P(\xi^{-1}((a; +\infty)))$, $P(\xi^{-1}((a; b)))$ и им подобных мы будем писать $P(\xi \leq b)$, $P(\xi \geq a)$ и $P(a \leq \xi \leq b)$.

Пример. В примере выше $P(\xi = 1) = \frac{3}{4}$ и $P(\xi = 0) = \frac{1}{4}$.

Пример. Дискретная случайная величина может быть задана и не на дискретном вероятностном пространстве. Пусть $\Omega = [-1; 1]$ — отрезок, $\mathfrak{F} = \mathcal{B}$ — сигма алгебра его борелевских подмножеств, а вероятностная мера — это умноженная на $\frac{1}{2}$ классическая мера Лебега μ . Тогда случайная величина $\xi(\omega) = \text{sgn}(\omega)$, заданная функцией знак, является дискретной случайной величиной.

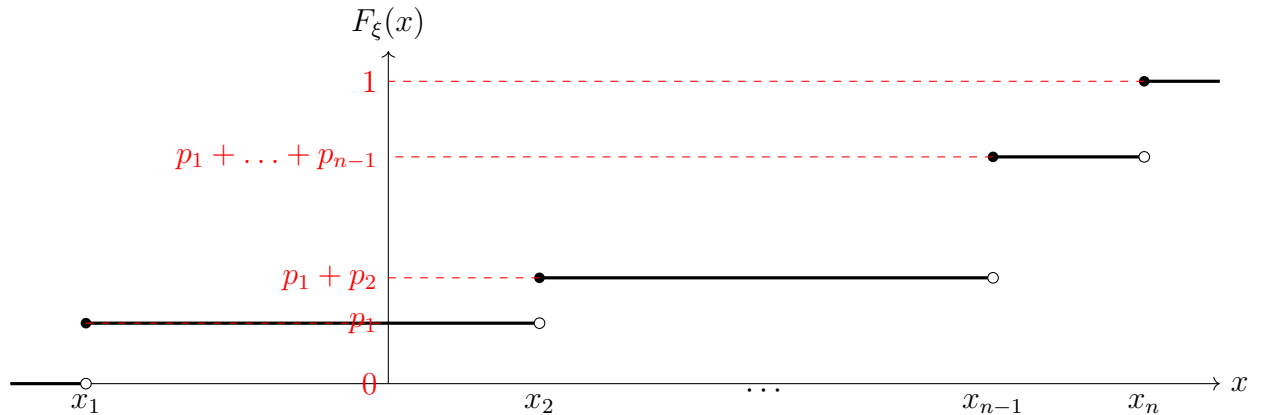
Пример. Вернёмся одному из первых рассмотренных примеров. Пусть $\Omega = [0, 2]$ и $\mathfrak{F} = M$ — σ -алгебра измеримых относительно меры Лебега подмножеств отрезка $[0, 2]$, $P = \frac{1}{2}\mu$ — мера, пропорциональная мере Лебега. Пусть $\xi(\omega) = \omega^2$. Эта функция непрерывна, поэтому является случайной величиной (измеримой функцией). Найдём $P(\frac{1}{4} \leq \xi < \frac{16}{9})$. Так как неравенства $\frac{1}{2} \leq \xi(\omega) < \frac{16}{9}$ выполнены тогда и только тогда, когда $\frac{1}{2} \leq \omega < \frac{4}{3}$. Тогда $P(\frac{1}{4} \leq \xi < \frac{16}{9}) = \frac{1}{2}(\frac{4}{3} - \frac{1}{2}) = \frac{5}{12}$.

Введём некоторые объекты, характеризующие случайную величину ξ через её вероятность принять некоторые значения. **Функцией распределения случайной величины ξ** будем называть функцию $F_\xi: \mathbb{R} \rightarrow \mathbb{R}$, заданную по правилу $F_\xi(x) = P(\xi \leq x)$. Заметим, что построенная функция F_ξ корректно определена на всей числовой прямой, так как ξ — случайная величина и прообраз бесконечного полуинтервала относительно ξ является событием (измерим).

Пример. Функция распределения дискретной случайной величины ξ имеет «ступенчатую форму». Пусть ξ принимает конечное число значений $x_1 < x_2 < \dots < x_n$ и $P(\xi = x_i) = p_i$. Тогда функция распределения F_ξ может быть задана как

$$F_\xi(x) = \begin{cases} 0, & x < x_1; \\ p_1, & x_1 \leq x < x_2; \\ p_1 + p_2, & x_2 \leq x < x_3; \\ \dots & \\ p_1 + p_2 + \dots + p_n = 1, & x \geq x_n. \end{cases}$$

График F_ξ будет выглядеть следующим образом.



Выделим некоторые свойства этой функции.

Предложение 3.3. Пусть ξ — случайная величина и F_ξ — её функция распределения. Тогда выполнены следующие свойства

- (1) функция F_ξ неубывает;
- (2) функция F_ξ непрерывна справа;
- (3) $F_\xi(+\infty) := \lim_{x \rightarrow +\infty} F_\xi(x) = 1$;
- (4) $F_\xi(-\infty) := \lim_{x \rightarrow -\infty} F_\xi(x) = 0$.

Доказательство. Для доказательства неубывания (1) воспользуемся свойствами вероятностной меры. Пусть $x < y$. Тогда

$$F_\xi(y) = P(\xi \leq y) = P(\xi \leq x) + P(x < \xi \leq y) \geq P(\xi \leq x) = F_\xi(x).$$

Проверим непрерывность справа (2). Пусть последовательность $\{x_n\}$ сходится к x_0 справа (можно считать, что последовательность убывает). Тогда $\bigcap_{n=1}^{+\infty} (-\infty; x_n] = (-\infty; x_0]$ и по непрерывности вероятностной меры

$$\lim_{x \rightarrow x_0+0} F_\xi(x) = \lim_{n \rightarrow +\infty} F_\xi(x_n) = \lim_{n \rightarrow +\infty} P(\xi \leq x_n) = P(\xi \leq x_0) = F_\xi(x_0).$$

Свойства (3) и (4) также следуют из непрерывности

$$\lim_{x \rightarrow +\infty} F_\xi(x) = \lim_{n \rightarrow +\infty} F_\xi(x_n) = \lim_{n \rightarrow +\infty} P(\xi \leq x_n) = P(\xi \in \mathbb{R}) = 1;$$

$$\lim_{x \rightarrow -\infty} F_\xi(x) = \lim_{n \rightarrow +\infty} F_\xi(x_n) = \lim_{n \rightarrow +\infty} P(\xi \leq x_n) = P(\xi \in \emptyset) = 0.$$

Здесь в первом случае $\bigcup_{n=1}^{+\infty} (-\infty; x_n] = (-\infty; +\infty) = \mathbb{R}$, а во втором $\bigcap_{n=1}^{+\infty} (-\infty; x_n] = (-\infty; -\infty] = \emptyset$. \square

Будем обозначать через $F_\xi(x-)$ или $F_\xi(x-0)$ левосторонний предел функции F_ξ в точке x . Тогда $P(\xi = x) = F_\xi(x) - F_\xi(x-)$.

Случайная величина ξ называется **непрерывной**, если её функция распределения F_ξ непрерывна на \mathbb{R} .

Предложение 3.4. Случайная величина ξ непрерывна тогда и только тогда, когда для любого числа $a \in \mathbb{R}$ вероятность того, что ξ примет значение a равна нулю: $P(\xi = a) = 0$.

Доказательство. Функция F_ξ монотонна и поэтому у неё существуют левые пределы во всех точках. Так как F_ξ ещё и непрерывна справа, то непрерывность в точке a для неё равносильна равенству $F_\xi(a) = F_\xi(a-)$, что в свою очередь равносильно равенству $P(\xi = a) = F_\xi(a) - F_\xi(a-) = 0$. \square

Теорема 3.5. Пусть функция F удовлетворяет свойствам из предложения 3.3. Тогда существует единственная вероятностная мера P_F на борелевской σ -алгебре \mathcal{B} подмножеств \mathbb{R} такая, что функция F является функцией распределения случайной величины $\text{id}: (\mathbb{R}, \mathcal{B}, P_F) \rightarrow (\mathbb{R}, \mathcal{B})$.

Доказательство. Существование следует из конструкции меры Лебега-Стилтьеса μ_F для функции F . Для неё функцией распределения тождественного отображения автоматически становится функция F (так как буквально теми же формулами определяется эта мера по полуинтервалам).

Пусть теперь P — некоторая вероятностная мера и функция F оказывается функцией распределения для тождественного отображения $\text{id}: x \mapsto x$ (далее мы будем также обозначать его через x). Тогда имеем

$$P((a, b]) = P(a < x \leq b) = P(x \leq b) - P(x \leq a) = F(b) - F(a);$$

$$P((-\infty, b]) = P(x \leq b) - 0 = F(b) - F(-\infty);$$

$$P((a, +\infty)) = P(x > a) = 1 - P(x \leq a) = 1 - F(a) = F(+\infty) - F(a).$$

Мы показали, что меры P и μ_F совпадают на полуинтервалах, следовательно, они совпадают на минимальной алгебре \mathcal{A} , порождённой ими. Так как борелевская σ -алгебра \mathcal{B} является минимальной σ -алгеброй, содержащей \mathcal{A} , то по теореме Каратеодори 0.17 меры P и μ_F совпадают на ней. \square

Для вероятностной меры P на измеримом пространстве $(\mathbb{R}, \mathcal{B})$ функцию распределения случайной величины тождественного отображения id мы будем иногда называть **функцией распределения вероятностной меры P** .

Следующая теорема не формулируется в основном курсе, но даёт ещё одно описание для меры P_F .

Прямой образ меры $\xi_* P$ (напомним, что он определяется правилом $\xi_* P(B) = P(\xi^{-1}(B)) = P(\xi \in B)$) на борелевской σ -алгебре \mathcal{B} подмножеств \mathbb{R} называется **распределением (вероятностей) случайной величины ξ** .

Теорема 3.6. Пусть ξ — случайная величина и F_ξ её функция распределения. Пусть μ_ξ — мера Лебега-Стилтьеса на подмножествах \mathbb{R} , порождённая функцией F_ξ . На борелевской σ -алгебре \mathcal{B} мера μ_ξ совпадает с распределением $\xi_* P$ вероятностной меры P .

Доказательство. Достаточно проверить, что эти меры совпадают на минимальном кольце, порождённом всеми полуинтервалами вида $(a, b]$, $(-\infty, b]$ и $(a, +\infty)$. Тогда равенство на борелевской σ -алгебре будет следовать из теоремы Каратеодори 0.17.

Поскольку мера на минимальном кольце, порождённом данным полукольцом, однозначной определяется по мере на этом полукольце, то достаточно проверить равенства на самих полуинтервалах.

Все равенства немедленно следуют из определений функции распределения, меры Лебега-Стилтьеса и прямого образа меры, а также свойств меры и функции распределения:

$$\begin{aligned}\mu_f((a, b]) &= F_\xi(b) - F_\xi(a) = P(\xi \leq b) - P(\xi \leq a) = P(a < \xi \leq b) = P(\xi^{-1}((a, b])) = \xi_* P((a, b]); \\ \mu_f((a, +\infty)) &= F_\xi(+\infty) - F_\xi(a) = 1 - P(\xi \leq a) = P(\xi > a) = P(\xi^{-1}((a, +\infty))) = \xi_* P((a, +\infty)); \\ \mu_f((-\infty, b]) &= F_\xi(b) - F_\xi(-\infty) = P(\xi \leq b) - 0 = P(\xi \leq b) = P(\xi^{-1}((-\infty, b])) = \xi_* P((-\infty, b]).\end{aligned}$$

□

В условиях последних двух теорем $\mu_F = \xi_* P_F$. Если случайная величина ξ имеет распределение Γ мы будем писать $\xi \sim \Gamma$.

Распределение дискретной случайной величины ξ называется **дискретным**. Если случайная величина ξ может принимать только значения x_1, x_2, \dots с вероятностями $p_i = P(\xi = x_i)$ (при этом мы будем считать, что вероятности $P(\xi = x_i)$ положительны, чтобы не рассматривать вырожденные случаи), то можно удобно изобразить это с помощью таблицы

ξ	x_1	x_2	\dots	или $\xi \sim$	x_1	x_2	\dots
	p_1	p_2	\dots		p_1	p_2	\dots

Поскольку $\sum_{i=1}^{+\infty} p_i = \sum_{i=1}^{+\infty} P(\xi = x_i) = P(\Omega) = 1$, то дискретное распределение $\{p_i\}$ можно задать на дискретном вероятностном пространстве.

Распределение случайной величины ξ и она сама называются **абсолютно непрерывными**, если существует неотрицательная функция $p_\xi: \mathbb{R} \rightarrow \mathbb{R}$ такая, что для любого борелевского множества $B \in \mathcal{B}$ выполнено равенство

$$P(\xi \in B) = \xi_* P(B) = \int_B p_\xi d\mu,$$

где μ — классическая мера Лебега. С точки зрения теории меры и интеграла Лебега это определение эквивалентно тому, что мера $\xi_* P$ абсолютно непрерывна относительно классической меры Лебега μ на прямой (это значит, что случайная величина ξ попадает в множество меры нуль с вероятностью нуль). Теорема Радона-Никодима даёт формулу для $\xi_* P$, принятую нами за определение. Функция p_ξ называется **функцией плотности** абсолютно непрерывной случайной величины ξ .

В силу доказанных выше теорем мы можем дать эквивалентное определение в терминах функции распределения. Будем говорить, что распределение случайной величины ξ и она сама называются **абсолютно непрерывными**, если функция распределения F_ξ может быть выражена как интеграл

$$F_\xi(x) = \int_{(-\infty; x]} p_\xi d\mu.$$

В действительном анализе такая функция F_ξ называется абсолютно непрерывной, почти всюду существует производная F'_ξ , которая почти всюду совпадает с $p(x)$, а также имеет место аналог формулы Ньютона-Лейбница: $F_\xi(b) - F_\xi(a) = \int_a^b p d\mu$. Можно доказать, что функция абсолютно непрерывна тогда и только тогда, когда она непрерывна, имеет ограниченную вариацию и переводим множества меры нуль по Лебегу в множества меры нуль.

Во многих рассматриваемых случаях (в частности, во всех примерах следующего раздела) функция p_ξ оказывается интегрируемой по Риману на любом отрезке и абсолютно интегрируема на \mathbb{R} , поэтому можно записать равенства

$$F_\xi(x) = \int_{-\infty}^x p_\xi(t) dt, \quad \int_{-\infty}^{+\infty} p_\xi(t) dt = 1.$$

4 Классические примеры распределений

В этом разделе мы рассмотрим классические примеры распределений, повсеместно встречающихся в теории вероятностей и в её приложениях.

Важно помнить, что мы описываем распределение случайной величины, а не непосредственно её саму (случайные величины с одинаковыми распределениями могут быть заданы на разных вероятностных пространствах). Вместо данной случайной величины всегда можно рассмотреть случайную величину id на вероятностном пространстве $(\mathbb{R}, \mathcal{B}, P)$ для некоторого P (или на пространстве $(U, \mathcal{B} \cap U, P)$ для некоторого подмножества $U \subset \mathbb{R}$), то есть распределение некоторой вероятностной меры на \mathbb{R} или подмножестве \mathbb{R} . Случайную величину, обладающую распределением Γ также называют Γ -ой (бернуллиевской, геометрической, биномиальной и т.д.) случайной величиной.

НУЖНО: вписать описания для всех классических распределений

4.1 Дискретные распределения.

Приведём примеры дискретных распределений. Каждый пример мы будем снабжать описывающей его таблицей.

4.1.1 Распределение константы

Пусть случайная величина ξ принимает значение C с вероятностью 1: $P(\xi = C) = 1$. Тогда распределение такой случайной величины называется **распределением константы** или «вырожденным распределением». Используя табличку можем записать

$$\xi \sim \begin{matrix} C \\ 1 \end{matrix}.$$

4.1.2 Распределение Бернулли

Пусть случайная величина ξ принимает значения 1 и 0 с вероятностями, соответственно p и q ($p + q = 1$, $p, q \geq 0$). Её распределение называется **распределением Бернулли** и обозначается $\text{Be}(p)$. И снова в виде таблицы

$$\xi \sim \begin{array}{cc} 0 & 1 \\ q & p \end{array} \sim \text{Be}(p).$$

Пример. Рассмотрим дискретное вероятностное пространство $\Omega = \{1, 2, 3, 4, 5, 6\}$ результатов броска игральной кости. Будем считать, что наша игральная кость «идеальная», то есть вероятности всех элементарных событий равны $\frac{1}{6}$. Пусть $\xi: \Omega \rightarrow \mathbb{R}$ — случайная величина, принимающая значение 1 на числах множества $\{1, 3\}$ и 0 иначе и выражающая смысл «выпала степень тройки». Тогда ξ обладает биномиальным распределением и

$$\xi \sim \begin{array}{cc} 0 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{array}.$$

4.1.3 Дискретное равномерное распределение

Рассмотрим случайную величину ξ принимающую значения из набора x_1, x_2, \dots, x_n , каждое с вероятностью $\frac{1}{n}$. Распределение случайной величины ξ называется **дискретным распределением** и обозначается $R\{1, \dots, n\}$. В виде таблицы:

$$\xi \sim \begin{array}{ccc} x_1 & \cdots & x_n \\ \frac{1}{n} & \cdots & \frac{1}{n} \end{array} \sim R\{1, \dots, n\}.$$

Пример. В качестве примера случайной величины с равномерным распределением можно взять случайную величину ξ на дискретном пространстве $\Omega = \{1, 2, 3, 4, 5, 6\}$ (вероятности всех элементарных событий снова равны $\frac{1}{6}$), сопоставляющую числу из Ω его же само.

4.1.4 Биномиальное распределение

Представим, что происходит n последовательных **независимых испытаний**, в каждом из которых с вероятностью $0 < p < 1$ происходит «успех», а с вероятностью $q = 1 - p$ «неудача». Формально, имеется пространство элементарных исходов $\Omega = \{0, 1\}^n$ с заданным на нём распределением вероятностей, полученное как **произведение** вероятностных пространств $(\{0, 1\}, 2^{\{0, 1\}}, P)$, где $P(1) = p$, $P(0) = q$. Из построения получаем вероятность элементарного исхода $(\varepsilon_1, \dots, \varepsilon_n)$ равной $p^k q^{n-k}$, где k — это «число успехов», то есть число единиц в кортеже $(\varepsilon_1, \dots, \varepsilon_n)$.

Рассмотрим случайную величину ξ принимающую на Ω значения, равные «числу успехов» (сумме единиц в кортеже). Тогда ξ принимает значения от 0 до n ,

причём значение k принимается ей с вероятностью $C_n^k p^k q^{n-k}$. Распределение такой случайной величины ξ называется **биномиальным распределением** и обозначается $B(n, p)$. В виде таблицы:

$$\xi \sim \begin{array}{cccccc} 0 & 1 & \dots & k & \dots & n \\ q^n & npq^{n-1} & \dots & C_n^k p^k q^{n-k} & \dots & p^n \end{array}.$$

4.1.5 Распределение Пуассона

Теперь мы введём счётный аналог биномиального распределение. Воспользуемся разложением экспоненты в ряд

$$e^\lambda = \sum_{k=1}^{+\infty} \frac{\lambda^k}{k!}.$$

Пусть случайная величина ξ принимает целые неотрицательные значения и принимает значение $k \in \mathbb{Z}_{\geq 0}$ с вероятностью $\frac{\lambda^k}{k!} e^{-\lambda}$, $\lambda > 0$ (это условие нужно для того, чтобы все вероятности были положительными). Распределение такой случайной величины называется **распределение Пуассона** и обозначается $\Pi(\lambda)$. Выписывая это в виде таблицы получаем

$$\xi \sim \begin{array}{cccccc} 0 & 1 & 2 & \dots & k & \dots \\ e^{-\lambda} & \lambda e^{-\lambda} & \frac{\lambda^2}{2} e^{-\lambda} & \dots & \frac{\lambda^k}{k!} e^{-\lambda} & \dots \end{array} \sim \Pi(\lambda).$$

4.1.6 Геометрическое распределение

Рассмотрим эксперимент с двумя исходами «успех» и «неудача», который повторяется до достижения «успеха». Будем считать, что в каждом отдельном испытании «успех» достигается с вероятностью $0 < p < 1$, а «неудача» с вероятностью $q = 1 - p$. Рассмотрим вероятностное пространство $\Omega = \mathbb{N}$ с распределением $P(n) = pq^{n-1}$ — вероятность того эксперимент закончится на n -м этапе (произошли $n - 1$ «неудача» и затем «успех»). Пусть случайная величина ξ сопоставляет числу из Ω его же само. Таким образом, ξ несёт в себя смысл «количество испытаний, проведённых до первого успеха». Распределение такой случайной величины называется **геометрическим распределением** и обозначается $G(p)$. Изобразим таблицу:

$$\xi \sim \begin{array}{cccccc} 1 & 2 & 3 & \dots & k & \dots \\ p & pq & pq^2 & \dots & pq^{k-1} & \dots \end{array} \sim G(p).$$

Пример. Пусть $\Omega = \{0, 1\}^{\mathbb{N}}$ — множество всех последовательностей из нулей и единиц. Пусть \mathfrak{F} — σ -алгебра, полученная лебеговским продолжением с полукольца последовательностей, имеющих фиксированное общее начало (в нём содержатся, например, последовательности, начинающиеся с 1, с 1101 и т.д., также будем включать в полукольцо всё Ω), по мере m заданной по правилу $m(S) = p^k q^m$ для множества S последовательностей с общим началом $\varepsilon_1 \varepsilon_2 \dots \varepsilon_{k+m}$. Здесь k равно

числу единиц в начальном префиксе, а m — числу нулей. Мы опустим проверку того, что это последовательности с общим фиксированным началом образуют полукольцо, а заданная на них функция m является σ -аддитивной мерой. Теперь пусть случайная величина ξ равна номеру позиции на которой встретилась первая единица последовательности и 1 на последовательности из одних нулей. Тогда ξ обладает биномиальным распределением.

4.1.7 Гипергеометрическое распределение

Теперь рассмотрим эксперимент с шарами. Пусть в мешке лежат N шаров (уникальных), каждый из которых покрашен в чёрный или в белый цвет. Пусть число белых шаров в мешке равно M . Вытаскивается $1 \leq n \leq N$ шаров (считается, что вероятности вытащить все возможные наборы шаров одинаковы). Обозначим через ξ случайную величину, равную числу вытащенных белых шаров и выясним, чему равна вероятность вытащить m белых шаров среди всех n шаров. Формально пространство элементарных исходов Ω состоит из всевозможных подмножеств мощности n множества всех шаров, а ξ принимает значение равное количеству белых шаров в этом подмножестве. Вероятность будет равна 0, если $m > n$ или $m > M$ или $n - m > N - M$. Разберём остальные случаи. Имеем $|\Omega| = C_N^n$. Число способов выбрать m белых шаров из общего числа M и $n - m$ чёрных шаров из общего числа $N - M$ равно C_M^m и C_{N-M}^{n-m} , соответственно. Поскольку вероятность выбрать данные n шаров равна $\frac{1}{C_N^n}$, то $P(\xi = m) = \frac{C_M^m \cdot C_{N-M}^{n-m}}{C_N^n}$ при всех наложенных ограничениях. Распределение такой случайной величины называется **гипергеометрическим распределением**

4.1.8 Отрицательное биномиальное распределение

Кроме подсчёта числа «успехов», как это делается биномиальной случайной величиной, и подсчёта числа попыток, после которых достигается «успех», как в случае с геометрической случайной величиной, можно считать число попыток, после которых достигается r успехов. Снова будем считать, что в каждом отдельном испытании «успех» достигается с вероятностью p , а неудача происходит с вероятностью q . В качестве вероятностного пространства можно снова взять пространство $\Omega = \mathbb{N}$ с распределением $P(k) = C_{k-1}^{r-1} p^r q^{k-r}$ для $k \geq r$ — вероятность того эксперимент закончится на k -м этапе (то есть в этот момент будет достигнут r -й «успех», а среди предыдущих $k - 1$ будут как-то разбросаны ещё $r - 1$ «успех»). Пусть случайная величина ξ сопоставляет числу из Ω его же само. Величина ξ несёт в себя смысл «количество испытаний, проведённых до r -го успеха». Распределение такой случайной величины называется **геометрическим распределением** и обозначается $NB(r, p)$. Изобразим таблицу:

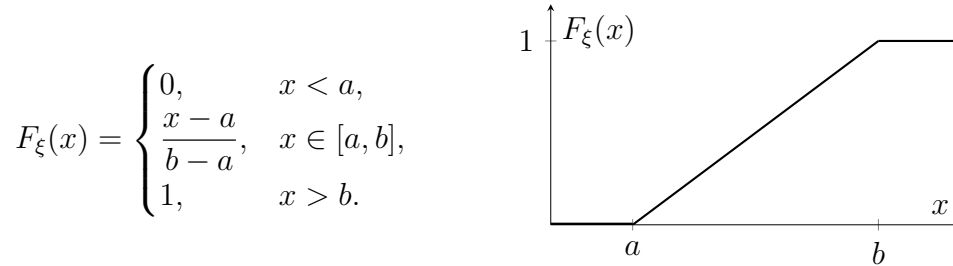
$$\xi \sim \begin{matrix} r & r+1 & r+2 & \dots & k & \dots \\ p^r & C_r^{r-1} p^r q & C_{r+1}^{r-1} p^r q^2 & \dots & C_{k-1}^{r-1} p^r q^{k-r} & \dots \end{matrix} \sim NB(r, p).$$

Пример. Пусть как и в примере с геометрической случайной величиной $\Omega = \{0, 1\}^{\mathbb{N}}$ — множество всех последовательностей из нулей и единиц. Пусть \mathfrak{F} — σ -алгебра, полученная лебеговским продолжением с полукольца последовательностей, имеющих фиксированное общее начало (в нём содержатся, например, последовательности, начинающиеся с 1, с 1101 и т.д., также будем включать в полукольцо всё Ω), по мере m заданной по правилу $m(S) = p^k q^m$ для множества S последовательностей с общим началом $\varepsilon_1 \varepsilon_2 \dots \varepsilon_{k+m}$. Здесь k равно числу единиц в начальном префиксе, а m — числу нулей. Мы опустим проверку того, что это последовательности с общим фиксированным началом образуют полукольцо, а заданная на них функция m является σ -аддитивной мерой. Теперь пусть случайная величина ξ равна номеру позиции на которой встретилась r -я единица последовательности и r , если в последовательности содержится меньше r единиц. Тогда ξ обладает отрицательным биномиальным распределением.

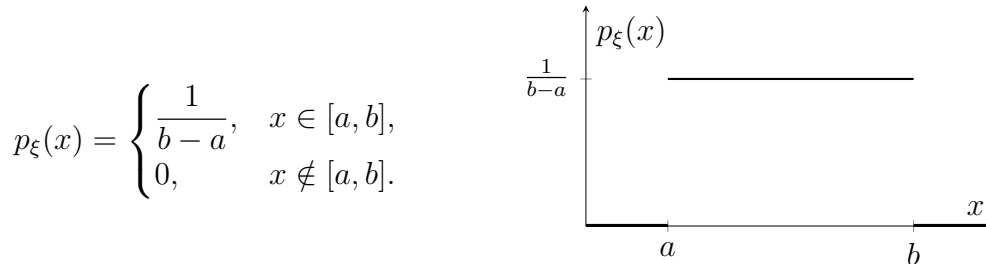
4.2 Абсолютно непрерывные случайные величины

4.2.1 Равномерное распределение

Абсолютно непрерывным аналогом дискретного равномерного распределения выступает равномерное распределение. Пусть $\Omega = [a, b]$ — отрезок, $\mathfrak{F} = \mathcal{B} \cap [a, b]$ — борелевская σ -алгебра и вероятностная мера задана как $P = \frac{1}{b-a} \mu$ — нормированная мера Лебега. Рассмотрим распределение случайной величины ξ , которая сопоставляет точке на отрезке её координату. Можно думать об этом, как об эксперименте в котором на отрезок «бросается» точка и изучаются её вероятности попасть в разные подмножества отрезка. Распределение случайной величины ξ называется **равномерным распределением** и обозначается $R[a, b]$. Функция распределения F_ξ этой случайной величины может быть задана кусочно.



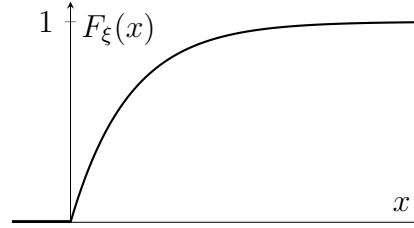
Данная случайная величина обладает функцией плотности.



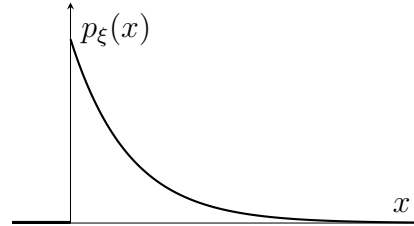
4.2.2 Экспоненциальное (показательное) распределение

Как равномерное распределение выступает аналогом дискретного равномерного распределения, так и экспоненциальное распределение является абсолютно непрерывным аналогом геометрического распределения. Пусть $\Omega = [0, +\infty)$ — бесконечный полуинтервал, $\mathfrak{F} = \mathcal{B} \cap [0, +\infty)$ — борелевская σ -алгебра и вероятностная мера задана на бесконечных полуинтервалах как $P((u; +\infty)) = e^{-\alpha u}$, где $\alpha > 0$ — некоторое число. Рассмотрим распределение случайной величины ξ , которая сопоставляет точке на отрезке её координату. В реальной жизни данная конструкция соответствует, например, ситуации, когда измеряется время до первой поломки электроприбора или до первого прихода автобуса, если такие события происходят часто и независимо, а частота их наступления («интенсивность» α) постоянна. Распределение случайной величины ξ называется **экспоненциальным (или показательным) распределением** и обозначается $E(\alpha)$. Экспоненциальная случайная величина обладает функцией плотности, которую, как и функцию распределения можно задать кусочно. Формулы для функции распределения и плотности приведены ниже.

$$F_{\xi}(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\alpha x}, & x \geq 0. \end{cases}$$



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



Пример. Будем рассматривать переменную x на прямой как переменную означающую «прошедшее время». Для абсолютно непрерывной случайной величины ξ введём **функцию надёжности** $\bar{F}_{\xi}(x) := 1 - F_{\xi}(x) = P(\xi > x)$ выражающую вероятность того, что нечто произойдёт (приедет автобус) или не произойдёт (поломка электроприбора) в течении времени x . Введём также функцию **интенсивности отказов** полагаемую равной $\lambda(x) = \lim_{\Delta x \rightarrow 0} \frac{\bar{F}_{\xi}(x+\Delta x) - \bar{F}_{\xi}(x)}{\bar{F}_{\xi}(x)\Delta x}$. Интенсивность отказов выражает «долю отказавших в единицу времени приборов». Почти всюду функции $\lambda(x)$ и $-\frac{\bar{F}'_{\xi}(x)}{\bar{F}_{\xi}(x)} = \frac{p_{\xi}(x)}{\bar{F}_{\xi}(x)}$ совпадают, поэтому будем также писать $\lambda(x) = \frac{p_{\xi}(x)}{\bar{F}_{\xi}(x)}$.

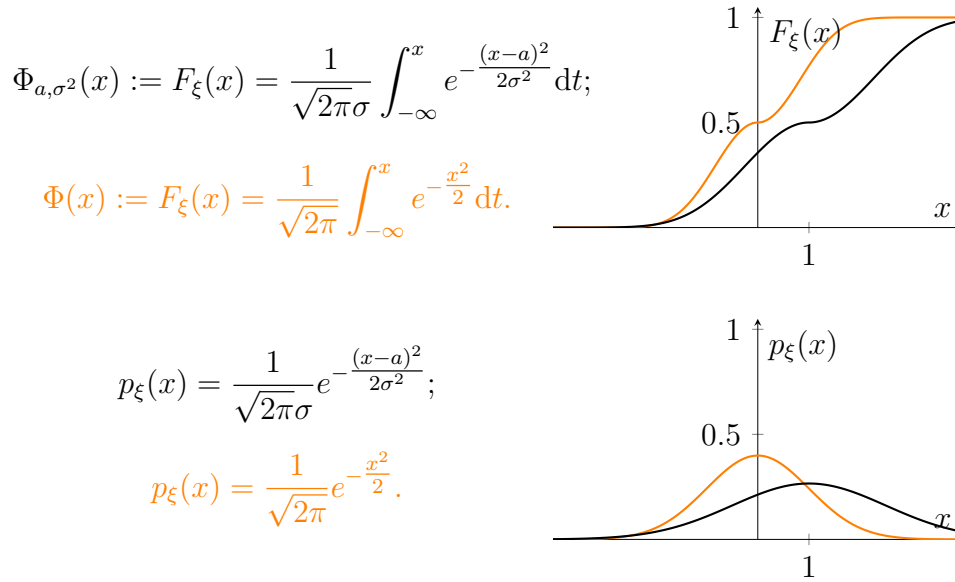
Пусть теперь интенсивность отказов постоянна на полуинтервале $[0; +\infty)$ и равна $\alpha > 0$. Из формулы выше получаем дифференциальное уравнение $\bar{F}'_{\xi} = -\alpha \bar{F}_{\xi}$, откуда $\bar{F}_{\xi}(x) = C e^{-\alpha x}$ на $[0; +\infty)$. Если считать, что $\bar{F}_{\xi}(0) = 1$ («изначально всё работает»), то $C = 1$ и $F_{\xi}(x) = 1 - e^{-\alpha x}$.

В общем случае, решив такую же задачу Коши с начальным условием $\overline{F}_\xi(0) = 1$, получаем формулу для функции распределения:

$$F_\xi(x) = 1 - e^{-\int_0^x \lambda(t) dt}.$$

4.2.3 Нормальное распределение (распределение Гаусса)

Распределение, которое мы рассмотрим сейчас возникает как предел распределения суммы большого количества малых, независимых случайных величин (см. Центральную предельную теорему ниже). Нормальное распределение $\mathcal{N}(a, \sigma^2)$ задаётся двумя параметрами a и $\sigma > 0$, которые, как мы выясним позже равны её математическому ожиданию и квадратному корню из дисперсии, соответственно. Если функция распределения случайной величины ξ имеет вид, указанный ниже, то её распределение называется **нормальным распределением**. Если к тому же параметры имеют особый вид: $a = 0$ и $\sigma^2 = 1$, то распределение называется **стандартным нормальным распределением**.

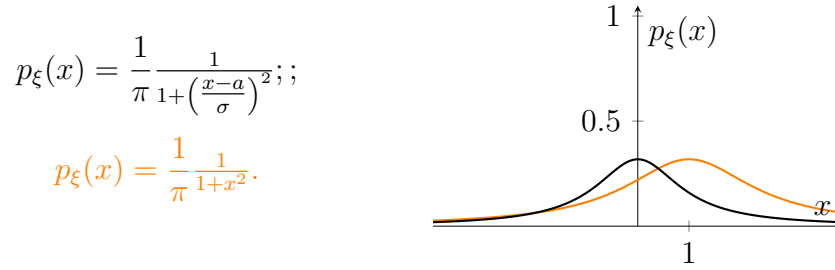
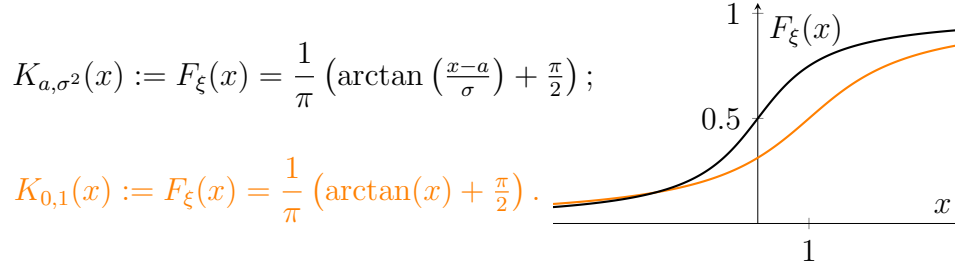


Функция распределения Φ_{a,σ^2} выражается через Φ по формуле $\Phi_{a,\sigma^2}(x) = \Phi(\frac{x-a}{\sigma})$.

4.2.4 Распределение Коши

Ещё одно абсолютно непрерывное распределение, которое мы рассмотрим, похоже на нормальное распределение. Как мы увидим позже, между ними есть существенные различия — распределение Коши является классическим примером распределения, которое не имеет матожидания. Если случайная величина имеет

описываемую формулой ниже функцию распределения, то её распределение называется **распределением Коши**.



Аналогично нормальному распределению имеет место формула $K_{a,\sigma^2}(x) = K_{0,1}\left(\frac{x-a}{\sigma}\right)$.

4.3 Сингулярные распределения

Распределение случайной величины ξ называется **сингулярным распределение**, если её функция распределения F_ξ непрерывна и множество её точек роста

$$\{x \in \mathbb{R} \mid \forall \varepsilon > 0 \ F_\xi(x - \varepsilon) < F_\xi(x + \varepsilon)\}$$

имеет меру нуль по Лебегу.

Пример. Стандартным примером сингулярного распределения служит распределение, соответствующее канторовой лестнице $c: [0, 1] \rightarrow [0, 1]$. Рассмотрим канторово множество

$$K = \bigcap_{i=1}^{+\infty} \bigcup_{j=0}^{2^i-1} \left[\frac{k_j}{3^i}, \frac{k_j+1}{3^i} \right] = \bigcap_{i=1}^{+\infty} \bigcup_{j=0}^{3^{i-1}-1} \left[\frac{3j}{3^i}, \frac{1+3j}{3^i} \right] \cup \left[\frac{2+3j}{3^i}, \frac{3+3j}{3^i} \right] = [0, 1] \setminus \bigcup_{i=1}^{+\infty} \bigcup_{j=0}^{3^{i-1}-1} \left(\frac{1+3j}{3^i}, \frac{2+3j}{3^i} \right),$$

где k_j — j -е число от 0 до 3^i , в троичной записи которого присутствуют только нули и двойки. Всякая точка канторова множеств в троичной системе счисления имеет вид бесконечной троичной дроби с нулями и двойками. Тогда канторова лестница задаётся следующими формулами

$$c(x) = \begin{cases} \frac{j}{2^i}, & x \in \left(\frac{1+3j}{3^i}, \frac{2+3j}{3^i} \right); \\ \sum_{i=1}^{+\infty} \frac{a_i/2}{2^i}, & x \in K, x = \sum_{i=1}^{+\infty} \frac{a_i}{3^i}. \end{cases}$$

Можно показать, что заданная такой формулой функция непрерывна. Множество точек роста s совпадает канторовым множеством, и поэтому имеет меру нуль по Лебегу.

В нашем примере оказалось так, что канторова лестница s имеет производную почти всюду, как и абсолютно непрерывные функции. Однако, в отличие от абсолютно непрерывных функций эта производная равна 0 во всех точках своего определения и поэтому канторову лестницу нельзя вычислить как интеграл от своей производной.

5 Совместные распределения случайных величин

Пусть ξ_1, \dots, ξ_n — случайные величины, заданные на вероятностном пространстве $(\Omega, \mathfrak{F}, P)$. Тогда функция $\xi: \Omega \rightarrow \mathbb{R}^n$, заданная по правилу $\xi(\omega) = (\xi_1(\omega), \dots, \xi_n(\omega))$ называется **случайным вектором** или **многомерной случайной величиной**.

Предложение 5.1. *Справедливы следующие утверждения о случайных величинах и многомерных случайных величинах.*

- (1) Многомерная случайная величина является измеримым отображением $\xi: (\Omega, \mathfrak{F}) \rightarrow (\mathbb{R}^n, \mathcal{B}_n)$, где \mathcal{B}_n — борелевская σ -алгебра подмножеств \mathbb{R} .
- (2) Если $\xi: (\Omega, \mathfrak{F}) \rightarrow (\mathbb{R}^n, \mathcal{B}_n)$ — измеримое отображение, то каждая композиция $\xi_i = \text{pr}_i \circ \xi$ ($\xi(\omega) = (\xi_1(\omega), \dots, \xi_n(\omega))$) является случайной величиной и тем самым ξ является многомерной случайной величиной.

Доказательство. Оба утверждения являются следствиями теоремы 0.38. □

Аналогично одномерным случайным величинам для случайного вектора определяются его **распределение** как **прямой образ** вероятностной меры $\xi_* P$. Снова напомним, что $\xi_* P(B) := P(\xi^{-1}(B)) = P(\xi \in B)$. Аналогично же определяется **функция распределения случайного вектора (функция совместного распределения)**

$$F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = P(\xi_1 \leq x_1, \dots, \xi_n \leq x_n).$$

Функция совместного распределения удовлетворяет свойствам, аналогичным свойствам одномерной функции распределения.

Предложение 5.2. *Пусть ξ_1, \dots, ξ_n — случайные величины и F_{ξ_1, \dots, ξ_n} — её функция распределения. Тогда выполнены следующие свойства*

- (1) функция F_{ξ_1, \dots, ξ_n} неубывает по каждому аргументу;
- (2) функция F_{ξ_1, \dots, ξ_n} непрерывна справа по каждому аргументу;
- (3) $\lim_{x_n \rightarrow +\infty} F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = F_{\xi_1, \dots, \xi_{n-1}}(x_1, \dots, x_{n-1})$;

$$(4) \quad \lim_{x_n \rightarrow -\infty} F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = 0;$$

$$(5) \quad \lim_{x_1 \rightarrow +\infty \dots x_n \rightarrow +\infty} F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = 1;$$

По функции F , удовлетворяющей свойствам из предложения 5.2 можно построить многомерную меру Лебега-Стилтьеса μ_F . Остаются верны следующие результаты.

Теорема 5.3. Пусть функция F удовлетворяет свойствам из предложения 5.2. Тогда существует единственная вероятностная мера P_F на борелевской σ -алгебре \mathcal{B}_n подмножеств \mathbb{R}^n такая, что функция F является функцией совместного распределения многомерной случайной величины $\text{id}: (\mathbb{R}^n, \mathcal{B}_n, P_F) \rightarrow (\mathbb{R}^n, \mathcal{B}_n)$.

Теорема 5.4. Пусть ξ — случайный вектор и F_ξ его функция распределения. Пусть μ_ξ — мера Лебега-Стилтьеса на подмножествах \mathbb{R}^n , порождённая функцией F_ξ . На борелевской σ -алгебре \mathcal{B}_n мера μ_ξ совпадает с распределением $\xi_* P$ вероятностной меры P .

Многомерная случайная величина называется **дискретной**, если она принимает не более, чем счётное число значений. Распределение такой случайной величины называется **дискретным**.

Распределение случайного вектора $\xi = (\xi_1, \dots, \xi_n)$ и он сам называются **абсолютно непрерывными**, если существует функция $p: \mathbb{R}^n \rightarrow \mathbb{R}$ такая, что для любого борелевского множества $B \in \mathcal{B}_n$ выполнено равенство

$$P(\xi \in B) = \xi_* P(B) = \int_B f d\mu^{(n)},$$

где $\mu^{(n)}$ — классическая мера Лебега на подмножествах \mathbb{R}^n . С точки зрения теории меры и интеграла Лебега это определение эквивалентно тому, что мера $\xi_* P$ абсолютно непрерывна относительно классической меры Лебега $\mu^{(n)}$ в многомерном пространстве. (это значит, что случайная величина ξ попадает в множество меры нуль с вероятностью нуль). Теорема Радона-Никодима даёт формулу для $\xi_* P$, принятую нами за определение. Функция $p_\xi = p_{\xi_1, \dots, \xi_n}$ называется **функцией совместной плотности** абсолютно непрерывного случайного вектора $\xi = (\xi_1, \dots, \xi_n)$.

В силу доказанных выше теорем мы можем дать эквивалентное определение в терминах функции совместного распределения. Будем говорить, что совместное распределение случайных величин ξ_1, \dots, ξ_n называется **абсолютно непрерывными**, если функция совместного распределения F_{ξ_1, \dots, ξ_n} может быть выражена как интеграл

$$F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = \int_{\prod_{i=1}^n (-\infty; x_i]} p_{\xi_1, \dots, \xi_n} d\mu^{(n)}.$$

Почти всюду существует частная производная $\frac{\partial^n F_{\xi_1, \dots, \xi_n}}{\partial x_1 \dots \partial x_n}(\tilde{x}_1, \dots, \tilde{x}_n)$, которая почти всюду совпадает с $p_{\xi_1, \dots, \xi_n}(\tilde{x}_1, \dots, \tilde{x}_n)$.

Во многих рассматриваемых случаях (в частности, во всех примерах выше) функция p_ξ оказывается интегрируемой по Риману на любом отрезке и абсолютно интегрируема на \mathbb{R} , поэтому можно записать равенства

$$F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = \int_{-\infty}^{x_n} \dots \int_{-\infty}^{x_1} p_{\xi_1, \dots, \xi_n}(t_1, \dots, t_n) dt_1 \dots dt_n,$$

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} p_{\xi_1, \dots, \xi_n}(t_1, \dots, t_n) dt_1 \dots dt_n = 1.$$

Также верна формула

$$p_{\xi_1, \dots, \xi_{n-1}}(t_1, \dots, t_{n-1}) = \int_{-\infty}^{+\infty} p_{\xi_1, \dots, \xi_n}(t_1, \dots, t_n) dt_n.$$

Пример. Пусть $a = (a_1, \dots, a_n)^T$ — вектор, $B = (\sigma_{ij})$ — положительно определённая симметрическая матрица. Будем говорить, что вектор $\xi = (\xi_1, \dots, \xi_n)$ имеет многомерное нормальное распределение, если это распределение имеет совместную плотность

$$p_\xi(x) = \frac{1}{(2\pi)^{n/2} \sqrt{\det B}} e^{-\frac{(x-a)^T B^{-1} (x-a)}{2}}.$$

Используя замену координат и приведение матрицы B к единичному виду, можно показать, что

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} p_{\xi_1, \dots, \xi_n}(t_1, \dots, t_n) dt_1 \dots dt_n = 1.$$

6 Независимость случайных величин

6.1 Независимость конечного набора случайных величин

Введём понятие независимости случайных величин, с которым мы будем в дальнейшем сталкиваться повсеместно. Интуиция говорит там, что случайные величины независимы, если вероятности их принять какое-то значения не связаны или другими словами независимы события вида $\xi_k \in B_k$, где B_k — **борелевские множества**. Случайные величины ξ_1, \dots, ξ_n называются **независимыми**, если для любого набора борелевских множеств B_1, \dots, B_n выполнено равенство

$$P(\xi_1 \in B_1, \dots, \xi_n \in B_n) = P(\xi_1 \in B_1) \dots P(\xi_n \in B_n).$$

По всякой случайной величине можно построить σ -алгебру событий \mathfrak{F}_ξ , состоящую из всевозможных прообразов борелевских множеств при ξ : $\mathfrak{F}_\xi = \{\xi^{-1}(B) \mid B \in \mathcal{B}\}$. Построенную σ -алгебру мы будем называть **σ -алгеброй, порождённой случайной величиной ξ** . Отметим, что данная σ -алгебра есть не что иное как, **обратный образ** борелевской σ -алгебры при отображении ξ . Проверка того, что это \mathfrak{F}_ξ действительно является σ -алгеброй осуществляется непосредственной проверкой всех аксиом. Кроме обозначения \mathfrak{F}_ξ , будем также писать $\sigma(\xi)$.

Предложение 6.1. Пусть ξ, η — две случайные величины и $\mathfrak{F}_\xi \subset \mathfrak{F}_\eta$ (в таких случаях ещё говорят, что ξ измерима относительно \mathfrak{F}_η , поскольку включение означает, что прообразы борелевских множеств при отображении ξ лежат в \mathfrak{F}_η). Тогда существует такая борелевская функция g , что $\xi = g(\eta)$.

Доказательство. Положим $A_{k,n} = \xi^{-1}\left(\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)\right)$. Так как $A_{k,n} \in \mathfrak{F}_\xi \subset \mathfrak{F}_\eta$, то найдутся борелевские множества $B_{k,n}$ такие, что $\eta^{-1}(B_{k,n}) = A_{k,n}$. Положим $g_n(x) = \frac{k}{2^n}$ на $B_{k,n} \setminus \bigcup_{k' \neq k, n' \neq n} B_{k',n'}$ и нулём там, где хотя бы два из множеств $B_{k,n}$ пересекаются (все такие точки лежат вне образа случайной величины η , так как множества $A_{k,n}$ попарно не пересекаются). Тогда g_n является борелевской и $g_n(\eta(\omega)) = \frac{\lfloor 2^n \xi(\omega) \rfloor}{2^n}$. Кроме того, функции g_n образуют поточечно возрастающую последовательность, поточечно ограниченную ξ или нулём и такую, что композиции $g_n \circ \eta$ стремятся к ξ . Следовательно, во всех точках существует предел $g(x) = \lim_{n \rightarrow +\infty} g_n(x)$, являющийся борелевской функцией и выполнено равенство $g(\eta) = \xi$. \square

Две (σ) -алгебры $\mathfrak{F}_1, \dots, \mathfrak{F}_n \subset \mathfrak{F}$ будем называть **независимыми**, если события из любого набора $A_1 \in \mathfrak{F}_1, \dots, A_n \in \mathfrak{F}_n$ **независимы**. Поскольку всегда в качестве некоторых A_i можно взять Ω , то независимость σ -алгебр $\mathfrak{F}_1, \dots, \mathfrak{F}_n \subset \mathfrak{F}$ равносильна истинности для всех наборов $A_1 \in \mathfrak{F}_1, \dots, A_n \in \mathfrak{F}_n$ равенства

$$P(A_1 \dots A_n) = P(A_1) \dots P(A_n).$$

Ввиду данного определения возникает описание независимости случайных величин в терминах порождённых ими σ -алгебр.

Предложение 6.2. Случайные величины независимы тогда и только тогда, когда порождённые ими σ -алгебры независимы.

Доказательство. Взяв в качестве $A_i = \xi^{-1}(B_i)$, где B_i борелевское, дословно перепишем друг в друга равенства из обоих определений. \square

Следствие 6.3. Пусть g, h — борелевские функции, а ξ, η — независимые случайные величины. Тогда случайные величины $g(\xi)$ и $h(\eta)$ независимы.

Доказательство. Имеем включения σ -алгебр $\mathfrak{F}_{g(\xi)} \subset \mathfrak{F}_\xi$ и $\mathfrak{F}_{h(\eta)} \subset \mathfrak{F}_\eta$. По предложению 6.2 большие σ -алгебры \mathfrak{F}_ξ и \mathfrak{F}_η независимы. Тогда из определения независимости алгебр следует, что содержащиеся в них σ -алгебры $\mathfrak{F}_{g(\xi)}$ и $\mathfrak{F}_{h(\eta)}$ тоже независимы. Вновь по предложению 6.2 случайные величины $g(\xi)$ и $h(\eta)$ независимы. \square

Теорема 6.4 (Об аппроксимации). Пусть $(\Omega, \mathfrak{F}, P)$ — вероятностное пространство и \mathfrak{A} — σ -алгебра, порождённая алгеброй $\mathcal{A} \subset \mathfrak{F}$. Тогда для всякого элемента $U \in \mathfrak{A}$ существует последовательность элементов $\{U_i\} \subset \mathcal{A}$ такая, что $\lim_{i \rightarrow +\infty} P(U_i \Delta U) = 0$, и, следовательно, $\lim_{i \rightarrow +\infty} P(U_i) = P(U)$.

Доказательство. Построим по алгебре \mathcal{A} и ограничению на неё меры P их лебеговское продолжение, как это делалось в предварительных сведениях. Тогда, новая σ -алгебра измеримых множеств будет содержать в себе алгебру \mathfrak{A} . По теореме Каратеодори 0.17 меры P и продолжение по Лебегу этой же меры с \mathcal{A} совпадают на \mathfrak{A} . Из измеримости множества U и конструкции лебеговского продолжения следует существование множеств $U_i \in \mathcal{A}$, удовлетворяющих условию. Второе выражение с пределом следует из неравенства $|P(U) - P(U_i)| \leq P(U \Delta U_i)$. \square

Теорема 6.5. Пусть $\mathfrak{F}_1, \dots, \mathfrak{F}_n \subset \mathfrak{F}$ — σ -алгебры и σ -алгебра σ_i порождается алгеброй \mathcal{A}_i . Тогда независимость данных σ -алгебр равносильна независимости всех возможных наборов $A_i \in \mathcal{A}_i$ (независимости алгебр \mathcal{A}_i).

Доказательство. Независимость наборов из условия следует из определения независимости σ -алгебр.

Пусть теперь все алгебры \mathcal{A}_i из условия независимы. Пусть $U_i \in \mathfrak{F}_i$ — набор событий. По лемме об аппроксимации 6.4 существуют последовательности $U_{ij} \in \mathcal{A}_i$ такие, что $\lim_{j \rightarrow +\infty} P(U_{ij}) = P(U_i)$. Так как

$$P((U_1 \dots U_n) \Delta (U_{1j} \dots U_{nj})) \leq \sum_{i=1}^n P(U_i \Delta U_{ij}),$$

то существует и предел $\lim_{j \rightarrow +\infty} P(U_{1j} \dots U_{nj}) = P(U_1 \dots U_n)$. Тогда

$$P(U_1) \dots P(U_n) = \lim_{j \rightarrow +\infty} P(U_{1j}) \dots P(U_{nj}) = \lim_{j \rightarrow +\infty} P(U_{1j} \dots U_{nj}) = P(U_1 \dots U_n).$$

\square

Следующая теорема демонстрирует связь между независимостью, функциями распределения и совместным распределением.

Теорема 6.6. Случайные величины ξ_1, \dots, ξ_n независимы тогда и только тогда, когда для любого набора (x_1, \dots, x_n) имеет место равенство

$$F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = F_{\xi_1}(x_1) \cdot \dots \cdot F_{\xi_n}(x_n).$$

Доказательство. Если случайные величины независимы, то взяв в качестве $B_i = (-\infty; x_i]$ получим требуемое равенство.

Предположим, что выполнено равенство из условия. В силу предложения 6.2 достаточно доказать, что независимы σ -алгебры $\mathfrak{F}_{\xi_1}, \dots, \mathfrak{F}_{\xi_n}$. Обозначим через

$B_i(x)$ множество $\xi_i^{-1}((-\infty; x])$ и через $C_i(a, b]$ множество $\xi_i^{-1}((a, b])$. По условию, для любых x_1, \dots, x_n события $B_1(x_1), \dots, B_n(x_n)$ независимы. Так как $C_1(a, b] = B_1(b) \setminus B_1(a)$ и выполнены равенства

$$\begin{aligned} P(C_i(a, b] B_2(x_2) \dots B_n(x_n)) &= P(B_1(b) B_2(x_2) \dots B_n(x_n)) - P(B_1(a) B_2(x_2) \dots B_n(x_n)) = \\ &= P(B_1(b)) P(B_2(x_2)) \dots P(B_n(x_n)) - P(B_1(a)) P(B_2(x_2)) \dots P(B_n(x_n)) = \\ &= P(C_i(a, b]) P(B_2(x_2)) \dots P(B_n(x_n)), \end{aligned}$$

по индукции, а также аналогично для интервалов $(a; +\infty)$ можем вывести, что для любых $a_1, \dots, a_n, b_1, \dots, b_n$ независимы события A_1, \dots, A_n , где $A_i \in S_i$ — принадлежат полукольцу полуинтервалов. Из построения минимального кольца 0.7 и свойств независимости событий 2.7 вытекает, что алгебры, порождённые прообразами всех полуинтервалов независимы. Так как минимальная сигма-алгебра, порождённая полуинтервалами совпадает с борелевской σ -алгеброй, то по лемме 0.37 их прообразы порождают σ -алгебры \mathfrak{F}_{ξ_i} . По теореме 6.5 эти σ -алгебры независимы. \square

Следствие 6.7. Пусть абсолютно непрерывные случайные величины ξ_1, \dots, ξ_n имеют плотности $p_{\xi_1}, \dots, p_{\xi_n}$. Тогда их независимость равносильна тому, что их совместное распределение абсолютно непрерывно с плотностью

$$p_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = p_{\xi_1}(x_1) \dots p_{\xi_n}(x_n).$$

Доказательство. В доказательстве мы будем писать интегралы в нотации интеграла Римана, подразумевая, что они вычисляются любым допустимым способом (например, по Лебегу).

По теореме 6.6 из независимости имеем равенство

$$F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = F_{\xi_1}(x_1) \cdot \dots \cdot F_{\xi_n}(x_n).$$

Тогда

$$\begin{aligned} F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) &= \int_{-\infty}^{x_1} p_{\xi_1}(t_1) dt_1 \dots \int_{-\infty}^{x_n} p_{\xi_n}(t_n) dt_n = \\ &= \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} p_{\xi_1}(t_1) \dots p_{\xi_n}(t_n) dt_n \dots dt_1. \end{aligned}$$

Обратно, если верна формула выше, то проведя выкладки в обратном порядке и воспользовавшись теоремой 6.6 в другую сторону, получим независимость. \square

6.2 Последовательность независимых случайных величин

Последовательность случайных величин $\{\xi_i\}$ будем называть **последовательностью независимых случайных величин**, если для всякого $n \in \mathbb{N}$ случайные величины ξ_1, \dots, ξ_n независимы.

7 Численные характеристики случайных величин

В этом разделе мы введём различные величины, характеризующие случайные величины, такие как математическое ожидание, выражающее «усреднённое значение» случайной величины, дисперсию, показывающую, насколько случайная величина отклоняется от своего «среднего» и другие.

7.1 Математическое ожидание, моменты и абсолютные моменты

Пусть ξ — случайная величина. Если существует интеграл Лебега $\int_{\Omega} \xi dP$, то говорят, что ξ **обладает математическим ожиданием**, величина интеграла называется **математическим ожиданием случайной величины ξ** и обозначается $E\xi$. Интегрируемость случайной величины ξ также будем записывать формулой $\xi \in \mathcal{L}^1(\Omega)$ или $\xi \in \mathcal{L}^1$.

Для дискретной случайной величины ξ , принимающей значения x_1, x_2, \dots с вероятностями p_1, p_2, \dots , было бы естественно определить её математическое ожидание как сумму ряда $\sum_{i=1}^{+\infty} x_i p_i$, а его существование как его абсолютную сходимость.

И следующее предложение позволяет нам сделать это

Предложение 7.1. Пусть ξ — дискретная случайная величина, принимающая значения x_1, x_2, \dots с вероятностями p_1, p_2, \dots , соответственно. Тогда математическое ожидание существует тогда и только тогда, когда абсолютно сходится ряд $\sum_{i=1}^{+\infty} x_i p_i$, и в случае сходимости равно его сумме.

Доказательство. Дискретная случайная величина является **обобщённой простой функцией** на Ω , поэтому утверждение следует из предложения 0.27. \square

Пример. Есть N лотерейных билетов, причём на m_1, \dots, m_n из них приходятся соответственно выигрыши a_1, \dots, a_n . Разыгрывается денежная сумма

$$S = a_1 m_1 + a_2 m_2 + \dots + a_n m_n.$$

Тогда средний выигрыш, приходящийся на один билет, равен

$$\frac{S}{N} = \frac{a_1 m_1 + \dots + a_n m_n}{N} = a_1 \frac{m_1}{N} + \dots + a_n \frac{m_n}{N} = a_1 p_1 + \dots + a_n p_n,$$

где $p_i = \frac{m_i}{N}$.

Введём в данном примере вероятностное пространство Ω , состоящее из N точек, и распределение вероятностей, сопоставляющие точкам из непересекающихся подмножеств по m_i точек вероятности p_i . Тогда, если случайная величина ξ сопоставляет точкам из i -го подмножества число («выигрыш») a_i , то $E\xi$ существует и равно $\frac{S}{N}$.

Пример. Пусть на прямой в точках x_1, \dots, x_n сосредоточены массы p_1, \dots, p_n . Тогда центр тяжести этой системы равен

$$\bar{x} = \frac{x_1 p_1 + x_2 p_2 + \dots + x_n p_n}{p_1 + p_2 + \dots + p_n}.$$

Формально можно рассмотреть меру введенную на числовой прямой и сосредоточенную в точках с координатами x_1, \dots, x_n , а также случайную величину, сопоставляющую точке её координату. Тогда матожидание оказывается равным в точности \bar{x} .

Пример. Если $\xi = c$ — константа, то $\mathbb{E}\xi = c$. Если $\xi = \mathbb{1}_A$, то $\mathbb{E}\xi = P(A)$.

Таким образом, для бернуллиевской случайной величины $\xi \sim \text{Be}(p)$ с параметром p выполнено равенство:

$$\mathbb{E}\xi = p.$$

Предложение 7.2. Случайная величина ξ обладает математическим ожиданием тогда и только тогда, когда $|\xi|$ обладает математическим ожиданием.

Доказательство. Пусть $\{\xi_n\}$ — последовательность интегрируемых обобщённых простых функций, равномерно сходящаяся к ξ . Тогда функции $|\xi_n|$ интегрируемы и равномерно сходятся к $|\xi|$ из чего следует, что $|\xi|$ интегрируема.

Обратно, пусть $\{\xi_n\}$ — произвольная последовательность, равномерно сходящаяся к ξ . Снова $\{|\xi_n|\}$ равномерно сходится к $|\xi|$. Поскольку $|\xi|$ интегрируема, то начиная с некоторого члена функции $|\xi_n|$ интегрируемы по 0.30. Отсюда следует интегрируемость ξ . \square

Предложение 7.3. Имеют место следующие свойства.

- (1) $\mathbb{E}c = c$ для любой константы c ;
- (2) $\mathbb{E}\mathbb{1}_A = P(A)$ для любого события A ;
- (3) если η, ξ — неотрицательные случайные величины, для всякого $\omega \in \Omega$ имеем место $\eta(\omega) < \xi(\omega)$ ($=: \eta \leq \xi$) и $\xi \in \mathcal{L}^1$, то $\eta \in \mathcal{L}^1$;
- (4) если $\xi, \eta \in \mathcal{L}^1$, то $a\xi, \xi + \eta \in \mathcal{L}^1$ для любой константы $a \in \mathbb{R}$;
- (5) если $\xi, \eta \in \mathcal{L}^1$, то $\mathbb{E}(a\xi) = a\mathbb{E}\xi$ и $\mathbb{E}(\xi + \eta) = \mathbb{E}\xi + \mathbb{E}\eta$ для любой константы $a \in \mathbb{R}$;
- (6) если $\xi, \eta \in \mathcal{L}^1$ и $\xi \leq \eta$, то $\mathbb{E}\xi \leq \mathbb{E}\eta$.

Доказательство. Свойства (1) и (2) следуют из предложения о дискретных случайных величинах.

Свойство (3) следует из признака сравнения абсолютной сходимости рядов (мы применяем его к последовательности равномерно сходящихся к данной обобщённых простых функций).

Свойства (4) и (5) также следуют из абсолютной сходимости суммы двух абсолютно сходящихся рядов и абсолютно сходящегося ряда, умноженного на коэффициент.

Свойство (6) следует из того, что аналогичное неравенство возникает для абсолютно сходящихся рядов — интегралов обобщённых простых функций. \square

Теперь покажем, как с помощью функций распределения и плотностей можно посчитать математическое ожидание.

Лемма 7.4. Пусть λ — конечная абсолютно непрерывная относительно меры μ мера на измеримом пространстве (Ω, \mathfrak{F}) и $\lambda(A) = \int_A p d\mu$ для некоторой интегрируемой по μ функции p . Пусть $\xi: (\Omega, \mathfrak{F}) \rightarrow (\mathbb{R}, \mathcal{B})$ — измеримая функция. Тогда ξ интегрируема относительно меры λ тогда и только тогда, когда $\xi \cdot p$ интегрируема относительно меры μ и в случае интегрируемости равны интегралы

$$\int_{\Omega} \xi d\lambda = \int_{\Omega} \xi \cdot p d\mu.$$

Доказательство. Докажем утверждение для случая конечной меры μ . Если мера μ является σ -конечной, то сначала воспользуемся теоремой для множеств Ω_k из исчерпания Ω , а после воспользуемся сходимостью интегралов по исчерпанию для одной из сторон, чтобы доказать сходимую другую.

Пусть сначала $\xi = \sum_{i=1}^{+\infty} a_i \mathbb{1}_{A_i}$ — обобщённая простая функция. Имеем цепочку равенств

$$\int_{\Omega} \xi d\lambda = \sum_{i=1}^{+\infty} a_i \lambda(A_i) = \sum_{i=1}^{+\infty} a_i \int_{A_i} p d\mu = \sum_{i=1}^{+\infty} \int_{A_i} \xi \cdot p d\mu = \int_{\Omega} \xi \cdot p d\mu,$$

где последнее равенство и одновременная сходимость выполнены в силу σ -аддитивности интеграла Лебега. \square

Теорема 7.5 (О способах вычисления математического ожидания). Пусть ξ — случайная величина. Тогда ξ обладает математическим ожиданием тогда и только тогда, когда сходится интеграл Лебега-Стилтьеса

$$\int_{-\infty}^{+\infty} x dF_{\xi}.$$

Кроме того, когда интеграл сходится, он равен математическому ожиданию $E\xi$. Если ξ абсолютно непрерывна с плотностью p_ξ , то существование математического ожидания $E\xi$ равносильно сходимости интеграла Лебега

$$\int_{-\infty}^{+\infty} x \cdot p_\xi d\mu,$$

а в случае сходимости равно этому интегралу.

Доказательство. Поскольку $x \circ \xi = \xi$, то по теореме о прямом образе меры 0.32 имеем одновременную интегрируемость и равенство

$$\int_{\Omega} \xi dP = \int_{\mathbb{R}} x d\xi_* P.$$

Поскольку мы рассматриваем для вычисления второго интеграла борелевскую σ -алгебру и на ней по теореме 3.6 совпадают мера Лебега-Стилтьеса, порождённая F_ξ , и прямой образ меры $\xi_* P$, то справедливо первое утверждение теоремы.

Эквивалентность второму способу следует из леммы 7.4. \square

Теорема 7.6. Пусть g — борелевская функция и ξ — случайная величина. Тогда математическое ожидание $g(\xi)$ существует тогда и только тогда, когда сходится интеграл

$$\int_{-\infty}^{+\infty} g(x) dF_\xi,$$

и равно этому интегралу в случае сходимости. Если ξ абсолютно непрерывна с плотностью p_ξ , то существование математического ожидания $E\xi$ равносильно сходимости интеграла Лебега

$$\int_{-\infty}^{+\infty} g \cdot p_\xi d\mu,$$

а в случае сходимости равно этому интегралу.

Доказательство. Полностью повторяет доказательство теоремы 7.5, с точностью до замены x на $g(x)$. \square

Поскольку функция распределения монотонна и ограничена, то она является функцией ограниченной вариации и для непрерывных g первый интеграл равен соответствующему (существующему) интегралу Римана-Стилтьеса. Если функция плотности оказывается интегрируемой по Риману на всех промежутках, а g непрерывна, то второй интеграл в условии теоремы можно заменить на интеграл Римана.

7.2 Моменты и абсолютные моменты случайной величины

Величины $E\xi^n$ обозначаются α_n и называются **моментами случайной величины ξ** . Величины $E|\xi|^n$ обозначаются β_n и называются **абсолютными моментами случайной величины ξ** . Доказанное выше предложение говорит нам, что k -й момент существует тогда и только тогда, когда существует k -й абсолютный момент и $\alpha_n \leq \beta_n$.

Лемма 7.7. Пусть у случайной величины ξ существует k -й момент $E\xi^k$. Тогда у ξ существуют все моменты с меньшими номерами.

Доказательство. По сказанному выше, у ξ существует $E|\xi|^k$. Достаточно доказать, что для всякого $j < k$ существует момент $E|\xi|^j$. Воспользуемся теоремой 7.6 и рассмотрим интеграл:

$$\int_{\mathbb{R}} |x|^j dF_{\xi} = \int_{[-1,1]} |x|^j dF_{\xi} + \int_{\mathbb{R} \setminus [-1,1]} |x|^j dF_{\xi}.$$

Первый интеграл всегда сходится, так как $|x|^n$ непрерывна, а F_{ξ} монотонна. Вторым же интеграл существует, так как неотрицательная функция $|x|^j$ не превосходит $|x|^k$ при $j < k$ и $|x| > 1$. \square

7.3 Дисперсия, ковариация и корреляция

Кроме «среднего значения» нам также хотелось бы знать абсолютное отклонение от него, для чего мы и вводим понятие **дисперсии случайной величины ξ** определяемой по формуле $D\xi = E(\xi - E\xi)^2$.

Предложение 7.8. Пусть у случайной величины ξ существует математическое ожидание $E\xi$. Тогда дисперсия тогда и только тогда, когда существует математическое ожидание $E\xi^2$ (далее будет сказано, что эта величина называется вторым моментом случайной величины ξ) и в случае существования верна формула

$$D\xi = E\xi^2 - (E\xi)^2.$$

Доказательство. Имеем равенство $(\xi - E\xi)^2 = \xi^2 - 2\xi \cdot E\xi + (E\xi)^2$. Поскольку $E\xi$ есть константа и ξ обладает математическим ожиданием, то сумма обладает им тогда и только тогда, когда существует $E\xi^2$. По линейности имеем

$$D\xi = E(\xi^2 - 2\xi \cdot E\xi + (E\xi)^2) = E\xi^2 - 2(E\xi)(E\xi) + (E\xi)^2 = E\xi^2 - (E\xi)^2.$$

\square

Предложение 7.9. Дисперсия удовлетворяет следующим свойствам:

- (1) $Dc = 0$ для любой константы c ;

- (2) если существует $D\xi = 0$, то для некоторой константы c выполнено $P(\xi = c) = 1$ (в частности $c = E\xi$);
- (3) если существует дисперсия $D\xi$, то для любой константы c существует $D(c\xi) = c^2 D\xi$.
- (4) если у случайных величин ξ и η существуют матожидания и дисперсии, то

$$D(\xi + \eta) = D(\xi) + D(\eta) + 2E(\xi - E\xi)(\eta - E\eta).$$

Доказательство. Все свойства следуют из определения и формулы доказанной выше. Мы докажем только второе и третье. Если $D\xi = 0$, то $\xi - E\xi$ равно нулю почти наверное ($:=$ почти всюду). Положим $c = E\xi$. Тогда $P(\xi = c) = 1$.

$$\begin{aligned} D(\xi + \eta) &= E(\xi + \eta - E(\xi + \eta))^2 = E((\xi - E\xi) + (\eta - E\eta))^2 = \\ &= E(\xi - E\xi)^2 + E(\eta - E\eta)^2 + 2E(\xi - E\xi)(\eta - E\eta) = D(\xi) + D(\eta) + 2E(\xi - E\xi)(\eta - E\eta). \end{aligned}$$

□

Последнее доказанное свойство показывает, что дисперсия является квадратичной формой на пространстве случайных величин (и даже более того, положительно определённой формой, но на факторпространстве случайных величин по величинам почти всюду равным константе). Соответствующую ей (поляризованную) билинейную функцию $E(\xi - E\xi)(\eta - E\eta)$ называют **ковариацией случайных величин ξ и η** и обозначают $\text{Cov}(\xi, \eta)$.

Предложение 7.10. Пусть L — векторное пространство случайных величин обладающих матожиданием и дисперсией рассмотренных с точностью до случайных величин почти наверное равных константе. Тогда Cov на L является положительно определённой квадратичной формой (скалярным произведением), а D — соответствующей ей квадратичной формой.

Доказательство. Если случайная величина ξ почти наверное равна константе, то $\xi - E\xi$ почти наверное равно нулю и $\text{Cov}(\xi, \eta) = 0$ для любой η , так что Cov корректно определена на данном векторном пространстве. Билинейность Cov следует из линейности интеграла и билинейности произведения. Положительная определённость Cov доказана в предложении 7.9. □

7.4 Связь с независимостью случайных величин

7.5 Мода, медиана

7.6 Квантили, асимметрия и эксцесс

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

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

7.7 Вычисление матожиданий и дисперсий некоторых случайных величин

8 Сходимости случайных величин

8.1 Сходимость почти наверное

8.2 Сходимость по вероятности

8.3 Пространство \mathcal{L}_p и сходимость в нём

8.4 Сходимость по распределению

8.5 Связь сходимостей

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

9 Производящие функции

НУЖНО: записать определение

10 Характеристические функции

Теорема 10.1 (Бохнер, Хинчин). Пусть $f: \mathbb{R} \rightarrow \mathbb{C}$ — непрерывная функция и $f(0) = 1$. Тогда следующие утверждения о функции f эквивалентны

- (1) функция $f = f_\xi$ является характеристической функцией некоторой случайной величины ξ ;
- (2) функция f неотрицательно определена, то есть для любого натурального n и любых наборов $t_1, \dots, t_n \in \mathbb{R}$, $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ вещественных и комплексных чисел неотрицательна сумма

$$\sum_{i,j=1}^n f(t_i - t_j) \lambda_i \overline{\lambda_j} \geq 0.$$

11 Пределные теоремы

НУЖНО: дописать ниже доказательства теорем

11.1 Неравенства

Теорема 11.1 (Неравенства Чебышёва и Маркова). Пусть ξ — неотрицательная случайная величина. Тогда при дополнительных условиях имеют место следующие неравенства.

- (1) Если существует математическое ожидание $E\xi$, то справедливо первое неравенство Чебышёва $P(\xi \geqslant E\xi) \leqslant \frac{E\xi^2}{(E\xi)^2}$.

11.2 Закон больших чисел

Теорема 11.2 (Закон больших чисел в форме Бернулли).

Теорема 11.3 (Закон больших чисел в форме Чебышёва).

Теорема 11.4 (Усиленный закон больших чисел).

Теорема 11.5 (Закон больших чисел в форме Хинчина). *content*

11.3 Теорема Муавра-Лапласа

Теорема 11.6 (Теорема Пуассона).

Теорема 11.7 (Формула Стирлинга).

Теорема 11.8 (Муавр, Лапласа).

11.4 Закон нуля или единицы

Лемма 11.9 (Борель, Кантелли).

Лемма 11.10 (Борель, Кантелли).

Теорема 11.11 (Закон нуля или единицы Колмогорова).

11.5 Закон повторного логарифма

Теорема 11.12 (Закон повторного логарифма).

11.6 Закон арксинуса

Теорема 11.13 (Закон арксинуса).

11.7 Правило трёх сигм

Теорема 11.14 (Правило трёх сигм).

11.8 Центральная предельная теорема

Теорема 11.15 (Центральная предельная теорема).

Теорема 11.16 (Оценка Берри-Эссена).

12 Свёртки случайных величин

d

13 Указатель терминов

d

14 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.