Московский физико-технический институт
Физтех-школа прикладной математики и информатики

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

IV СЕМЕСТР

Лектор:

h\nu

Автор: *Киселев Николай*
Репозиторий на Github

весна 2025

Содержание

1 Введение

1.1 Предисловие

Ну чето рассказали там про принцип устойчивости частот, про то, что ля-ля-ля тополя

1.2 Напоминание с ОВиТМа

Определение 1.1. $\mathcal{F}$ — алгебра над Ω, если

1. $\Omega \in \mathcal{F}$
2. $A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$
3. $A \in \mathcal{F} \Rightarrow \overline{A} \in \mathcal{F}$

Замечание. Элементы Ω называются элементарными событиями.

Замечание. Элементы $\mathcal{F}$ называются событиями.

Определение 1.2. $\mathcal{F}$ — σ-алгебра над Ω, если

1. $\mathcal{F}$ — алгебра
2. $A_1, \dots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

Определение 1.3. Функция $P : \mathcal{F} \to [0, 1]$ называется вероятностной мерой, если $P(\Omega) = 1, P$ — σ-аддитивна.

Замечание. 1. $P(\varnothing) = 0$

2. $A \subset B \Rightarrow P(A) \leqslant P(B)$ — **монотонность меры**
3. P конечно аддитивна
4. Для P верна формула включений-исключений.
5. $P\left(\bigcup_{i=1}^{\infty} A_i\right) =\leqslant \sum_{i=1}^{\infty} P(A_i)$

Теорема 1.1 (О непрерывности вероятностной меры)**.** *Пусть* $(\Omega, \mathcal{F})$ *таково, что* $\mathcal{F}$ — *алгебра над* Ω, P — *мера и* $P(\Omega) = 1$. *Тогда следующие условия эквивалентны:*

1. P — σ*-аддитивна на* $\mathcal{F}$
2. P *непрерывна в нуле, т.е.* $\bigcap A_i = \varnothing \Rightarrow P(A_i) \to P(A)$
3. P *непрерывна сверху, т.е.* $\bigcap A_i = A \Rightarrow P(A_i) \to P(A)$
4. P *непрерывна снизу, т.е.* $\bigcup A_i = A \Rightarrow P(A_i) \to P(A)$

Определение 1.4. Вероятностное пространство — это измеримое пространство $(\Omega, \mathcal{F}, P)$, т.е. Ω — множество, $\mathcal{F}$ — некая σ-алгебра над Ω, P — вероятностная мера на $(\Omega, \mathcal{F})$.

Определение 1.5. Система $\mathcal{M}$ подмножеств Ω называется π-системой, если $\mathcal{M}$ замкнуто относительно счетного пересечения

Определение 1.6. Система $\mathcal{M}$ подмножеств Ω называется λ-системой, если

1. $\Omega \in \mathcal{M}$
2. $A, B \in \mathcal{M}, A \subset B \Rightarrow B \setminus A \in \mathcal{M}$
3. $A_i \in \mathcal{M} \Rightarrow \bigcup A_i \in \mathcal{M}$

Теорема 1.2 (Первая теорема о λ-системах)**.** *Система $\mathcal{F}$ является σ-алгеброй над Ω тогда и только тогда, когда она является λ-системой и π-системой.*

Утверждение 1.1. *Для любой системы $\mathcal{M}$ подмножеств Ω существует минимальная по включению, содержащаяся в $\mathcal{M}$*

Замечание. $\sigma(\mathcal{M}), \alpha(\mathcal{M}), \pi(\mathcal{M}), \lambda(\mathcal{M})$ — порожденные (минимальные) σ-алгебра, алгебра, π-система, λ-система соответственно.

Теорема 1.3 (Вторая теорема о λ-системах)**.** *Если $\mathcal{M}$ — π-система на Ω, то $\sigma(\mathcal{M}) = \lambda(\mathcal{M})$*

Доказательство. См. доказательство из курса ОВиТМа □

Пример (Борелевская σ-алгебра)**.** $\mathcal{B}(\mathbb{R})$ — σ-алгебра, порожденная всеми открытыми множествами (или, что равносильно, всеми полуинтервалами)

Пример (Борелевская σ-алгебра в $\mathbb{R}^n$)**.** $\mathcal{B}(\mathbb{R}^n)$ — σ-алгебра, порожденная всеми открытыми множествами (или, что равносильно, всеми кубами, где куб — декартово произведение борелевских множеств).

Пример. $\mathcal{B}(\mathbb{R}^\infty)$ — σ-алгебра, порожденная всеми цилиндрами. Цилиндр — множество $x \in \mathbb{R}^\infty : (x_1, \ldots x_n) \in B, B \in \mathcal{B}(\mathbb{R}^n)$

Пример (Борелевская σ-алгебра в общем случае)**.** Пусть (S, ρ) — метрическое пространство, тогда $\mathcal{B}(S)$ — σ-алгебра, порожденная всеми открытыми множествами.

2 Вероятностная мера на прямой

Пусть P — вероятностная мера на $\mathbb{R}, \mathcal{B}(\mathbb{R})$

Определение 2.1. Функцией распределения называется функция $F(x) = P((-\infty, x]), x \in \mathbb{R}$.

Лемма 2.1 (О свойствах функции распределения)**.**

1. *F не убывает*
2. *$\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to+\infty} F(x) = 1$*
3. *F непрерывна справа*

Доказательство.

1. $x \leqslant y \Rightarrow (-\infty, x] \subset (-\infty, y] \Rightarrow F(x) \leqslant F(y)$

2. $x_n \to -\infty \Rightarrow (-\infty, x_n] \to \varnothing \Rightarrow \lim_{n\to\infty} F(x_n) = 0 = \lim_{x\to+\infty} F(x)$, $x_n \to +\infty \Rightarrow (-\infty, x_n] \to \mathbb{R} \Rightarrow \lim_{n\to\infty} F(x_n) = 1 = \lim_{x\to+\infty} F(x)$
3. Если $x_n \searrow x$, то $(-\infty, x_n] \searrow (-\infty, x]$ и $F(x_n) \to F(x)$

$\square$

Теорема 2.1 (О взаимно-однозначном соответствии функции распределения и вероятностной меры)**.** *Пусть F удовлетворяет условиям 1-3 из предыдущей теоремы. Тогда существует единственная вероятностная мера P на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, т.ч. F — её функция распределения.*

2.1 Классификация вероятностных мер

Далее мы будем отождествлять понятия распределения и вероятностной меры.

2.1.1 Дискретные распределения

Определение 2.2. Вероятностная мера P на $(R, \mathcal{B}(R))$ называется дискретной, если $\exists X$ — не более. чем счетое множество на $\mathbb{R}$, такое, что $P(\mathbb{R} \setminus X)$ и $\forall x \in X P(\{x\}) > 0$.

Примеры

1. Распределение Бернулли: $P \sim Bern(p)$, если $X = \{0, 1\}, P(\{1\}) = p, P(\{0\}) = 1 - p$
2. Равномерное распределение: $X = 1, 2, \dots n$, $P(\{i\}) = \frac{1}{n}$
3. Биномиальное распределение: $P \sim Bin(n, p)$, если $X = \mathbb{N}, P(\{k\}) = C_n^k p^k (1-p)^{n-k}, P(\{0\})$, моделирует количество успехов среди n испытаний.
4. Пуассоновское распределение: $P \sim Pois(n, p)$, если $X = \mathbb{N}, P(\{k\}) = \frac{\lambda^k}{k!} e^{-\lambda}, P(\{0\})$, моделирует редкие события
5. Геометрическое распределение: $P \sim Geom(p)$, если $X = \mathbb{N}, P(\{k\}) = p(1-p)^{k-1}, P(\{0\})$, моделирует первый момент удачи в бесконечной схеме испытаний Бернулли

2.1.2 Абсолютно непрерывные распределения

Определение 2.3. $F(x)$ называется абсолютно непрерывной, если $\exists p(t) \geqslant 0 : \forall x \in \mathbb{R} F(x) = \int_{-\infty}^{x} p(t)dt$. В таком случае мы говорим, что $p(t)$ является плотностью функции F или соответствующего распределения (вероятностной меры).

Замечание. В таком случае $F'(x) = p(x)$ почти всюду по мере Лебега.

Примеры

1. Равномерное распределение на отрезке $[a, b]$ — $U(a, b)$:

$$p(x) = \begin{cases} \frac{1}{b-a}, a \leqslant x \leqslant b \\ 0, \text{иначе} \end{cases}, F(x) = \begin{cases} 0, x < a \\ \frac{x-a}{b-a}, a \leqslant x \leqslant b \\ 1, x > b \end{cases}$$

   Моделирует случайную точку из отрезка $[a, b]$

2. Нормальное распределение — $N(a, \sigma^2)$:

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

   Моделирует измерение с ошибкой

3. Экспоненциальное распределение — $N(a, \sigma^2)$:

$$p(x) = \lambda e^{\lambda x} I\{x > 0\}$$

$$F(x) = \begin{cases} 0, x < 0 \\ 1 - e^{-\lambda x}, x \geqslant 0 \end{cases}$$

   Моделирует время ожидание

4. Гамма-распределение: $\Gamma(\lambda, \alpha), \lambda, \alpha > 0$.

$$p(x) = \frac{x^{\alpha-1}\lambda^\alpha}{\Gamma(\alpha)} e^{-\lambda x} I\{x > 0\}$$

   Где:

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx$$

   Нам в дальнейшем потребуются различные свойства Γ-функции: $\Gamma(n) = (n-1)!, \Gamma(\alpha+1) = \alpha\Gamma(\alpha)$

5. Распределение Коши $K(\sigma), \sigma > 0$

$$p(x) = \frac{\sigma}{\pi(x^2 + \sigma^2)}$$

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan \frac{x}{\sigma}$$

   Модель

2.1.3 Сингулярные распределения

Определение 2.4. Точка x называется точкой роста фукнции распределения $F(x)$, если $\forall \varepsilon > 0 : F(x + \varepsilon) - F(x - \varepsilon) > 0$

Определение 2.5. Функция распределения $F(x)$ наызвается сингулярной, если она непрерывна, и множество точек ее роста имеет $\mu = 0$.

Примеры

1. Канторова лестница — ее точками роста являвется канторово множество.

Теорема 2.2 (Лебега о разложении)**.** *Если $F(x)$ — функция распределения на прямой, тогда $\exists!$ разложение вида: $F(x) = \alpha_1 F_1(x) + \alpha_2 F_2(x) + \alpha_3 F_3(x)$, где $\alpha_i \geqslant 0, \alpha_1 + \alpha_2 + \alpha_3 = 1$ и F_1 — дискретная, F_2 — абсолютно непрерывная, F_3 — сингулярная.*

3 Вероятностные меры в $\mathbb{R}^n$

Пусть P — вероятностная мера на $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

Определение 3.1. Функцией распределения P называется $F(x_1, \dots x_n) = P((-\infty, x_1] \times \dots \times (-\infty, x_n])$

Замечание (Обозначения)**.**

1. $\vec{x} = (x_1, \dots x_n)$
2. $\vec{x} \geqslant \vec{y} \Leftrightarrow x_i \geqslant y_i$
3. $(-\infty, \vec{x}] = (-\infty, x_1] \times \dots \times (-\infty, x_n]$
4. $\vec{x}_n \downarrow \vec{x}$ если $\vec{x} = \lim_{n\to\infty} \vec{x}_n$ и $\vec{x}_n \geqslant \vec{x}_{n+1}$.

Определение 3.2. Для $i = 1, \dots n, a_i < b_i$ введем:

$$\Delta^i_{a_i,b_i} F(x_1, \dots x_n) = F(x_1, x_{i-1}, b_i, x_{i+1}, \dots x_n) - F(x_1, x_{i-1}, a_i, x_{i+1}, \dots x_n)$$

Лемма 3.1 (Свойства многомерных функций распределения)**.**

1. *Если $\vec{x}_n \downarrow \vec{x}$, то $F(\vec{x}_n) \to F(\vec{x})$*
2. $\lim\limits_{\substack{x_1 \to \infty \\ \vdots \\ x_n \to \infty}} F(x_1, \dots x_n) = 1$
3. $\forall i = 1, \dots n : \lim_{x_i \to -\infty} F(x_1, \dots x_n) = 0$.
4. *Для любых $a_i < b_i, i = 1, \dots n$:*

$$\Delta^1_{a_1,b_1} \circ \Delta^2_{a_2,b_2} \circ \dots \circ \Delta^n_{a_n,b_n} F(x_1, \dots x_n) \geqslant 0$$

Доказательство.

1. Если $\vec{x}_n \downarrow \vec{x}$, то $(-\infty, \vec{x}_n] \downarrow (-\infty, \vec{x}] \Rightarrow$ в силу непрерывности вероятностной меры, получаем $F(\vec{x}_n) \to F(\vec{x})$
2. Если $\vec{x}_n \uparrow (+\infty, \dots +\infty)$, то $(-\infty, \vec{x}_n] \uparrow \mathbb{R}^n \Rightarrow$ в силу непрерывности вероятностной меры, получаем $F(\vec{x}_n) \to P(\mathbb{R}^n) = 1$
3. Если $x_i \downarrow -\infty$, то $(-\infty, \vec{x}_n] \downarrow \varnothing \Rightarrow$ в силу непрерывности вероятностной меры, получаем $F(\vec{x}_n) \to P(\varnothing) = 0$

4. Для $n = 2$:

$$\Delta^1_{a_1,b_1}\circ\Delta^2_{a_2,b_2}F(x_1,x_2) = F(b_1,b_2)-F(b_1,a_2)-F(a_1,b_2)+F(a_1,a_2) = P((a_1,b_1]\times(a_2,b_2]) \geqslant 0$$

В общем случае:

$$\Delta^i_{a_i,b_i}P(A_1\times\dots\times A_{i-1}\times(-\infty,x_i)\times A_{i+1}\times\dots\times A_n) =$$

$$= P(A_1\times\dots\times A_{i-1}\times(a_i,b_i]\times A_{i+1}\times\dots\times A_n)$$

Получаем, что

$$\Delta^1_{a_1,b_1}\circ\Delta^2_{a_2,b_2}\circ\dots\circ\Delta^n_{a_n,b_n}F(x_1,\dots x_n) = P((a_1,b_1]\times\dots\times(a_n,b_n]) \geqslant 0$$

□

Теорема 3.1 (О взаимно однозначном соответствии)**.** *Если F удовлетворяет свойствами 1-3 из леммы, то $\exists!$ вероятностная мера P на $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$, такая, что F — ее функция распределения*

Доказательство. См. ОВиТМ □

Замечание. Свойство 3 нельзя заменить на неубывание по каждой из переменнх.

Доказательство. Рассмотрим $F(x,y) = \begin{cases} 1, x+y\geqslant 0 \\ 0, x+y<0\end{cases}$. Заметим, что F удовлетворяет свойствам 1, 2 и не убывает по обеим переменным. При этом, если мы возьем:

$$\Delta^x_{-1,1}\circ\Delta^y_{-1,1}F(x,y) = F(1,1)-F(-1,1)-F(1,-1)+F(-1,-1) = 1-2+0<0$$

Получаем, что F — не двумерная функция распределения. □

3.1 Примеры

1. Пусть $F_1, F_2, \dots F_n$ — одномерные функции распределения. Рассмотрим $F(x_1,\dots x_n) = F_1(x_1)\dots F_n(x_n)$ — многомерная функция распределения. Свойства 1, 2 очевидны, проверим 3:

$$\Delta^1_{a_1,b_1}\circ\Delta^2_{a_2,b_2}\circ\dots\circ\Delta^n_{a_n,b_n}F(x_1,\dots x_n) = \prod_{k=1}^{n}(F_k(b_k)-F_k(a_k)) \geqslant 0$$

2. Пусть $p(t_1,\dots t_n)\geqslant 0$, т.ч.

$$\int_{\mathbb{R}^n} p(t_1,\dots t_n)dt_1dt_2\dots dt_n = 1$$

Тогда:

$$(*)\quad F(x_1,\dots x_n) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n}p(t_1,\dots t_n)dt_1dt_2\dots dt_n$$

Свойства 1, 2 очевидны, проверим 3:

$$\Delta^1_{a_1,b_1} \circ \Delta^2_{a_2,b_2} \circ \dots \circ \Delta^n_{a_n,b_n} F(x_1, \dots x_n) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \dots \int_{a_n}^{b_n} p(t_1, \dots t_n) dt_1 dt_2 \dots dt_n$$

Определение 3.3. Если имеет место представление $(*)$, то $p(t_1, \dots t_n)$ называется плотностью функции распределения F.

4 Вероятностные меры в $\mathbb{R}^\infty$

Определение 4.1. Пусть P — вероятностная мера на $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$. Рассмотрим для $n \in \mathbb{N}$ вероятностную меру P_n на $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, т.ч.

$$P_n(B) = P(Cyl(n, B)), B \in \mathcal{B}(\mathbb{R}^n)$$

Тогда можно заметить, что $P_{n+1}(B \times \mathbb{R}) = P_n(B)$.

Определение 4.2. Свойство выше называется согласованностью для последовательности вероятностных мер $\{P_n\}$

Теорема 4.1 (Колмогорова, о мерах в $\mathbb{R}^\infty$)**.** *Пусть $\{P_n, n \in \mathbb{N}\}$ — последовательность согласованных вероятностных мер, P_n — мера на $\mathbb{R}^n$. Тогда $\exists!$ вероятностная мера P на $(\mathbb{R}^\infty, \mathcal{B}(R^\infty))$, такая, что $\forall n \in \mathbb{N} : \forall B_n \in \mathcal{B}(\mathbb{R}^n)$:*

$$(*) \quad P_n(B_n) = P(Cyl(n, B_n))$$

Доказательство. Зададим меру P на цилиндрах по правилу $(*)$. Цилиндры образуют алгебру $\mathcal{A}$. Проверим корректность задания P. Если $Cyl(n, B_n) = Cyl(n + k, B_{n+k})$, то $B_{n+k} = B_n \times \mathbb{R}^k$. Тогда в силу согласованности:

$$P_n(B_n) = P_{n+k}(B_{n+k})$$

Проверим, что P — конечно аддитивна на $\mathcal{A}$. Если $\tilde{B}_1, \dots \tilde{B_N} \in \mathcal{A}$ — непересекаются, то будем считать, что $\exists n : \tilde{B}_i = Cyl(n, B_i), i = 1, \dots N, B_i \in \mathcal{B}(\mathbb{R}^n)$. Тогда:

$$P\left(\bigsqcup_{i=1}^N \tilde{B}_i\right) = P\left(Cyl\left(n, \bigsqcup_{i=1}^N \tilde{B}_i\right)\right) = P_n\left(\bigsqcup_{i=1}^N B_i\right) = \sum_{i=1}^N P_n(B_i) = \sum_{i=1}^N P(\tilde{B}_i)$$

Проверим, что P непрерывна в нуле (на $\mathcal{A}$). Пусть $\tilde{B_n} \downarrow \varnothing, \tilde{B_n} \in \mathcal{A}$. Без ограничения общности, считаем, что $\tilde{B_n} = Cyl(n, B_n)$.

От противного. Пусть $\lim_{n\to\infty} P(\tilde{B_n}) = \delta > 0$. Тогда для $\forall n \in \mathbb{N}$ выберем компактные $A_n \subset \mathbb{R}^n$, такие, что $A_n \subset B_n$ и $P(B_n \setminus A_n) \leqslant \frac{\delta}{2^{n+1}}$. Обозначим $\tilde{A}_n = Cyl(n, A_n)$. Тогда: $P(\tilde{B_n} \setminus \tilde{A}_n) = P_n(B_n \setminus A_n) \leqslant \frac{\delta}{2^{n+1}}$. Введем $\tilde{C}_n = \bigcap_{i=1}^n \tilde{A}_i$. Тогда $\tilde{C}_n \downarrow \varnothing, \tilde{C}_n = Cyl(n, C_n)$, где $C_n = A_n \cap (A_{n-1} \times \mathbb{R}) \cap (A_{n-2} \times \mathbb{R}^2) \cap \dots \cap (A_1 \times \mathbb{R}^{n-1})$ — тоже компакт в $\mathbb{R}^n$. Далее:

$$P(\tilde{B_n} \backslash \tilde{C_n}) \leqslant \sum_{i=1}^n P(\tilde{B_n} \backslash \tilde{A_i}) \leqslant |B_i \subset B_n, i \geqslant n| \leqslant \sum_{i=1}^n P(\tilde{B_i} \backslash \tilde{A_i}) \leqslant \frac{\delta}{2} \Rightarrow \lim_{n\to\infty} P(\tilde{C_n}) \geqslant \frac{delta}{2} > 0$$

Возьмем в каждом $\tilde{C}_n$ по точке $(x_1^{(n)}, x_2^{(n)}, \dots) \in \tilde{C}_n$. Тогда $(x_1^{(n)}, \dots x_n^{(n)}) \in C_n$. Рассмотрим последовательность $x_1^{(n)}$. Эти все точки лежат в C_1. Выберем подпоследовательность n_{k_i} $x_1^{(n_{k_i})} \to x_1^0 \in C_1$. Для $k \in \mathbb{N}$ построим последовательности $(x_1^{(n_{k_i})}, x_2^{(n_{k_i})}, \dots x_k^{(n_{k_i})})$, которая сходится к $x_k^{(0)}, \dots x_k^{(0)}$ в $_k$. Возьмем диагональ, т.е положим $m_k = n_{k_k}$. Тогда $\forall s \in \mathbb{N} : (x_1^{(m_k)}, \dots x_s^{(m_k)}) \to (x_1^{(0)}, \dots x_s^{(0)}) = x^{(0)}$. Получаем, что P счетно-аддитивна на $\mathcal{A}$. В таком случае по теореме о продолжении меры, P единственным образом продолжается до вероятностной меры на $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R}^\infty)$ □

5 Случайные величины и векторы

Пусть $(\Omega, \mathcal{F}, P)$ — вероятностное пространство, $(E, \mathcal{E})$ — измеримое пространство

Определение 5.1. Отображение $X : \Omega \to E$ называется случайным элементом, если X — измеримо, т.е. $\forall B \in \mathcal{E} : X^{-1}(B) = \omega : X(\omega) \in B \in \mathcal{F}$

Определение 5.2. Если $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, то X называется случайной величиной.

Определение 5.3. Если $(E, \mathcal{E}) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, то X называется случайным вектором.

Напоминание (Критерий измеримости отображения)**.** Если $\mathcal{M} \subset \mathcal{E}$ и $\sigma(\mathcal{M}) = \mathcal{E}$, то X — случайный элемент $\Leftrightarrow \forall B \in \mathcal{M} X^{-1}(B) \in \mathcal{F}$

Следствие (Эквивалентные определения случайных величин и векторов)**.**

1. $\xi : \Omega \Rightarrow \mathbb{R}$ — случайная величина $\Leftrightarrow \{\xi \leqslant x\} \in \mathcal{F} \Leftrightarrow \{\xi < x\} \in \mathcal{F}$
2. $\xi = (\xi_1, \dots, \xi_n)$ — случайный вектор $\Leftrightarrow \forall i : \xi_i$ измеримо

Доказательство.

1. Очевидно следует из критерия, т.к. лучи (замкнутые или открытые) порождают $\mathcal{B}(\mathbb{R})$.
2. В одну сторону. возьмем $B_i \in \mathcal{B}(\mathbb{R})$. Т.к. $\underbrace{\mathbb{R} \times \dots \times \mathbb{R} \times}_{i-1} B_i \times \dots \times \mathbb{R} = B \in \mathcal{B}(\mathbb{R}^n)$, то $\{\xi_i \in B_i\} = \{\xi \in B\} \in \mathcal{F}$. В другую: $\{\xi \in B_1 \times \dots \times B_n\} = \bigcap_{i=1}^n \{\xi_i \in B_i\}$. Далее, по критерию измеримости, следствие верно (т.к. полуалгебра кубов порождает $\mathcal{B}(\mathbb{R}^n)$).

□

5.1 Действия со случайными величинами

1. Если $f : \mathbb{R}^n \to \mathbb{R}^l$ — борелевская, ξ — случайная, то $f(\xi)$ — случайная.
2. Арифметические операции
3. Пусть ξ_n — последовательность случайных величин, тогда $\sup_n \xi_n, \inf_n \xi_n, \overline{\lim} \xi_n, \underline{\lim} \xi_n$ — тоже случайные величины

5.2 Характеристики случайных величин и векторов

Определение 5.4. Распределением случайной величины ξ на $(\Omega, \mathcal{F}, P)$ называется вероятностная мера P_ξ на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, определяемая по правилу $\forall B \in \mathcal{B}(\mathbb{R}) : P_\xi(B) = P(\xi \in B)$

Определение 5.5. Распределением случайного вектора ξ на $(\Omega, \mathcal{F}, P)$ называется вероятностная мера P_ξ на $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, определяемая по правилу $\forall B \in \mathcal{B}(\mathbb{R}^n) : P_\xi(B) = P(\xi \in B)$

Упражнение. *P_ξ действительно является вероятностной мерой.*

Определение 5.6. Функция распределения случайной величины ξ — это функция $F_\xi(x) = P_\xi((-\infty, x]) = P(\xi \leqslant x)$.

Определение 5.7. Функция распределения случайного вектора $\vec{\xi}$ — это функция $F_{\vec{\xi}} = P_{\vec{\xi}}((-\infty, \vec{x}]) = P(\vec{\xi} \leqslant \vec{x}) = P(\xi_1 \leqslant x_1, \xi_2 \leqslant x_2, \ldots \xi_n \leqslant x_n)$.

Замечание. Обозначение $\{\xi \leqslant x, \eta \leqslant y\}$ обозначает $\{\xi \leqslant x\} \cap \{\eta \leqslant y\}$.

Замечание. Случайные величины и векторы наследуют классификацию распределений.

Определение 5.8. σ-алгебра, порожденная случайной величиной ξ называется $\mathcal{F}_\xi = \{\{\xi \in B\} : B \in \mathcal{B}(\mathbb{R})\}$.

Определение 5.9. σ-алгебра, порожденная случайным вектором ξ называется $\mathcal{F}_\xi = \{\{\xi \in B\} : B \in \mathcal{B}(\mathbb{R}^n)\}$.

Упражнение. *Проверить, что $\mathcal{F}_\xi$ действительно является σ-алгебой.*

Определение 5.10. Пусть ξ — случайный вектор на $(\Omega, \mathcal{F}, P)$. Пусть $\mathcal{C} \subset \mathcal{F}$ — некая σ-алгебра. Тогда ξ является $\mathcal{C}$-измеримой, если $\mathcal{F}_\xi \subset \mathcal{C}$.

Возникает вопрос: а что, если $\mathcal{C} = \mathcal{F}_\eta$?

Лемма 5.1. *Случайная величина ξ является $\mathcal{F}_\eta$-измеримой $\Leftrightarrow \exists$ борелевская функция f, такая, что $\xi = f(\eta)$.*

Доказательство.

- $\Leftarrow$ Пусть $B \in \mathcal{B}(\mathbb{R})$. Тогда $\{\xi \in B\} = \{f(\eta) \in B\} = \{\eta \in \underbrace{f^{-1}(B)}_{\in \mathcal{B}(\mathbb{R})}\} \in \mathcal{F}_\eta$.

- $\Rightarrow$ Пусть $\xi = I_A, A \in \mathcal{F}_\eta$. Тогда $\exists B \in \mathcal{B}(\mathbb{R}^n)$, т.ч. $A = \{\eta \in B\}$. Тогда $\xi = f(\eta), f(x) = I_B$. Следовательно, все простые $\mathcal{F}_\eta$-измеримые случайные величины тоже являются борелевскими функциями от η. Если ξ — произвольная случайная величина, то рассмотрим последовательность последовательность величин $\{\xi_n\}$, такую, что $\forall \omega \in \Omega :$ $\xi_n(\omega) \to \xi(\omega)$, где ξ_m — простые случайные величины и $\forall m \in \mathbb{N} : \xi_m = g_m(\xi)$, где g — борелевская. Далее, $\mathcal{F}_{\xi_m} \subset \mathcal{F}_\xi \subset \mathcal{F}_\eta \Rightarrow \exists$ борелевская функция f_m, т.ч. $\xi_m = f_m(\eta)$. Введем $B = \{x \in \mathbb{R}^n : \exists \lim_{m\to\infty} f_m(x)\}$. Положим $f(x) = \begin{cases} \lim_{m\to\infty} f_m(x), x \in B \\ 0, x \notin B. \end{cases}$.
Заметим, что $B \in \mathcal{B}(\mathbb{R}^n) \Rightarrow f$ — тоже борелевская функция. По построению $\forall \omega \in \Omega :$ $\xi(\omega) = \lim_{m\to\infty} \xi_m(\omega) = \lim_{m\to\infty} f_m(\eta(\omega)) = f(\eta)$.

$\square$

6 Математическое ожидание

Пусть ξ — случайная величина на $(\Omega, \mathcal{F}, P)$

Определение 6.1. Математическое ожидание случайной величины ξ называется интеграл Лебега:

$$\mathbb{E}\xi = \int_\Omega \xi(\omega)P(d\omega) = \int_\Omega \xi dP$$

Напоминание. Напомним, как мы строили интеграл Лебега:

1. Для простой функции как сумму
2. Для неотрицательной функции как предел
3. Для функции разных знаков, разбиваем $\xi = \xi^+ - \xi^-$

Далее говорим следующее:

1. Если $\mathbb{E}\xi^\pm$ конечно, то $\mathbb{E}\xi = \mathbb{E}\xi^+ - \mathbb{E}\xi^-$.
2. Если одно из $\mathbb{E}\xi^\pm$ равно $+\infty$, то $\mathbb{E}\xi = \pm\infty$ (ставим соответствующий знак).
3. Если $\mathbb{E}\xi^+ = \mathbb{E}\xi^- = +\infty$, то говорим, что $\mathbb{E}\xi$ не определено.

6.1 Свойства матожидания

Из свойств интеграла:

1. Линейность: $\mathbb{E}(\alpha\xi + \beta\eta) = \alpha\mathbb{E}\xi + \beta\mathbb{E}\eta$
2. Сохраняет отношение порядка: $\xi \leqslant \eta \Rightarrow \mathbb{E}\xi \leqslant \mathbb{E}\eta$.
3. $|\mathbb{E}\xi| \leqslant \mathbb{E}|\xi|$

Определение 6.2. Событие $A \in \mathcal{F}$ происходит почти наверное, если $P(A) = 1$. Мы будем писать это как "A п.н.".

И еще свойства математического ожидания:

4. Если $\xi \geqslant 0, \mathbb{E}\xi = 0$, то $\xi = 0$ п.н.
5. Если $\xi = \eta$ п.н., то $\mathbb{E}\xi = \mathbb{E}\eta$.
6. Если $\forall A \in \mathcal{F} : \mathbb{E}(\xi \cdot I_A) \leqslant \mathbb{E}(\eta \cdot I_A) \Rightarrow \xi \leqslant \eta$ п.н.

Теорема 6.1 (Замена переменных в интеграле Лебега)**.** *Пусть ξ — случайный вектор из $\mathbb{R}^n$ на $(\Omega, \mathcal{F}, P)$. Пусть $f : \mathbb{R}^n \to \mathbb{R}$ — борелевская функция. Тогда $\mathbb{E}f(\xi) = \int_{\mathbb{R}^n} f(x)P_\xi(dx)$.*

Доказательство. Пусть сначала $f(x) = I_B(x), B \in \mathcal{B}(\mathbb{R}^n)$. Тогда

$$\mathbb{E}f(\xi) = E \cdot I\{\xi \in B\} = P(\xi \in B) = P_\xi(B) = \int_B P\xi(dx) = \int_{\mathbb{R}^n} I_B(x)P_\xi(dx) = \int_{\mathbb{R}^n} f(x)P_\xi(dx)$$

Заключаем, что теорема верна для простых функцйи в силу линейности матожидания и интеграла Лебега. Далее осуществляем приближение простыми функциями в общем случае. $\square$

Замечание. Матожидание функций от случайной величины или случайного вектора зависит только от распределения.

Определение 6.3. Случайные величины (векторы) ξ, η наызваются одинаково распределенными, если их распределения равны, т.е. $P_\xi = P\eta$. В таком случае пишут $\xi \overset{d}{=} \eta$.

Следствие. Если $\xi \overset{d}{=} \eta$, то $\forall f : \mathbb{E}f(\xi) = \mathbb{E}f(\eta)$

Следствие. Если ξ — случайная величина, то $\mathbb{E}\xi = \int_{\mathbb{R}} xP_\xi(dx)$.

Доказательство. Подставляем $f(x) = x$ □

6.2 Обобщенная плотность

Но как посчитать последний интеграл? Здесь нам поможет обобщенная плотность. Пусть P — вероятностная мера на $(\Omega, \mathcal{F}), \mu$ — σ-конечная мера на $(\Omega, \mathcal{F})$.

Определение 6.4. Мера P имеет обобщенную плотность по мере $p(x)$ по мере μ, если

$$\forall B \in \mathcal{F} : P(B) = \int_B p(x)\mu(dx)$$

Смысл данного определения состоит в том, что если P — абсолютно непрерывна относительно μ, то $p = \frac{dP}{d\mu}$ — производная Радона-Никодима.

6.3 Примеры

1. μ — мера Лебега на $\mathbb{R} \Rightarrow$ для абсолютно-непрерывного распределения обобщенная плотность совпадает с обычной плотностью

2. μ — считающая мера на $\mathbb{Z}$ (дискретное распределение на $\mathbb{Z}$). Тогда обобщенная плотность вычилсяется по формуле $p(x) = P(\{x\})$

Теорема 6.2 (О вычислении через обобщенную плотность). *Пусть P — вероятностная мера на $(\mathbb{R}^n, \mathbb{R}^\setminus)$, имеющая плотность $p(x)$ по мере μ. Тогда $\forall$ борелевской функции $f : \mathbb{R}^n \to \mathbb{R}$ выполнено:*

$$\int_{\mathbb{R}^n} f(x)P(dx) = \int_{\mathbb{R}^n} f(x)p(x)dx$$

Доказательство. Пусть $f(x) = I_B(x), B \in \mathcal{B}(\mathbb{R}^n)$. Тогда:

$$\int_{\mathbb{R}^n} f(x)P(dx) = \int_B P(dx) = P(B) = \int_B p(x)\mu(dx) = \int_{\mathbb{R}^n} f(x)p(x)\mu(dx)$$

Т.к. обе части требуемого линейны по f, получаем, что равенство верно для простых функций. Для произвольной f осуществляем приближение простыми по стандартной схеме. □

Следствие. Пусть ξ – абсолютно непрерывный случайный вектор. с плотностью $p(x)$. Тогда $\forall f$ — борелевской:

$$\mathbb{E}f(\xi) = \int_{\mathbb{R}^n} f(x)p(x)dx$$

Следствие. Пусть ξ — дискретная случайная величина, сосредоточенная на некотором множестве $\mathcal{X}$. Тогда $\forall f$ — борелевской:

$$\mathbb{E}f(\xi) = \sum_{x\in\mathcal{X}}(f(x)P(\xi = x))$$

Доказательство.

$$\mathbb{E}f(\xi) = \int_{\mathbb{R}} f(x)P_\xi(dx) = \sum_{x\in\mathcal{X}} f(x) \underbrace{P(\xi = x)}_{\text{обобщ. плотность}} \underbrace{\mu(\{x\})}_{\text{считающая}} = \sum_{x\in\mathcal{X}} f(x)P(\xi = x)$$

$\square$

Пример. Пусть ξ имеет следующую функцию распределения:

$$F(x) = \begin{cases} 0, x < 0 \\ 1 - e^{-x}, x \in [0, 1) \\ 0, \text{иначе} \end{cases}$$

Требуется найти $\mathbb{E}\xi$

Решение.

$$\mathbb{E}\xi = 1\cdot P\left(\xi = \frac{1}{e}\right) + \int_0^1 xF'(x)dx = \frac{1}{e} + \int_0^1 xe^{-x}dx = \frac{1}{e} + (-xe^{-x} - e^{-x})|_0^1 = 1 - \frac{2}{e} + \frac{1}{e} = 1 - \frac{1}{e}$$

7 Независимость

7.1 Независимость событий

Пусть $(\Omega, \mathcal{F}, P)$ — вероятностное пространство.

Определение 7.1. События A, B независимы в $(\Omega, \mathcal{F}, P)$, если $P(A \cap B) = P(A)P(B)$.

Определение 7.2. События $A_1, \dots A_n$ называются попарно независимыми, если $\forall i \neq j$: A_i, A_j независимы.

Определение 7.3. События $A_1, \dots A_n$ называются независимыми в совокупности, если $\forall k \leqslant n \forall i_1 < \dots < i_k : P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \dots P(A_{i_k})$.

Далее, под независимостью будем подразумевать независимость в совокупности.

Определение 7.4. Системы событий $\mathcal{M}_1, \dots \mathcal{M}_n$ независимы в совокупности, если $\forall A_i \in \mathcal{M}_i$ события $\{A_i\}$ независимы в совокупности.

Теорема 7.1 (критерий независимости σ-алгебр)**.** *Пусть $\mathcal{M}_1, \dots \mathcal{M}_n$ — π- системы событий на $(\Omega, \mathcal{F}, P)$. Тогда $\mathcal{M}_1, \dots \mathcal{M}_n$ независимы $\Leftrightarrow \sigma(\mathcal{M}_1), \dots \sigma(\mathcal{M}_2)$*

Доказательство.

$\Leftarrow$ очевидно из определения

$\Rightarrow$ Докажем для $n = 2$ (для других аналогично). Рассмотрим

$$\mathcal{L}_1 = \{A \in \mathcal{F} : A \text{ независимо с } \mathcal{M}_2\}$$

По условию, $\mathcal{M}_1 \subset \mathcal{L}_1$. Проверим, что $\mathcal{L}_1$ — λ-система.

(a) $\Omega \in \mathcal{L}_1$, т.к. оно независимо со всеми событиями.

(b) Пусть $A, B \in \mathcal{L}_1, A \subset B$ докажем, что $B \setminus A \in \mathcal{L}_1$. Пусть $C \in \mathcal{M}_2$. Тогда:

$$P((B \setminus A) \cap C) = P((B \cap C) \setminus (A \cap C)) = P(B \cap C) - P(A \cap C) = (*)$$

Из того, что $B, A \in \mathcal{L}_1$ получаем, что они независимы с C. Тогда:

$$(*) = P(B)P(C) - P(A)P(C) = (P(B) - P(A))P(C) = P(B \setminus A)P(C)$$

(c) Пусть $\{A_n, n \in \mathbb{N}\}$ таковы, что $A_n \uparrow A$ и $\forall n : A_n \in \mathcal{L}_1$ докажем, что $A \in \mathcal{L}_1$. Пусть $C \in \mathcal{M}_2$. Тогда $A_n \cap C \uparrow A \cap C$. В силу непрерывности вероятностной меры:

$$P(A \cap C) = \lim_{n\to\infty} P(A_n \cap C) = \lim_{n\to\infty} P(A_n)P(C) = P(C) \lim_{n\to\infty} P(A_n) = P(C)P(A)$$

Получили, что A независимо с $\mathcal{M}_2 \Rightarrow A \in \mathcal{L}_1$.

То есть $\mathcal{L}_1$ — это λ-система, $\mathcal{M}_1 \subset \mathcal{L}_1$. По второй теореме о π-λ-системах:

$$\sigma(\mathcal{M}_1) \subset \mathcal{L}_1 \Rightarrow \sigma(\mathcal{M}_1) \text{ независимо с } \mathcal{M}_2$$

Теперь рассмотрим $\mathcal{L}_2 = \{A \in \mathcal{F} : A \text{ независимо с } \sigma(\mathcal{M}_1)\}$. По доказанному $\mathcal{M}_2 \subset \mathcal{L}_{\in\int}$. Далее точно так же проверяем, что $\mathcal{L}_2$ — λ-система.

□

Определение 7.5. Пусть $\{\mathcal{M}_\alpha, \alpha \in \mathcal{A}\}$ — системы событий на $(\Omega, \mathcal{F}, P)$. Этот набор называется независимым, если любой конечный поднабор $M_{\alpha_1}, \dots M_{\alpha_k}, \alpha_i \neq \alpha_j$ независим.

7.2 Независимость случайных величин и случайных векторов

Определение 7.6. Случайные векторы $\xi_1, \dots \xi_n$ называются независимыми в совокупности, если порожденные ими σ-алгебры являются независимыми.

Утверждение 7.1. *Случайные векторы* $\xi_1, \dots \xi_n, \xi_i \in \mathbb{R}^{k_i}$ *независимы в совокупности* $\Leftrightarrow \forall B_1, \dots B_n, B_i \in \mathcal{B}(\mathbb{R}^{k_i})$ *выполнено*

$$P(\xi_1 \in B_1, \dots \xi_n \in B_n) = \prod_{i=1}^{n} P(\xi_i \in B_i)$$

Доказательство. $\Rightarrow$ очевидно

$\Leftarrow$ Надо проверить, что $\forall B_1, \dots B_n$ события $\{\xi_i \in B_i\}$ независимы в совокупности. Если фиксирован набор $i_1, \dots i_k$, положим $B_j = \mathbb{R}^{k_j}$ для $j \neq i_s$. Тогда получаем:

$$P(\xi_1 \in B_1, \dots \xi_n \in B_n) = \prod_{i=1}^{n} P(\xi_i \in B_i) = \prod_{s=1}^{k} P(\xi_{i_s} \in B_{i_s})$$

□

Определение 7.7. Пусть $\xi_1, \dots \xi_n$ — случайные величины. Совместным распределением $\xi_1, \dots \xi_n$ называется распределение вектора $(\xi_1, \dots \xi_n)$. Соответствующие функции распределения и плотности тоже называются совместными

Утверждение 7.2. *Если случайные величины $\xi_1, \dots \xi_n$ независимы, то их совместное распределение есть прямое произведение распределений $\xi_1, \dots \xi_n$, т.е:*

$$P_{(\xi_1, \dots \xi_n)} = P_{\xi_1} \times \dots \times P_{\xi_n}$$

Доказательство. Для $\forall B_1, \dots B_n \in \mathcal{B}(\mathbb{R})$ имеем:

$$P_{(\xi_1, \dots \xi_n)}(B_1 \times \dots \times B_n) = P((\xi_1, \dots \xi_n) \in B_1 \times \dots \times B_n) =$$

$$P(\xi_1 \in B_1, \dots \xi_n \in B_n) = P(\xi_1 \in B_1) \dots P(\xi_n \in B_n) = P_{\xi_1} \times \dots \times P_{\xi_n}$$

□

Теорема 7.2 (Критерий независимости для функции распределения)**.** *Пусть $\xi_1, \dots \xi_n$ — случайные величины на $(\Omega, \mathcal{F}, P)$. Они независимы в совокупности $\Leftrightarrow \forall x_1, \dots x_n \in \mathbb{R}$ выполнено:*

$$P(\xi_1 \leqslant x_1, \dots \xi_n \leqslant x_n) = \prod_{i=1}^{n} \underbrace{P(\xi_i \leqslant x_i)}_{F_{\xi_i}(x_i)}$$

Доказательство. $\Rightarrow$ следует из определения

$\Leftarrow$ Рассмотрим

$$G(x_1, \dots x_n) = \prod_{i=1}^{n} F_{\xi_i}(x_i) = P(\xi_1 \leqslant x_1, \dots \xi_n \leqslant x_n)$$

Это многомерная функция распределения. Заметим, что ей соответствует

□