Забавные факты по теории вероятностей

Источник: Чернова Н.И., Теория вероятностей

Авторы заметок: Хоружий Кирилл

От: *18 мая 2021 г.*

Содержание

1 Основные понятия теории вероятностей **3**
- 1.1 Элементы комбинаторики . . . 3
- 1.2 События и операции над ними . . . 3
- 1.3 Дискретное пространство элементарных исходов . . . 3
- 1.4 Дискретное пространство элементарных исходов . . . 4
- 1.5 Геометрическая вероятность . . . 5

2 Аксиоматика теории вероятностей **5**
- 2.1 Алгебра и σ-алгебра событий . . . 5
- 2.2 Мера и вероятностная мера . . . 5

3 Условная вероятность и независимость **6**
- 3.1 Условная вероятность . . . 6
- 3.2 ($\mathfrak{B}$) Операции с вероятностями . . . 7
- 3.3 Независимость событий . . . 7
- 3.4 Формула полной вероятности . . . 7
- 3.5 Формула Байеса . . . 7

4 Схема Бернулли **8**
- 4.1 Распределение числа успехов в n испытаниях . . . 8
- 4.2 Номер первого успешного испытания . . . 8
- 4.3 Независимые испытания с несколькими исходами . . . 8
- 4.4 Теорема Пуассона для схемы Бернулли . . . 9

5 Случайные величины и их распределения **9**
- 5.1 Случайные величины . . . 9
- 5.2 Распределения случайных величин . . . 9
- 5.3 Функция распределения . . . 10
- 5.4 ($\mathfrak{Z}$) Примеры дискретных распределений . . . 10
- 5.5 ($\mathfrak{Z}$) Примеры абсолютно непрерывных распределений . . . 11
- 5.6 Свойства функций распределения . . . 12
- 5.7 Свойства нормального распределения . . . 12

6 Преобразования случайных величин **13**
- 6.1 Измеримость функций от случайных величин . . . 13
- 6.2 Распределения функций от случайных величин . . . 13

7 ✗Многомерные распределения **13**
- 7.1 Совместное распределение . . . 13
- 7.2 Типы многомерных распределений . . . 14
- 7.3 Примеры многомерных распределений . . . 14
- 7.4 Независимость случайных величин . . . 14
- 7.5 Функции от двух случайных величин . . . 15

8 Числовые характеристики распределений **15**
8.1 Математическое ожидание случайной величины . . . 15
8.2 Свойства математического ожидания . . . 15
8.3 Дисперсия и моменты старших порядков . . . 15
8.4 Свойства дисперсии . . . 16
8.5 Математические ожидания и дисперсии стандартных распределений . . . 16
8.6 Другие числовые характеристики распределений . . . 16
8.7 Производящие функции . . . 17
8.8 Вычисление моментов через производящие функции . . . 17

9 Числовые характеристики зависимости **18**
9.1 Ковариация двух случайных величин . . . 18
9.2 Коэффициент корреляции . . . 18

10 Характеристические функции **19**
10.1 Определение и примеры . . . 19
10.2 Свойства характеристических функций . . . 19

11 Сходимость последовательностей случайных величин **20**
11.1 Определение и примеры . . . 20

12 Контрольная работа №1 **21**

13 Контрольная работа №2 **24**

Контрольная работа №1 по теории вероятностей

Автор работы: Хоружий Кирилл

От: *18 мая 2021 г.*

1 Основные понятия теории вероятностей

1.1 Элементы комбинаторики

Для начала подружимся с комбинаторикой, взяв некоторую её проекцию на теорвер

Thr 1.1. *Пусть множества $A = \{a_1, \ldots, a_k\}$ состоит из k элементов, а множество $B = \{b_1, \ldots, b_m\}$ – из m элементов. Тогда можно образовать равно $k \cdot m$ пар (a_i, b_j).*

Thr 1.2. *Общее количество различных наборов при выборе k элементов из n **без** возвращения и **с** учётом порядка равняется*

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

где A_n^k называется числом размещений из n элементов по k элементов.

Thr 1.3. *Общее количество различных наборов при выборе k элементов из n **без** возвращения и **без** учета порядка равняется*

$$C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!},$$

где число C_n^k называется числом сочетаний из n элементов по k элементов.

Thr 1.4. *Общее количество различных наборов при выборе k элементов из n с возвращением и без учёта порядка равняется*

$$C_{n+k-1}^k = C_{n+k-1}^{n-1}.$$

1.2 События и операции над ними

Def 1.5. *Пространством элементарных исходов* называют множество Ω, содержащее все возможные взаимоисключающие результаты данного случайного эксперимента. Элементы множества Ω называются *элементарными исходами* и обозначаются ω.

Def 1.6. *Событиями* называются подмножества Ω. Говорят, что *произошло событие* A, если эксперимент завершился одним из элементарных исходов, входящих в множество A.

Вообще в силу таких определений события и множества оказываются очень похожими, так что определены операции *объединения*, *пересечения*, *дополнения*, а также взятия *противоположного* $\bar{A} = \Omega \backslash A$. Также можно выделить достоверное событие Ω и невозможное $\varnothing$.

События A и B называются *несовместными*, если они не могут произойти одновременно: $A \cap B = \varnothing$. События $A_1, \ldots, A_n$ называются *попарно несовместными*, если несовместны любые два из них: $A_i \cap A_j = \varnothing$, $\forall i \neq j$. Говорят, что событие A *влечет* событие B ($A \subseteq B$), если $A \Rightarrow B$.

1.3 Дискретное пространство элементарных исходов

Пространство элементарных исходов назовём дискретным, если множество Ω конечно или счётно: $\Omega = \{\omega_1, .., \omega_n, \ldots\}$.

Def 1.7. Сопоставим каждому элементарному исходу ω_i число $p_i \in [0, 1]$ так, чтобы $\sum p_i = 1$. Вероятностью события A называют число

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

где в случае $A = \varnothing$ считаем $\mathrm{P}(A) = 0$.

Def 1.8 (Классическое определение вероятности)**.** Говорят, что эксперимент описывается *классической вероятностной моделью*, если пространство его элементарных исходов состоит из конечного числа равновозможных исходов. Для любого события верно, что

$$\mathrm{P}(A) = \frac{\operatorname{card} A}{\operatorname{card} \Omega}. \tag{1.1}$$

Эту формулу называют *классическим определением вероятности.*

Тут стоит вспомнить три схемы из модели с урнами: схема выбора с возвращением и с учётом порядка (n^k), выбора без возвращения и с учётом порядка (A_n^k), а также выбора без возвращения и без учёта порядка (C_n^k), описываются классической вероятностной моделью. А вот схема выбора с возвращением и без учёта порядка уже не описывается классической вероятностью.

Пример с гипергеометрическим распределением

Из урны, в которой K белых и $N-K$ чёрных шаров, наудачу и без возвращения вынимают n шаров, где $n \leqslant N$. Термин «наудачу» означает, что появление любого набора из n шаров равновозможно. Найти вероятность того, что будет выбрано k белых и $n-k$ чёрных шаров.

Результат – набор из n шаров. Общее число $\operatorname{card}\Omega = C_N^n$. Пусть A_k – событие, состоящее в том, что в наборе окажется k белых и $n-k$ черных. Есть ровно C_K^k способов выбрать k белых шаров из K, и C_{N-K}^{n-k} способов выбрать $n-k$ черных шаров из $N-K$. Тогда $\operatorname{card} A_k = C_K^k C_{N_K}^{n-k}$,

$$\mathrm{P}(A_k) = \frac{\operatorname{card} A_k}{\operatorname{card}\Omega} = \frac{C_K^k C_{N_K}^{n-k}}{C_N^n}.$$

Этот набор вероятностей называется *гипергеометрическим распределением* вероятностей.

1.4 Дискретное пространство элементарных исходов

Пространство элементарных исходов назовём дискретным, если множество Ω конечно или счётно: $\Omega = \{\omega_1, .., \omega_n, \ldots\}$.

Def 1.9. Сопоставим каждому элементарному исходу ω_i число $p_i \in [0,1]$ так, чтобы $\sum p_i = 1$. Вероятностью события A называют число

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

где в случае $A = \varnothing$ считаем $\mathrm{P}(A) = 0$.

Def 1.10 (Классическое определение вероятности)**.** Говорят, что эксперимент описывается *классической вероятностной моделью*, если пространство его элементарных исходов состоит из конечного числа равновозможных исходов. Для любого события верно, что

$$\mathrm{P}(A) = \frac{\operatorname{card} A}{\operatorname{card} \Omega}. \tag{1.2}$$

Эту формулу называют *классическим определением вероятности.*

Тут стоит вспомнить три схемы из модели с урнами: схема выбора с возвращением и с учётом порядка (n^k), выбора без возвращения и с учётом порядка (A_n^k), а также выбора без возвращения и без учёта порядка (C_n^k), описываются классической вероятностной моделью. А вот схема выбора с возвращением и без учёта порядка уже не описывается классической вероятностью.

Пример с гипергеометрическим распределением

Из урны, в которой K белых и $N-K$ чёрных шаров, наудачу и без возвращения вынимают n шаров, где $n \leqslant N$. Термин «наудачу» означает, что появление любого набора из n шаров равновозможно. Найти вероятность того, что будет выбрано k белых и $n-k$ чёрных шаров.

Результат – набор из n шаров. Общее число $\operatorname{card}\Omega = C_N^n$. Пусть A_k – событие, состоящее в том, что в наборе окажется k белых и $n-k$ черных. Есть ровно C_K^k способов выбрать k белых шаров из K, и C_{N-K}^{n-k} способов выбрать $n-k$ черных шаров из $N-K$. Тогда $\operatorname{card} A_k = C_K^k C_{N_K}^{n-k}$,

$$\mathrm{P}(A_k) = \frac{\operatorname{card} A_k}{\operatorname{card}\Omega} = \frac{C_K^k C_{N_K}^{n-k}}{C_N^n}.$$

Этот набор вероятностей называется *гипергеометрическим распределением* вероятностей.

1.5 Геометрическая вероятность

Def 1.11. Пусть некоторая область $\Omega \subset \mathbb{R}^k$ такая, что $\mu(\Omega)$ конечна. Пусть эксперимент состоит из равновероятного выбора случайной точки в области Ω. *Геометрическое определение вероятности*:

$$\mathrm{P}(A) = \frac{\mu(A)}{\mu(\Omega)}.$$

Если для точки выполнены условия геометрического определения, то говорят, что точка *равномерно распределена* в Ω.

2 Аксиоматика теории вероятностей

2.1 Алгебра и σ-алгебра событий

Def 2.1. Множество $\mathcal{A}$, элементами которого являются некоторые подмножества Ω называют *алгеброй*, если оно удовлетворяет следующим условиям:

A1) $\Omega \in \mathcal{A}$ (алгебра содержит достоверные события);

A2) если $A \in \mathcal{A}$, то $\bar{A} \in \mathcal{A}$ (вместе с любым множеством алгебра содержит противоположное к нему);

A3) если $A \in \mathcal{A}$ и $B \in \mathcal{A}$, то $A \cup B \in \mathcal{A}$ (вместе с любыми двумя множествами алгебра содержит их объединение).

Вообще из A1 и A2 следует, что $\varnothing = \bar{\Omega} \in \mathcal{A}$. Пункт A3 экстраполируется на любой конечный набор. Кстати, объединение можно заменить (в силу закона де Моргана) на пересечение:

$$xy \in \mathcal{A} \quad \Leftrightarrow \quad \overline{xy} \in \mathcal{A} \quad \Leftrightarrow \quad \overline{x} + \overline{y} \in \mathcal{A}.$$

Thr 2.2 (закон де Моргана)**.** *Для множеств x, y верно, что*

$$\overline{x+y} = \overline{x} \cdot \overline{y}, \qquad \overline{xy} = \overline{x} + \overline{y},$$

где $xy = x \cap y$, $x + y = x \cup y$.

В случае счётного пространства элементарных исходов A3 алгебры оказывается недостаточно, так приходим к σ-алгебре:

Def 2.3. Множество $\mathcal{F}$, элементами которого являются некоторые подмножества Ω называется σ-алгеброй, если выполнены следующий условия:

S1) $\Omega \in \mathcal{F}$ (алгебра содержит достоверные события);

S2) если $A \in \mathcal{F}$, то $\bar{A} \in \mathcal{F}$ (вместе с любым множеством алгебра содержит противоположное к нему);

S3) если $\{A_i\} \in \mathcal{F}$, то $\cup_i A_i \in \mathcal{F}$ (вместе с любым *счетным* набором событий σ-алгебра содержит их объединение).

Def 2.4. Минимальной σ-алгеброй, содержащей набор множеств $\mathcal{U}$, называется пересечение всех σ-алгебр, содержащих $\mathcal{U}$.

Def 2.5. Минимальная σ-алгебра, содержащая множество $\mathcal{U}$ всех интервалов на вещественной прямой называется *борелевской сигма-алгеброй* в $\mathbb{R}$ и обозначается $\mathfrak{B}(\mathbb{R})$.

Итак, оказался определен специальный класс $\mathcal{F}$ подмножеств Ω, названный σ-алгеброй событий. Применение счетного числа любых операция к множествам из $\mathcal{F}$ снова дает множество из $\mathcal{F}$. *Событиями* будем называть только множества $A \in \mathcal{F}$.

2.2 Мера и вероятностная мера

Def 2.6. Пусть Ω – некоторое непустое множество $\mathcal{F}$ – σ-алгебра его подмножеств. Функция

$$\mu \colon \mathcal{F} \mapsto \mathbb{R} \cap [0, +\infty) \cup \{+\infty\}$$

называется *мерой* на $(\Omega, \mathcal{F})$, если она удовлетворяет условиям

μ1) $\mu(A) \geqslant 0$ для любого множества $A \in \mathcal{F}$;

$\mu 2)$ $\forall$ счетного $\{A_i\} \in \mathcal{F}$ таких, что $A_i \cap A_j = \varnothing$, $\forall i \neq j$ мера их объединения равна сумме их мер:

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

Последнее свойство называют *счётное аддитивностью* или σ-аддитивностью меры.

Thr 2.7 (свойство непрерывности меры)**.** *Пусть дана убывающая последовательность* $B_1 \supseteq B_2 \supseteq B_2 \supset B_3 \supset \ldots$ *множеств из* $\mathcal{F}$*, причем* $\mu(B_1) < \infty$*. Пусть* $B = \cap_i^{\infty} B_i$*. Тогда* $\mu(B) = \lim\limits_{n\to\infty} \mu(B_n)$*.*

Def 2.8. Пусть Ω – непустое множество, $\mathcal{F}$ – σ-алгебра его подмножеств. Мера $\mu\colon \mathcal{F} \mapsto \mathbb{R}$ называется *нормированной*, если $\mu(\Omega) = 1$. Другое название нормированной меры – *вероятность*.

Def 2.9. Пусть Ω – пространство элементарных исходов, $\mathcal{F}$ – σ-алгебра его подмножеств (событий). *Вероятностью* или *вероятностной мерой* на $(\Omega, \mathcal{F})$ называется функция

$$\mathrm{P}\colon \mathcal{F} \mapsto \mathbb{R}$$

обладающая свойствами

P1) $\mathrm{P}(A) \geqslant 0$ для любого события $A \in \mathcal{F}$;

P2) для любого счётного набора *попарно несовместных* событий $\{A_i\} \in \mathcal{F}$ имеет равенство

$$\mathrm{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{k=1}^{\infty} \mathrm{P}(A_i);$$

P3) вероятность достоверного события равна единице: $\mathrm{P}(\Omega) = 1$.

Свойства (P1) – (P3) называют *аксиомами вероятности*.

Def 2.10. Тройка $\langle \Omega, \mathcal{F}, \mathrm{P} \rangle$, в которой Ω – пространство элементарных исходов, $\mathcal{F}$ – σ-алгебра его подмножеств и P – вероятная мера на $\mathcal{F}$, называется *вероятностным пространством*.

Вообще, для вероятности верны следующие свойства

1. $\mathrm{P}(\varnothing) = 0$.
2. Для любого конечного набора попарно несовместных событий $A_1, \ldots, A_n \in \mathcal{F}$ имеет место равенство $\mathrm{P}(A_1 \cup \ldots \cup A_n) = \mathrm{P}(A_1) + \ldots + \mathrm{P}(A_n)$.
3. $\mathrm{P}(\bar{A}) = 1 - \mathrm{P}(A)$.
4. Если $A \subseteq B$, то $\mathrm{P}(B \backslash A) = \mathrm{P}(B) - \mathrm{P}(A)$.
5. $A \subseteq B$, то $\mathrm{P}(A) \leqslant \mathrm{P}(B)$.
6. $\mathrm{P}(A_1 \cup \ldots \cup A_n) \leqslant \sum_{i=1}^{n} P(A_i)$.

И это всё, конечно, хорошо, но если мы хотим что-то посчитать, то

Thr 2.11 (Формула включения-исключения)**.** *Для вероятности, в частности для двух событий, верно, что*

$$\mathrm{P}(A \cup B) = P(A) + P(B) - P(A \cap B)$$

и, обобщая, для объединения n *множеств*

$$\mathrm{P}(A_1 \cup \ldots \cup A_n) = \sum_{i=1}^{n} \mathrm{P}(A_i) - \sum_{i<j} \mathrm{P}(A_i A_j) + \sum_{i<j<m} \mathrm{P}(A_i A_j A_m) - \ldots + (-1)^{n-1}\, \mathrm{P}(A_1 A_2 \ldots A_n).$$

3 Условная вероятность и независимость

3.1 Условная вероятность

Def 3.1. Условной вероятностью события A при условии, что произошло событие B, называется число

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

которое само собой определено только при $\mathrm{P}(B) = 0$.

Thr 3.2. *Если* $\mathrm{P}(B) > 0$ *и* $\mathrm{P}(A) > 0$, *то*

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

Thr 3.3. *Для любых событий* $A_1, \dots, A_n$ *верно равенство:*

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

если все участвующие в нём условные вероятности определены.

3.2 ($\mathfrak{B}$) Операции с вероятностями

Thr 3.4. *Вероятность произведения (совмещения) двух зависимых событий равна произведению вероятностей одного из них на условную вероятность другого, вычисенную в предположении, что первое событие уже наступило:*

$$\mathrm{P}(AB) = \mathrm{P}(A)\,\mathrm{P}_A(B) = \mathrm{P}(A)\,\mathrm{P}(B|A).$$

Con 3.5. *Для конечного числа зависимых событий верна формула:*

$$\mathrm{P}(A_1 \cdot A_2 \cdot \ldots \cdot A_k) = \mathrm{P}(A_1)\,\mathrm{P}_{A_1}(A_2)\,\mathrm{P}_{A_1,A_2}(A_3) \dots \mathrm{P}_{A_1 \dots A_{k-1}}(A_k).$$

3.3 Независимость событий

Def 3.6. События A и B называются *независимыми*, если $\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B)$.

Из этого определения вытекают следующие леммы.

Lem 3.7. *Пусть* $\mathrm{P}(B) > 0$. *Тогда события* A *и* B *независимы тогда и только, когда* $\mathrm{P}(A|B) = \mathrm{P}(A)$.

Lem 3.8. *Пусть* A *и* B *несовместны. Тогда независимыми они будут только в том случае, если* $\mathrm{P}(A) = 0$ *или* $\mathrm{P}(B) = 0$.

Другими словами несовместные события не могут быть независимыми. Зависимость между ними – просто причинно-следственная: если $A \cap B = \varnothing$, то $A \subseteq \bar{B}$, т.е. при выполнении A события B *не происходит.*

Lem 3.9. *Если события* A *и* B *независимы, то независимы и события* A *и* $\bar{B}$, $\bar{A}$ *и* B, $\bar{A}$ *и* $\bar{B}$.

Def 3.10. События $A_1, \dots, A_n$ называются *независимыми в совокупности*, если для любого $1 \leqslant k \leqslant n$ и любого набора различных меж собой индекс $1 \leqslant i_1 < \ldots < i_k \leqslant n$ имеет место равенство

$$\mathrm{P}(A_{i_1} \cap \ldots \cap A_{i_k}) = \mathrm{P}(A_{i_1}) \cdot \ldots \cdot \mathrm{P}(A_{i_k}).$$

3.4 Формула полной вероятности

Def 3.11. Конечный или счётный набор попарно несовместных событий $\{H_i\}$ таких, что $\mathrm{P}(H_i) > 0\ \forall i$ и $\cup_i H_i = \Omega$, называется *полной группой событий* или разбиением пространства Ω. Также события, образующие полную группу событий, часто называют *гипотезами.*

При подходящем выборе гипотез для любого события A могут быть сравнительно просто вычислены $\mathrm{P}(A|H_i)$ и, собственно, $\mathrm{P}(H_i)$. Как посчитать вероятность события A?

Thr 3.12 (формула полной вероятности)**.** *Пусть дана полная группа событий* $\{H_i\}$. *Тогда вероятность любого события* A *может быть вычислена по формуле*

$$\mathrm{P}(A) = \sum_{i=1}^{\infty} \mathrm{P}(H_i) \cdot \mathrm{P}(A|H_i).$$

3.5 Формула Байеса

Thr 3.13 (формула Байеса)**.** *Пусть* $\{H_i\}$ *– полная группа событий, и* A *– некоторое событие,* $\mathrm{P}(A) > 0$. *Тогда условная вероятность того, что имело место событие* H_k, *если в рещультате эксперимента наблюдалось событие* A, *может быть вычислена по формуле*

$$\mathrm{P}(H_k|A) = \frac{\mathrm{P}(H_k) \cdot \mathrm{P}(A|H_k)}{\sum\limits_{i=1}^{\infty} \mathrm{P}(H_i) \cdot \mathrm{P}(A|H_i)}. \tag{3.1}$$

Def 3.14. Вероятности $\mathrm{P}(H_i)$, вычисленные заранее, до проведения эксперимента, называют[1] *априорными* вероятностями. Условные вероятности $\mathrm{P}(H_i|A)$ называют[2] *апостериорными* вероятностями.

Формула Байеса позволяет переоценить заранее известные вероятности после того, как получено знание о результате эксперимента. `Эта формула находит многочисленные применения в экономике, статистике, социлогии и т.п`

4 Схема Бернулли

4.1 Распределение числа успехов в n испытаниях

Def 4.1. *Схемой Бернулли* называется последовательность независимых в совокупности испытаний, в каждом из которых возможны лишь два исхода – «успех» и «неудача», при этом успех ✓в одном испытании происходит с вероятностью $p \in (0, 1)$, а неудача ✗– с вероятностью $q = 1 - p$.

В испытаниях схемы Бернулли независимость в совокупности испытаний означает, что при любом n независимы в совокупности события успехов в каждом событие.

Эти события принадлежат одному и тому же пространству элементарных исходов, полученному декартовым произведением бесконечного числа двухэлементных множеств {✓,✗}:

$$\Omega = \{(a_1, a_2, \ldots, a_n) \mid a_i \in \{✓,✗\}, n \in \mathbb{Z}_+\}.$$

Далее количество успехов для n испытаний схемы Бернулли будем называть ν_n. Заметим, что $\nu_n \in \mathbb{Z}_+ \cap [0, n]$.

Thr 4.2 (формула Бенулли)**.** *При любом $k = 0, 1, \ldots, n$ имеет место равенство:*

$$\mathrm{P}(\nu_n = k) = C_n^k p^k q^{n-k}.$$

Def 4.3 ($\mathfrak{D}$). Набор чисел $\{C_n^k p^k q^{n-k},\ k = 0, 1, \ldots, n\}$ называется *биномиальным* распределением.

4.2 Номер первого успешного испытания

Далее, для схемы Бернулли, введем величину $\tau \in \mathbb{Z}_+ \cap [1, +\infty)$ равную номеру перого успешного испытания.

Thr 4.4. *Вероятность того, что первый успех произойдёт в испытании с номером $k \in \mathbb{N} \cap [1, +\infty)$, равна*

$$\mathrm{P}(\tau = k) = pq^{k-1}.$$

Def 4.5 ($\mathfrak{D}$). Набор чисел $\{pq^{k-1} \mid k = 1, 2, \ldots\}$ называется *геометрическим* распределением вероятностей.

Thr 4.6 («Нестарение» геометрического распределения)**.** *Пусть $\mathrm{P}(\tau = k) = pq^{k-1}$ $\forall k \in \mathbb{N}$. Тогда для любых неотрицательных целых n и k имеет место равенство:*

$$\mathrm{P}(\tau > n + k \mid \tau > n) = \mathrm{P}(\tau > k).$$

Другими название – свойство отсутствия последствия.

4.3 Независимые испытания с несколькими исходами

Теперь рассмотрим схему независимых испытаний независимых испытаний уже не с двумя, а с болбшим количество возможных результатов в каждом испытании.

Пусть возможны m исходов, i-й исход в одном испытании случается с вероятностью p_i, где $\sum_i p_i = 1$. Через $\mathrm{P}(n_1, \ldots, n_m)$ обозначим вероятность того, что в n независимых испытаниях первый исход случится n_1 раз, ..., m-исход – n_m раз.

Thr 4.7. *Для любого n и любых неотрицательных целых чисел $\{n_i\}$, сумма которых равна n, верна формула*

$$\mathrm{P}(n_1, \ldots, n_m) = \frac{n!}{n_1! \ldots n_m!} p_1^{n_1} \cdot \ldots \cdot p_m^{n_m}.$$

[1] *a'priori* – « до опыта ».
[2] *a'priori* – « после опыта ».

Def 4.8 ($\mathfrak{D}$). Набор чисел

$$\left\{ \frac{n!}{n_1! \dots n_m!} p_1^{n_1} \cdot \dots \cdot p_m^{n_m} \;\middle|\; n = 1, 2, \dots \right\}$$

называется *мультиномиальным (полиномиальным)* распределением.

4.4 Теорема Пуассона для схемы Бернулли

Сформулируем теорему о приближенном вычислении вероятности иметь k успехов в большом числе испытаний Бернулли с маленькой вероятностью успеха p.

Thr 4.9 (теорема Пуассона). *Пусть $n \to \infty$ и $p_n \to 0$ так, что $np_n \to \lambda > 0$. Тогда для любого $k \geqslant 0$ вероятность получить k успехов в n испытаниях схемы Бернулли с вероятностью успеха p_n*

$$\mathrm{P}(\nu_n = k) = C_n^k p_n^k (1 - p_n)^{n-k} \quad \to \quad \frac{\lambda^k}{k!} e^{-\lambda}. \tag{4.1}$$

то есть стремится к величине $\lambda^k e^{-\lambda}/k!$.

Def 4.10 ($\mathfrak{D}$). Набор чисел

$$\left\{ \frac{\lambda^k}{k!} e^{-\lambda} \;\middle|\; k = 0, 1, 2, \dots \right\}$$

называется *распределением Пуассона* с параметром $\lambda > 0$.

Для всех этих распределений можно посчитать вектора средних и матрицы ковариации.

5 Случайные величины и их распределения

5.1 Случайные величины

Пусть задано вероятностное пространство $\langle \Omega, \mathcal{F}, \mathrm{P} \rangle$.

Def 5.1. Функция $\xi \colon \Omega \mapsto \mathbb{R}$ называется *случайное величиной*, если для любого борелевского множества $B \in \mathfrak{B}(\mathbb{R})$ множество $\xi^{-1}(B)$ является событием, т.е принадлежит σ-алгебре $\mathcal{F}$.

Множество $\xi^{-1}(B) = \{\omega \mid \xi(\omega) \in B\}$, состоящее из элементарных исходов ω, называется *полным прообразом множества B*. Можно немного другим способом сформулировать требования к величине:

Def 5.2. Функция $\xi \colon \Omega \mapsto \mathbb{R}$ называется случайной величиной, если для любых веществнных $a < b$ множество

$$\{\omega \colon \xi(\omega) \in (a, b)\} \in \mathcal{F}$$

принадлежит σ-алгебре.

5.2 Распределения случайных величин

Def 5.3. *Распределением* случайной величины ξ называется вероятностная мера $\mu(B) = \mathrm{P}(\xi \in B)$ на множестве борелевских подмножеств $\mathbb{R}$.

Можно представить себе распределение случайной величины ξ как соответствие между множествами $B \in \mathfrak{B}(\mathbb{R})$ и вероятностями $\mathrm{P}(\xi \in B)$.

Def 5.4. Если две функции ξ и η отличаются на множестве меры нуль, при этом имеют одинаковое распределение, то говорят, что ξ и η совпадают *почти наверное*: $\mathrm{P}(\xi = \eta) = 1$.

Def 5.5. Случайная велчина ξ имеет *дискретное* распределение, если существует конечный, или счётный набор чисел $\{a_i\}$ такой, что

$$\mathrm{P}(\xi = a_i) > 0 \quad \forall i, \qquad \sum_{i=1}^{\infty} \mathrm{P}(\xi = \alpha_i) = 1.$$

Значения эти называют *атомами*: ξ имеет атом в точке x, если $\mathrm{P}(\xi = x) > 0$.

Если случайная величина ξ имеет дискретное распределение, то для любого $B \subseteq \mathbb{R}$

$$\mathrm{P}(\xi \in B) = \sum_{a_i \in B} \mathrm{P}(\xi = a_i).$$

Вообще дискретные распредления удобно задавать вероятностной таблицей

ξ	a_1	a_2	a_3	$\ldots$
P	p_1	p_2	p_3	$\ldots$

Def 5.6. Случайная величина ξ имеет *абсолютно непрерывно* распределение, если существует неотрицательная функция $f_\xi(x)$ такая, что для любого борелевского множества B имеет место равенство:

$$\mathrm{P}(\xi \in B) = \int_B f_\xi(x)\,dx.$$

Функцию $f_\xi(x)$ называют *плотностью распределения* величины ξ.

Thr 5.7. *Плотность распределения обладает свойствами:*

$$\text{(f1)} \qquad f_\xi(x) \geqslant 0 \quad \forall x, \qquad\qquad \text{(f2)} \qquad \int_{-\infty}^{+\infty} f_\xi(t)\,dt = 1.$$

Thr 5.8. *Если функция f обладает свойствами* (f1) *и* (f2)*, то существует вероятностное пространство и случаяная величина ξ на нём, для которой f является плотностью распределения.*

Ещё бывает сингулярное распределение[3], смешанные варианты, и всё (*Лебег approved*).

5.3 Функция распределения

Хотелось бы найти некоторый универсальный способ для описания распределения.

Def 5.9. *Функцией распределения* случайной величины ξ называется функция $F_\xi \colon \mathbb{R} \mapsto [0,1]$, при каждом $x \in \mathbb{R}$ равная вероятности случайной величине ξ принимать значения, меньшие x:

$$F_\xi(x) = \mathrm{P}(\xi < x) = \mathrm{P}\{\omega \mid \xi(\omega) < x\}.$$

Далее перечислены основные дискретные и абсолютно непрерывные распределения и найдены их функции распределения.

5.4 (З) Примеры дискретных распределений

Вырожденное распределение. Для удобства вводят *вырожденное распределение*, когда возможен единственный результат при $\mathrm{P}(\xi = c) = 1$, тогда функция распрееления имеет вид

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

В таком случае принято писать, что $\xi \in \mathrm{I}_c$.

Распределение Бернулли. Говорят про *распределение Бернулли* с параметром p ($\xi \in \mathrm{B}_p$), если ξ принимает значения 1 и 0 с вероятностью p и $1-p$ соответственно. Случайная величина ξ с таким распределением равна *числу упехов* в одном испытании схемы Бернулли с вероятностью успеха p. Функция распредления случайной величины ξ тогда равна

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

Биномиальое распределение. Говорят, что случайная величина ξ имеет биномиальное распределение с параметрами $n \in \mathbb{N}$ и $p \in (0,1)$, и пишут $\xi \in \mathrm{B}_{n,p}$, если ξ принимает значения $k = 0, \ldots, n$ с вероятностями $\mathrm{P}(\xi = k) = C_n^k p^k (1-p)^{n-k}$. Случайная величиная с таким распределением имеет смысл *числа успехов* в n исыпытаниях схемы Бернулли с вероятностью успеха p.

Геометрическое распределение. Говорят, что случайная величина τ имеет геометрическое распределение с параметром $p \in (0,1)$, и пишут $\tau \in \mathrm{G}_p$, если τ принимает значения $k = 1, 2, 3, \ldots$ с вероятностями $\mathrm{P}(\tau = k) = p(1-p)^{k-1}$. Случайная величина с таким распределением имеет смысл номера первого успешного испытания в схеме Бернулли с вероятностью успеха р.

[3] На континуальном множестве меры нуль.

Распреeление Пуассона. Говорят, что случайная величина ξ имеет распределение Пуассона с параметром $\lambda > 0$, и пишут $\xi \in \Pi_\lambda$, если ξ принимает значения $k = 0, 1, \ldots$ с вероятностью $\mathrm{P}(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$. Иначе распределение Пуассона называют *распределением числа редких событий*.

Гипергеметрическое распределение. Говорят, что случайная величина ξ имеет гипергеометрическое распределение с параметрами N, $n \leqslant N$ и $K \leqslant N$, если ξ принимает целые значения k такие, что $0 \leqslant k \leqslant K$, $0 \leqslant n - k \leqslant N_K$, с вероятностями $\mathrm{P}(\xi = k) = C_K^k C_{N_K}^{n-k} / C_N^n$. Случайная величина с таким распределением имеет смысл числа белых шаров среди n шаров, выбранных наудачу и без возвращения из урны, содержащей K белых и $N - K$ не белых.

5.5 (З) Примеры абсолютно непрерывных распределений

Равномерное распределение. Говорят, что ξ имеет равномерное распределение на отрезке $[a, b]$ ($\xi \in U_{a,b}$), если плотность распределения ξ постоянна на отрезке $[a, b]$ и равна нуля вне него:

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

Площадь под графиком этой функции равна единице, $f_\xi \geqslant 0$, так что $f_\xi(x)$ действительно плотность.

Легко теперь посчитать функцию распределения величины ξ:

$$F_\xi(x) = \mathrm{P}(\xi < x) = \int_{-\infty}^{x} f_\xi(t)\, dt = \begin{cases} 0, & x < a; \\ \dfrac{x-a}{b-a}, & a \leqslant x \leqslant b, \\ 1, & x > b, \end{cases}$$

что вполне логично. График функции распределения и плотности распределения приведен ниже.

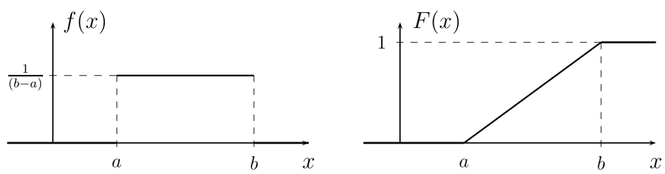

Рис. 1: Плотность и функция распределения $U_{a,b}$

Показательное распределение. Говорят, что ξ имеет показательное (экспоненциальное) распределение с параметром $\alpha > 0$ ($\xi \in \mathrm{E}_\alpha$), если ξ имеет следующую плотность распределения:

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

Функция распределения случайной величины ξ непрерывна:

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

Стоит заметить, что показательное распределение является единственным абсоютно непрерывным распределением, для которого выполнено свойство «нестарения» (а-ля геоетрическое):

Thr 5.10. *Пусть $\xi \in \mathrm{E}_\alpha$. Тогда для любых $x,\ y > 0$ верно, что* $\mathrm{P}(\xi > x + y \mid \xi > x) = \mathrm{P}(\xi > y)$.

Нормальное распределение. Говорят, что ξ имеет *нормальное* (*гауссовское*) распределение с параметрами a, σ^2, где $a \in \mathbb{R}$, $\sigma > 0$ ($\xi \in \mathrm{N}_{a,\sigma^2}$), если ξ имеет плотность распределения вида

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right), \qquad x \in \mathbb{R}. \tag{5.1}$$

Это действительно функция распределения, ведь вспоминая интеграл Пуассона

$$I = \int_{-\infty}^{+\infty} e^{-x^2/2}\, dx = \sqrt{2\pi},$$

нетрудно заменой переменных свести $\int f_\xi(x)\, dx$ к I.

Def 5.11. Нормальное распределение $\mathrm{N}_{0,1}$ называется *стандартным нормальным* распределением.

Для функции распределения нормального закона N_{a,σ^2} далее будет использоваться $\Phi_{a,\sigma^2}(x)$ для функции распределения нормального закона N_{a,σ^2}.

Распределение Коши. Говорят, что ξ имеет распределение Коши с параметрами $a \in \mathbb{R}$, $\sigma > 0$ ($\xi \in \mathrm{C}_{a,\sigma}$), если ξ имеет следующую плотность распределения:

$$f_\xi(x) = \frac{1}{\pi}\frac{\sigma}{\sigma^2 + (x-a)^2}, \qquad \forall x \in \mathbb{R}.$$

Плотность распределения Коши симметрична относительно $x = a$ и похожа на нормальное, но с более толыстыми хвостами на $\pm\infty$. Функция распределения случайной величины ξ с распределением Коши равна

$$F_\xi(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg}\left(\frac{x-a}{\sigma}\right).$$

Гамма-распределение.

Распределение Парето.

5.6 Свойства функций распределения

Общие свойства функций распределения. Функцией распределения случайной величины ξ мы назвали функцию $F_\xi(x) = \mathrm{P}(\xi < x)$.

Thr 5.12. *Любая функция распределения обладает свойствами*

F1) она не убвает;

F2) в прелелах $x \to -\infty$, и $x \to +\infty$ равна 0 и 1 соответственно;

F3) она в любой точке непрерывна слева.

Thr 5.13. *Если функция $F\colon \mathbb{R} \mapsto [0,1]$ удовлетворяет свойствам (F1)-(F3), то F есть функция распределения некоторой случайной величины ξ, т.е. найдётся вероятностное пространство $\langle \Omega, \mathcal{F}, \mathrm{P}\rangle$ и случайная величина ξ на нём такая, что $F(x) \equiv F_\xi(x)$.*

Lem 5.14. *любой точке x_0 разница $F_\xi(x_0+0) - F_\xi(x_0)$ равна $\mathrm{P}(\xi = x_0)$:*

$$F_\xi(x_0+0) = F_\xi(x_0) + \mathrm{P}(\xi = x_0) = P(\xi \leqslant x_0).$$

Lem 5.15. *Для любой случайной величины ξ*

$$\mathrm{P}(a \leqslant \xi < b) = F_\xi(b) - F_\xi(a).$$

Функция распределения дискретного распределения. Как мы поним, функция распределения может быть найдена по талице распределения, как сумма $F_\xi(x) = \mathrm{P}(\xi < x) = \sum_k \mathrm{P}(\xi = a_k)$, где $a_k < x$.

Lem 5.16. *Случайная величина ξ имеет дискретное распределение тогда и только тогда, когда функция распределения $F_\xi(x)$ имеет в точках a_i скачки с величиной $p_i = \mathrm{P}(\xi = a_i) = F_\xi(a_i+0) - F_\xi(a_i)$, и растёт только за счёт скачков.*

Свойства абсолютно непрерывного распределения. Пусть слу- чайная величина ξ имеет абсолюлютно непрерывное распределение с плотностью $f_\xi(t)$. Тогда функция распределения может быть найдена, как интеград.

Lem 5.17. *Если случайная величина ξ имеет абсолютно непрерывное распределение, то её функция распределения всюду непрерывна. Более того её функция распределенеия дифференцируема почти всюду: $f_\xi(x) = F'_\xi(x) = d_x F_\xi(x)$.*

Функция распределения сингулярного распределения.

Функция распределения смешанного распределения.

5.7 Свойства нормального распределения

Lem 5.18. *Для любого $x \in \mathbb{R}$ справедливо соотношение:*

$$\Phi_{a,\sigma^2}(x) = \Phi_{0,1}\left(\frac{x-a}{\sigma}\right).$$

Аналогичное утверждение для случайных величичн: если $\xi \in \mathrm{N}_{a,\sigma^2}$, то $\eta = \frac{\xi - a}{\sigma} \in \mathrm{N}_{0,1}$. Более того, если $\xi \in \mathrm{N}_{a,\sigma^2}$, то

$$\mathrm{P}(x_1 < \xi < x_2) = \Phi_{a,\sigma^2}(x_2) - \Phi_{a,\sigma^2}(x_2) = \Phi_{0,1}\left(\frac{x_2 - a}{\sigma}\right) - \Phi_{0,1}\left(\frac{x_1 - a}{\sigma}\right).$$

В общем вычисления любых вероятностей для нормального распределения сводятся к вычислению $\Phi_{0,1}(x)$, которое обладает следующими свойствами:

- $\Phi_{0,1}(0) = 0.5, \qquad \Phi_{0,1}(-x) = 1 - \Phi_{0,1}(x)$.
- Если $\xi \in \mathrm{N}_{0,1}$, то для любого $x > 0$, верно что $\mathrm{P}(|\xi| < x) = 1 - 2\Phi_{0,1}(-x) = 2\Phi_{0,1}(x) - 1$.

6 Преобразования случайных величин

6.1 Измеримость функций от случайных величин

Пусть на векторном пространстве $\langle \Omega, \mathcal{F}, \mathrm{P} \rangle$ задана случайная величина ξ.

Thr 6.1. *Пусть ξ – случайная величина, а $g\colon \mathbb{R} \mapsto \mathbb{R}$ – борелевская функция, т.е. такая, что для всякого борелевского множества B его прообраз $g^{-1}(B)$ есть снова борелевское множество. Тогда $g(\xi)$ – случайная величина.*

6.2 Распределения функций от случайных величин

Линейные и монотонные преобразования. Если с дискретными распределениями всё понятно, то с абсолютно непрерывными чуть интереснее, о них дальше и поговорим. Пусть случайная величина ξ имеет функцию распределения $F_\xi(x)$ и плотность распределения $f_\xi(x)$. Построим с помощью борелевской функции $g\colon \mathbb{R} \to \mathbb{R}$ случайную величину $\eta = g(\xi)$,и найдём плотность распределения (если она существует).

Thr 6.2. *Пусть ξ имеет функцию распределения $F_\xi(x)$ и плотность распределения $f_\xi(x)$, и постоянная a отлична от нуля. Тогда случайная величина $\eta = a\xi + b$ имеет плотность распределения*

$$f_\eta(x) = \frac{1}{|a|} f_\xi\left(\frac{x-b}{a}\right).$$

Квантильное преобразование. Полезно уметь строить случайные величины с заданным распределением по равномерно распределенной случайной величине.

Thr 6.3. *Пусть функция распределения $F(x) = F_\xi(x)$ непрерывна. Тогда случайная величина $\eta = F(\xi)$ имеет равномерное на отрезке $[0,1]$ распределение.*

Thr 6.4 (alarm). *Пусть $\eta \in U_{0,1}$, а F – произвольная функция распределения. Тогда случайная величина $\xi = F^{-1}(\eta)$ (**«квантильное преобразование»** над η) имеет функцию распределения F.*

Как следствие, для $\eta \in \mathrm{U}_{0,1}$,верны следующие утверждения:

$$-\frac{1}{\alpha}\ln(1-\eta) \in \mathrm{E}_\alpha, \qquad a + \sigma\,\mathrm{tg}(\pi\eta - \pi/2) \in \mathrm{C}_{\alpha,\sigma}, \qquad \Phi_{0,1}^{-1}(\eta) \in \mathrm{N}_{0,1}.$$

7 ✗Многомерные распределения

7.1 Совместное распределение

Пусть случайные величины $\xi_1, \ldots, \xi_n$ заданы на одном вероятностном пространстве $\langle \Omega, \mathcal{F}, \mathrm{P} \rangle$.

Def 7.1. Функция

$$F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = \mathrm{P}(\xi_1 < x_1, \ldots, \xi_n < x_n) \tag{7.1}$$

называется функцией распределения вектора $(\xi_1, \ldots, \xi_n)$ или функцией *совместного* распределения случайных величины $\xi_1, \ldots, \xi_n$.

7.2 Типы многомерных распределений

Далее рассмотрим два типичных случая, когда совместное распределение либо дискретно, либо непрерывно. `Сингулярное распределение не является редкостью: стоит выбрать отрезок на плоскости.`

Def 7.2. Случайные величины ξ_1, ξ_2 имеют *дискретное* совместное распределение, если существует конечный или счётный набор пар числе $\{a_i, b_j\}$ такой, что

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\mathrm{P}(\xi_1 = a_i,\ \xi_2 = b_j) = 1.$$

Таблицу, на пересечении i-й строки и j-го столбца которых стоит $\mathrm{P}(\xi_1 = a_i,\ \xi_2 = b_j)$, называют таблицей совместного распределения случайных величин ξ_1 и ξ_2.

Def 7.3. Случайные величины ξ_1, ξ_2 имеют *абсолютно непрерывное* совместное распредеение, если существует неотрицательная функция $f_{\xi_1,\xi_2}(x, y)$ такая, что для любого множества $B \in \mathfrak{B}(\mathbb{R}^2)$ имеет место равенство

$$\mathrm{P}\left((\xi_1, \xi_2) \in B\right) = \iint_B f_{\xi_1,\xi_2}(x, y)\,dx\,dy.$$

Функция $f_{\xi_1,\xi_2}(x, y)$ называется *плотностью совместного распределения* случайных величин ξ_1 и ξ_2.

Если случайные величины ξ_1 и ξ_2 имеют абсолютно непрерывное совместное распределение, то для любых x_1, x_2 имеет место равенство

$$F_{\xi_1,\xi_2}(x_1, x_2) = \mathrm{P}(\xi_1 < x_1,\ \xi_2 < x_2) = \int_{-\infty}^{x_1}\left(\int_{-\infty}^{x_2} f_{\xi_1,\xi_2}(x, y)\,dy\right)dx.$$

По функции совместного распределения его плотность находится как смешанная частная производная:

$$f_{\xi_1,\xi_2}(x, y) = \frac{\partial^2}{\partial x \partial y}F_{\xi_1,\xi_2}(x, y)$$

для почти всех (x, y).

Thr 7.4. *Если случайные величины ξ_1 и ξ_2 имеютабсолютно непреывное совместное распределение с плотностью $f(x, y)$, то ξ_1 и ξ_2 в отдельности также имеют абсолютно непрерывное распределение с плотностями:*

$$f_{\xi_1}(x) = \int_{-\infty}^{+\infty} f(x, y)\,dy; \qquad f_{\xi_2}(y) = \int_{-\infty}^{+\infty} f(x, y)\,dx.$$

7.3 Примеры многомерных распределений

Многомерное нормальное распределение. Пусть $\Sigma > 0$ – положительно определенная симметричная матрица. Говорят, что вектор $(\xi_1, \ldots, \xi_n)$ имеет многомерное нормально распределение $\mathrm{N}_{\vec{a},\Sigma}$ с вектором средних $\boldsymbol{a}$ и матрицей ковариации Σ, если плотность совместного распределения $f_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n)$ равна

$$f_\xi(\boldsymbol{x}) = \frac{1}{\sqrt{\det\Sigma}(\sqrt{2\pi})^n}\exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{a})^{\mathrm{T}} \cdot \Sigma^{-1} \cdot (\boldsymbol{x} - \boldsymbol{a})\right)$$

В частном случае, когда Σ – диагональная матрица с элементами $\sigma_1^2, \ldots, \sigma_n^2$ на диагонали, совместная плотность превращается в произведение плотностей нормальных величин. Вообще это равенство означает независимость величин $\xi_1, \ldots, \xi_n$.

7.4 Независимость случайных величин

Def 7.5. Случайные величины $\xi_1, \ldots, \xi_n$ называются *независимыми* (в совокупности), если *для любого* набора борелевских множеств $B_1, \ldots, B_n \in \mathfrak{B}(\mathbb{R})$ имеет место равенство

$$\mathrm{P}(\xi_1 \in B_1, \ldots, \xi_n \in B_n) = \mathrm{P}(\xi_1 \in B_1) \cdot \ldots \cdot \mathrm{P}(\xi_n \in B_n).$$

Определение независимости можно сформулировать в терминах функций распределения.

Def 7.6. Случайные величины $\xi_1, \ldots, \xi_n$ независимы (в совокупности), если для любых $x_1, \ldots, x_n$ имеет место равенство

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

Thr 7.7. *Случайные величины $\xi_1, \ldots, \xi_n$ с абсолютно непрерывными распределениями независимы (в совокупности) тогда и только тогда, когда плотность их совместного распределения существует и равна произведению плотностей, т.е.для любых $x_1, \ldots, x_n$ имеет место равенство:*

$$f_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = f_{\xi_1}(x_1) \cdot \ldots \cdot f_{\xi_n}(x_n).$$

7.5 Функции от двух случайных величин

Пусть ξ_1 и ξ_2 – случайные величины с плотностью совместного распределения $f_{\xi_1,\xi_2}(x_1, x_2)$, и задана борелевская функция $g\colon \mathbb{R}^2 \mapsto \mathbb{R}$. Требуется найти функцию (и плотность, если повезет) распределения случайной величины $\eta = g(\xi_1, \xi_2)$.

Thr 7.8. *Пусть $x \in \mathbb{R}$, и область $D_x \subseteq \mathbb{R}^2$ состоит из точек (u, v) таких, что $g(u, v) < x$. Тогда случайная величина $\eta = g(\xi_1, \xi_2)$ имеет функцию распределения*

$$F_\eta(x) = \mathrm{P}(g(\xi_1, \xi_2) < x) = \mathrm{P}\left((\xi_1, \xi_2) \in D_x\right) = \iint_{D_x} f_{\xi_1,\xi_2}(u, v)\, du\, dv.$$

Если ξ_1 и ξ_2 независимы, то распределение $g(\xi_1, \xi_2)$ полностью определяется частными распределениями величин ξ_1 и ξ_2.

Thr 7.9 (формула свёртки)**.** *Если случайные величины ξ_1 и ξ_2 независимы и имеют абсолютно непрерывные распределения с плотностями $f_{\xi_1}(u)$ и $f_{\xi_2}(v)$, то плотность распределения суммы $\xi_1 + \xi_2$ существует и равна «свёртке» плотностей f_{ξ_1} и f_{ξ_2}:*

$$f_{\xi_1+\xi_2}(t) = \int_{-\infty}^{+\infty} f_{\xi_1}(u) f_{\xi_2}(t-u)\, du = \int_{-\infty}^{+\infty} f_{\xi_2}(u) f_{\xi_1}(t-u)\, du. \tag{7.2}$$

8 Числовые характеристики распределений

8.1 Математическое ожидание случайной величины

Далее будет использовать термин *математического ожидания*, и также можно встретить наименования: *среднее значение*, *первый момент.*

Def 8.1. *Математическим ожиданием* $\mathrm{E}(\xi)$ случайной величины ξ с дискретным распределением называется *число*

$$\mathrm{E}(\xi) = \sum_k a_k p_k = \sum_k a_k\, \mathrm{P}(\xi = a_k),$$

если данный ряд абсолютно сходится, т.е. если $\sum_i |a_i|\, p_i < +\infty$. В противном случае говорят, что математическое ожидание *не существует.*

Def 8.2. *Математическим ожиданием* $\mathrm{E}(\xi)$ случайное величины ξ с абсолютно непрерывным распределением с плотностью распределения $f_\xi(x)$ называется *число*

$$\mathrm{E}(\xi) = \int_{-\infty}^{+\infty} x f_\xi(x)\, dx,$$

если этот интеграл абсолютно сходится, т.е. если $\int |x| f_\xi(x)\, dx < +\infty$.

8.2 Свойства математического ожидания

Далее всегда предполагается, что матожидание существует.

(E1) Для $\forall$ борелевской $g\colon \mathbb{R} \mapsto \mathbb{R}$, для дискретного и непрерывного распределения, при существующем E:

$$\mathrm{E}\, g(\xi) = \sum_k g(a_k)\, \mathrm{P}(\xi = a_k), \qquad \mathrm{E}\, g(\xi) = \int_{-\infty}^{+\infty} g(x) f_\xi(x)\, dx.$$

(E3) Матожидание линейно по константам: $\mathrm{E}(c\xi) = c\,\mathrm{E}(\xi)$.

(E4) Матожидание суммы *любых* случайных величин равно сумме их матожиданий: $\mathrm{E}(\xi + \eta) = \mathrm{E}(\xi) + \mathrm{E}(\eta)$.

(E7) Если ξ и η независимы и их матожидания существуют, то $\mathrm{E}(\xi\eta) = \mathrm{E}(\xi) \cdot \mathrm{E}(\eta)$.

8.3 Дисперсия и моменты старших порядков

Def 8.3. Пусть $\mathrm{E}\,|\xi|^k < +\infty$. Число $\nu_k = \mathrm{E}\,\xi^k$ называется *моментом порядка k*, или *k-м моментом* случайной величины ξ, число $\mathrm{E}\,|\xi|^k$ называется *абсолютным k-м моментом.* Число $\mathrm{E}\,[\xi - \mathrm{E}(\xi)]^k$ называется *центральным k-м моментом*, $\mathrm{E}\,|\xi - \mathrm{E}\,\xi|^2$ – *абсолютным центральным k-м моментом.*

Def 8.4. Число $\mathrm{D}(\xi) = \mathrm{E}(\xi - \mathrm{E}\,\xi)^2$ (центральный момент второго порядка) называется *дисперсией* случайной величины ξ. Другими словами, это «среднее значение квадрата отклонения случайной величины ξ от своего среднего».

Def 8.5. Число $\sigma = \sqrt{D\xi}$ называют *среднеквадратичным отклонением* случайной величины ξ.

Thr 8.6 (неравенство Йенсена)**.** *Пусть вещественнозначная функция g выпукла. Тогда для любой случайной величины ξ с конечным первым моментом верно неравенство*

$$\mathrm{E}\, g(\xi) \geqslant g(\mathrm{E}\,\xi),$$

где для вогнутых функций знак неравенства меняется на противоположный.

Lem 8.7. *Если* $\mathrm{E}\,|\xi|^t < \infty$*, то для любого* $0 < s < t$ *верно, что*

$$\sqrt[s]{\mathrm{E}\,|\xi|^s} \leqslant \sqrt[t]{\mathrm{E}\,|\xi|^t}.$$

Также из неравенства Йенсена вытекает ряд удобных неравенств:

$$\mathrm{E}\, e^{\xi} \geqslant e^{\mathrm{E}\,\xi}, \qquad \ldots$$

8.4 Свойства дисперсии

Во всех свойствах ниже предполагается существование вторых моментов случайных величин.

(D1) Дисперсия может быть вычислена по формуле $\mathrm{D}\,\xi = \mathrm{E}\,\xi^2 - (\mathrm{E}\,\xi)^2$.

$$\mathrm{D}\,\xi = \mathrm{E}(\xi - E\xi)^2 = \Big/ a = \mathrm{E}\,\xi \Big/ = \mathrm{E}(\xi^2) - 2a\,\mathrm{E}\,\xi + a^2 = \mathrm{E}\,\xi^2 - (\mathrm{E}\,\xi)^2.$$

(D2) Считая c константной: $\mathrm{D}(c\xi) = c^2\,\mathrm{D}\,\xi$.

(D3) Дисперсия нетрицательна: $\mathrm{D}\,\xi \geqslant 0$, более того обращается в ноль, только при $\xi = \mathrm{const}$ почти наверное.

(D4) $\mathrm{D}(\xi + c) = \mathrm{D}\,\xi$.

(D5) Если ξ и η независимы, то $\mathrm{D}(\xi+\eta) = \mathrm{D}\,\xi + \mathrm{D}\,\eta$. Вообще верна формула

$$D(\xi+\eta) = \mathrm{D}\,\xi + \mathrm{D}\,\eta + 2(\mathrm{E}(\xi\eta) - \mathrm{E}\,\xi\,\mathrm{E}\,\eta). \tag{8.1}$$

(D6) Минимум среднеквадратичного отклонения ξ от точек числовой прямой есть $\mathrm{D}\,\xi$:

$$\mathrm{D}\,\xi = \mathrm{E}(\xi - \mathrm{E}\,\xi)^2 = \min_a \mathrm{E}(\xi - a)^2.$$

8.5 Математические ожидания и дисперсии стандартных распределений

Посчитаем несколько характерных значений для различных распределений:

Имя	$\mathrm{E}\,\xi$	$\mathrm{E}\,\xi^2$	$\mathrm{D}\,\xi$
$\mathrm{B}_{1,p}$	p	p	pq
G_p	$1/p$		qp^{-2}
Π_λ	λ	$\lambda^2+\lambda$	λ
$\mathrm{U}_{a,b}$	$(a+b)/2$	$\frac{1}{3}(a^2+ab+b^2)$	$\frac{1}{12}(b-a)^2$
$\mathrm{N}_{0,1}$	0	1	1
$\mathrm{N}_{a,\sigma}$	a		σ^2
E_α	$1!/\alpha^1$	$2!/\alpha^2$	$1/\alpha^2$
$\mathrm{C}_{0,1}$	$\nexists$	$\nexists$	$\nexists$

8.6 Другие числовые характеристики распределений

Далее кратко познакомимся с другими показателями из статистики.

Def 8.8. *Медианой* распределения случайной величины ξ называется любое из чисел μ таких, что

$$\mathrm{P}(\xi \leqslant \mu) \geqslant \frac{1}{2}, \qquad \mathrm{P}(\xi \geqslant \mu) \geqslant \frac{1}{2}.$$

Обобщая, приходим к понятию *квантили* уровня $\delta \in (0,1)$, так называется решение уравнения $\mathrm{P}(x_\delta) = \delta$, где x_δ отрезает площадь δ слева от себя и $1-\delta$ справа.

Вообще ещё есть такой зоопарк, что квантили уровней кратных 0.01 в прикладной статистике называют *процентилями*, кратных 0.1 – *децилями*, кратных 0.25 – *квартилями*.

Def 8.9. *Модой* абсолютно непрерывного распределения называют любую точку локального максимума плотности распределения. Для дискретных распределений модой считают любое значение a_i, вероятность которого больше соседних.

Для описания *унимодеальных* распределений используют следующие величины:

Def 8.10. *Коэффициентом асимметрии* распределения с конечным третьим моментом называют число

$$\beta_1 = \mathrm{E}\left(\frac{x-a}{\sigma}\right)^3,$$

где $a = \mathrm{E}\,\xi$, а $\sigma = \sqrt{D\xi}$.

Для симметричных распределений коэффициент асимметрии равен нулю, если $\beta_1 > 0$, то график плотности имеет более крутой наклон слева, и более пологий справа.

Def 8.11. *Коэффициентом эксцесса* распределения с конечным четвертым моментом называется число

$$\beta_2 = \mathrm{E}\left(\frac{\xi-a}{\sigma}\right) - 3,$$

где $a = \mathrm{E}\,\xi$, а $\sigma = \sqrt{D\xi}$.

Для нормального распределения $\beta_2 = 0$, при $\beta_2 > 0$ плотность распределения имеет более острую вершину, чем у нормального распределения.

8.7 Производящие функции

Дискретные величины, рассмотренные раннее, принимают только целые значения $X = 0, 1, \ldots$. Нахождение числовых характеристик упрощается, если рассматреть *производящие функции.*

Def 8.12. *Производящей функцией* дискретной целочисленной случайной величины ξ с законом распределения $\mathrm{P}(\xi = k) = p_k$, где $k = 0, 1, \ldots$ называется функция, заданная степенным рядом

$$\mathrm{E}(s^\xi) = P(s) = p_0 + p_1 s + p_2 s^2 + \ldots, \tag{8.2}$$

который сходится по крайней мере для $|s| \leqslant 1$.

Thr 8.13. *Производящая функция суммы независимых случайных величин ξ и η равна произведению производящих функций слагаемых*

$$P_{\xi+\eta}(s) = P_\xi(s) \cdot P_\eta(s). \tag{8.3}$$

Так например для биномиального распределения производящая функция примет вид

$$P(s) = (q + ps)^n.$$

А для геометрического закона распределения

$$\mathrm{P}(s) = ps + pqs^2 + pq^2s^3 + \ldots = \frac{ps}{1-qs}.$$

В случае же Пуассона

$$P(s) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} s^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!} = e^{-\lambda} e^{\lambda s} = e^{\lambda(s-1)}.$$

Thr 8.14. *Сумма независимых случайных величин, распределенных по закону Пуассона, распределена по тому же закону.*

Thr 8.15. *Для дискретной случайной величины ξ с производящей функцией $\mathrm{P}(s)$ выполняются следующие требования:*

$$\mathrm{E}(\xi) = P'_s(1), \qquad \mathrm{D}(\xi) = P''_{s,s}(1) + P'_s(1) - [P'_s(1)]^2. \tag{8.4}$$

8.8 Вычисление моментов через производящие функции

Def 8.16. *Производящей функцией моментов* случайной величины ξ называют математическое ожидание случайной величины $e^{s\xi}$, где s – действительный параметр:

$$\psi_\xi(s) = \mathrm{E}(e^{s\xi}). \tag{8.5}$$

Thr 8.17. *Если случайная величина ξ имеет начальный момент порядка n, то производящая функция $\psi_\xi(s)$ n раз дифференцируема по s, и для всех $k \leqslant n$ выполняется соотношение*

$$\nu_k = \psi_\xi^{(k)}(0). \tag{8.6}$$

Действительно, разлагая функции моментов в ряд Маклорена, можно получить её разложение в ряд с начальными моментами

$$\psi_\xi(s) = 1 + \nu_1 s + \frac{\nu_2}{2!}s^2 + \dots$$

9 Числовые характеристики зависимости

9.1 Ковариация двух случайных величин

Дисперсия суммы двух случайных величин равна

$$\mathrm{D}(\xi+\eta) = \mathrm{D}\,\xi + \mathrm{D}\,\eta + 2\left(\mathrm{E}(\xi\eta) - \mathrm{E}(\xi)\,\mathrm{E}(\eta)\right).$$

Величина $\mathrm{E}(\xi\eta) - \mathrm{E}\,\xi\,\mathrm{E}\,\eta = 0$, если ξ и η независимы, но это верно только в одну сторону, поэтому эту величину используют как «индикатор наличия зависимости» между двумя случайными величинами.

Def 9.1. *Ковариацией* $\mathrm{cov}(\xi,\eta)$ случайных величин ξ и η называется число

$$\mathrm{cov}(\xi,\eta) = \mathrm{E}\left[(\xi - \mathrm{E}\,\xi)(\eta - \mathrm{E}\,\eta)\right]. \tag{9.1}$$

Для ковариации справедливы следующие равенства:

$$\mathrm{cov}(\xi,\eta) = \mathrm{E}(\xi\eta) - \mathrm{E}(\xi)\,\mathrm{E}(\eta); \quad \mathrm{cov}(\xi,\xi) = \mathrm{D}(\xi); \quad \mathrm{cov}(\xi,\eta) = \mathrm{cov}(\eta,\xi); \quad \mathrm{cov}(c\xi,\eta) = c\,\mathrm{cov}(\xi,\eta).$$

Lem 9.2. *Дисперсия суммы нескольких случайных величин вычисляется по формуле:*

$$\mathrm{D}(\xi_1 + \ldots + \xi_n) = \sum_{i=1}^n \mathrm{D}(\xi_i) + \sum_{i\neq j}\mathrm{cov}(\xi_i,\xi_j) = \sum_{i,j}\mathrm{cov}(\xi_i,\xi_j). \tag{9.2}$$

Если ковариация $\mathrm{cov}(\xi,\eta) \neq 0$, то ξ и η зависимы. Найти совместное распределение бывает сложнее, чем посчитать $\mathrm{E}(\xi\eta)$, поэтому, если повезет, и $\mathrm{E}(\xi\eta) \neq \mathrm{E}(\xi)\,\mathrm{E}(\eta)$, то, не находя совместное распределение, мы обнаружим зависимость ξ и η, не находя их совсметного распределения. `Это очень хорошо.`

Однако есть проблема – ковариация не безразмерно, поэтому большие значения коварициии не говорят о более сильной зависимости. Хотелось бы как-то отнормировать $\mathrm{cov}(\xi,\eta)$, получив «безразмерную» величину. Так мы приходим к коэффициенту корреляции.

9.2 Коэффициент корреляции

Def 9.3. *Коэффициентом корреляции* $\rho(\xi,\eta)$ случайных величин ξ и η, дисперсии которых существуют и отличны от нуля, называется число

$$\rho(\xi,\eta) = \frac{\mathrm{cov}(\xi,\eta)}{\sqrt{\mathrm{D}\,\xi}\sqrt{\mathrm{D}\,\eta}}. \tag{9.3}$$

Можно наполнить это достаточно глубоким смыслом. На самом деле это «косинус угла» между двумя элементами $\xi - \mathrm{E}\,\xi$ и $\eta - \mathrm{E}\,\eta$ гильбертова пространства, образованного случайными величинами с нулевым матожиданием и конечным вторым моментом. Пространство набжено скалярным произведением $\mathrm{cov}(\xi,\eta)$ и «нормой», равной корню из дисперсии, или $\sqrt{\mathrm{cov}(\xi,\xi)}$.

Thr 9.4. *Коэффициент корреляции обладает свойствами:*

1) *если ξ и η независимы, то $\rho(\xi,\eta) = 0$;*

2) *всегда $|\rho(\xi,\eta)| \leqslant 1$;*

3) *$|\rho(\xi,\eta)| = 1$ тогда и только тогда, когда ξ и η почти наверное линейно связаны.*

Def 9.5. *Стандартизацией* случайной величины называется преобразование

$$\hat{\xi} = \frac{\xi - \mathrm{E}(\xi)}{\sqrt{\mathrm{D}(\xi)}}. \tag{9.4}$$

В терминах стандартизации чуть проще записывается коэффициент корреляции:

$$\rho(\xi, \eta) = \mathrm{E}\left(\hat{\xi} \cdot \hat{\eta}\right).$$

Def 9.6. Говорят, что ξ и η *отрицательно коррелированы*, если $\rho(\xi, \eta) < 0$; *положительно коррелированы*, если $\rho(\xi, \eta) > 0$; *некоррелированы*, если $\rho(\xi, \eta) = 0$.

Lem 9.7. *Для любых случайных величин ξ и η с конечной и ненулевой дсперсией при любых постоянных $a \neq 0$ и b имеет место равенство*

$$\rho(\alpha\xi + b, \eta) = \mathrm{sign}(a) \cdot \rho(\xi, \eta). \tag{9.5}$$

Разобрать пример 67 и далее.

10 Характеристические функции

10.1 Определение и примеры

Def 10.1. Функция $\varphi_\xi(t) = \mathrm{E}\left(e^{it\xi}\right)$ вещественной переменной t называется *характеричтической функцией* случайной величины ξ.

Например, если характеристическая функция имеет стандратное нормальное распределение, то её характеристическая функция равна

$$\varphi_\xi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx} e^{-x^2/2}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} = \frac{e^{-t^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(x-it)^2/2}\, d(x - it) = e^{-t^2/2}.$$

10.2 Свойства характеристических функций

(Ф1). Характеристическая функция всегда существует: $|\varphi_\xi(t)| = |\,\mathrm{E}\, e^{it\xi}| \leqslant 1$.

(Ф2). По характетристической функции однозначно восстанавливается распределение. Например, если модуль характеристической функции интегрируем на всей прямой, то

$$f_\xi(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \varphi_x(t)\, dt.$$

(Ф3). Характерестическая функция случайной величины $a + b\xi$ связана с характеристической функцией случайной величины ξ равенством

$$\varphi_{a+b\xi}(t) = \mathrm{E}\, e^{it(a+b\xi)} = e^{ita}\, \mathrm{E}\,(i(tb)\xi) = e^{ita} \varphi_\xi(tb).$$

(Ф4). Характеристическая функция суммы независимых случайных величин равна произведению характеричтических функций слагаемых: если случайные величины ξ и η независимы, то

$$\varphi_{\xi+\eta}(t) = \mathrm{E}\, e^{it(\xi+\eta)} = \mathrm{E}(e^{it\xi})\, \mathrm{E}(e^{it\eta}) = \varphi_\xi(t) \varphi_\eta(t).$$

Собственно, это очень простой и приятный инструмент для доказательства *устойчивости* распределений. Чем надо было бы и воспользоваться.

(Ф5). Пусть существует момент порядка $k \in \mathbb{N}$ случайной величины ξ. Тогда характеристическая функция $\varphi_\xi(t)$ непрерывно дифференцируема k раз и её k-я производная в *нуле* связана с моментом порядка k равенством

$$\varphi_\xi^{(k)}(0) = \left(\frac{d^k}{dt^k}\, \mathrm{E}\, e^{it\xi}\right)\bigg|_{t=0} = \left(\mathrm{E}\, i^k \xi^k e^{it\xi}\right)\bigg|_{t=0} = i^k\, \mathrm{E}(\xi^k).$$

Lem 10.2. *Для случайной величины ξ со стандартным нормальным распределением момент чёного порядка $2k$ равен*

$$\mathrm{E}(\xi^{2k}) = (2k-1)!! = (2k-1) \cdot (2k-3) \cdot \ldots \cdot 3 \cdot 1.$$

Все моменты нечётных порядков существуют и равны нулю.

Как только появились производные высших порядков, самое время разложить функцию в ряд Тейлора:

(Ф6). Пусть существует момент порядка $k \in \mathbb{N}$ случайной величина ξ, тогда характеричтическая функция $\varphi_\xi(t)$ в окрестности точки $t = 0$ разлагается в ряд Тейлора

$$\varphi_\xi(t) = \varphi_\xi(0) + \sum_{j=1}^{k} \frac{t^j}{j!} \varphi_\xi^{(j)}(0) + o(|t^k|) = 1 + \sum_{j=1}^{k} \frac{i^j t^j}{j!}\, \mathrm{E}(\xi^j) + o(|t^k|).$$

Thr 10.3 (теорема о непрерывно соответствии)**.** *Случайные величины ξ_n слабо сходятся к случайной величине ξ тогда и только тогда, когда для любого t характеристические функции $\varphi_{\xi-b}(t)$ сходятся к характеристической функции $\varphi_\xi(t)$.*

11 Сходимость последовательностей случайных величин

11.1 Определение и примеры

Плотность многомерного нормального распределения:

$$f_\xi(\boldsymbol{x}) = \left[(\sqrt{2\pi})^n\sqrt{\det\Sigma}\right]^{-1}\cdot\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right),$$

где Σ – симметричная, положитеьно определенная матрица.

12 Контрольная работа №1

T1

Известно, что ξ нормально распределена, и $\mathrm{E}\,\xi = -11$, а $\mathrm{E}\left((2\xi+3)(\xi-2)\right) = 253$, найти $\mathrm{P}(\xi \in (-15, 0.16))$.
Итак, $\mathrm{P}(\xi)$ вида

$$\mathrm{P}(\xi) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\xi-a)^2}{2\sigma^2}\right).$$

Матожидание можем найти, как интеграл вида

$$\mathrm{E}\,\xi = \int_{-\infty}^{+\infty} \frac{\xi}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\xi-a)^2}{2\sigma^2}\right) d\xi = \Big/ \xi - a = t \Big/ = \frac{a}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{t^2}{2\sigma^2}\,dt\right) = a.$$

Аналогчино находим второй момент:

$$\mathrm{E}\,\xi^2 = \int_{-\infty}^{+\infty} \frac{\xi^2}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\xi-a)^2}{2\sigma^2}\right) d\xi = \Big/ \xi - a = t \Big/ = \int_{-\infty}^{+\infty} \frac{t^2}{\sigma\sqrt{2\pi}} \exp\left(-\frac{t^2}{2\sigma^2}\right) dt + \int_{-\infty}^{+\infty} \frac{a^2}{\sigma\sqrt{2\pi}} \exp\left(-\frac{t^2}{2\sigma^2}\right) dt.$$

Первый интеграл легко сводится к Гамма-функции, находим, что

$$\int_{-\infty}^{+\infty} \frac{t^2}{\sigma\sqrt{2\pi}} \exp\left(-\frac{t^2}{2\sigma^2}\right) dt = \Gamma\left(\frac{3}{2}\right) \cdot \frac{(2\sigma^2)^{3/2}}{\sqrt{2\pi}} = \sigma^2,$$

второй интеграл просто сводится к a^2, итого

$$\mathrm{E}\,\xi^2 = \sigma^2 + a^2.$$

Так приходим к системе вида

$$\begin{cases} 2\,\mathrm{E}(\xi^2) - \mathrm{E}(\xi) - 6 = 253 \\ \mathrm{E}(\xi) = a \end{cases} \quad \Rightarrow \quad \begin{cases} \sigma^2 = 3, \\ a = -11. \end{cases}$$

Искомая веростность $\mathrm{P}(\xi \in (-15, 0.16))$ в таком случае равна

$$\int \mathbf{N}(\xi,\,\sigma,\,a)\,d\xi = \frac{1}{2}\mathrm{Erf}\left(\frac{\xi - a}{\sqrt{2}\sigma}\right), \quad \Rightarrow \quad \mathrm{P}(\xi \in (-15, 0.16)) = \frac{1}{2}\left(\mathrm{Erf}\left(\frac{0.16-11}{\sqrt{6}}\right) - \mathrm{Erf}\left(\frac{-15+11}{\sqrt{6}}\right)\right) \approx 0.99$$

T2

Для функции вида

$$f(x) = \begin{cases} C/(x-3), & -5 < x < 2 \\ 0, & x \leqslant -5 \mid\mid 2 \leqslant x \end{cases}$$

являющейся функцией распределения некоторой случайной величины ζ найдём C и характеристики $\mathrm{E}(\zeta)$, $\mathrm{D}(\zeta)$.

Для начала найдём C:

$$1 = \int_{-\infty}^{+\infty} f(x)\,dx = -\int_{-5}^{2} \frac{C}{3-x} = C\ln(3-x)\Big|_{-5}^{2} = -C\ln(8), \quad \Rightarrow \quad C = -\frac{1}{\ln 8}.$$

Теперь найдём характеристики ζ:

$$\mathrm{E}(\zeta) = -\int_{-5}^{2} x\frac{C}{3-x} = \Big/ -\frac{x}{3-x} = \frac{3}{x-3} + 1 \Big/ = \frac{7 - 3\ln(8)}{-\ln 8} \approx -0.37$$

$$\mathrm{E}(\zeta^2) = -\int_{-5}^{2} x^2\frac{C}{3-x} = \Big/ -\frac{x^2}{3-x} = x + \frac{9}{x-3} + 3 \Big/ = \frac{18\ln 2 - 7}{\ln 4},$$

тогда дисперсия ζ:

$$\mathrm{D}(\zeta) = \mathrm{E}(\zeta^2) - \mathrm{E}^2(\zeta) = \frac{7(27\ln(2) - 14)}{18\ln^2(2)} \approx 3.82.$$

T3

Введем для каждого места величину ξ_i, равную 1 в случае нечётного числа и 0 иначе. Число, наверное, подразумевает, что на первом месте не может стоять ноль, но на всякий случай пока обозначим вероятность быть первой цифре нечетной за γ, остальных местах равновероятны значения 0 и 1.

Вероятность существования хотя бы одной нечётной цифры найдём через вероятность их отсутсвия:

$$\mathrm{P}(\exists a_i \in \mathrm{Odd}) = 1 - \left(\frac{1}{2}\right)^2 \cdot (1-\gamma).$$

Матожидание же величины $\xi = \sum_{i=1}^{8} \xi_i$ легко найти, в силу независимости ξ:

$$\mathrm{E}(\xi) = \sum_{i=1}^{8} \mathrm{E}(\xi_i) = 7\,\mathrm{E}(\xi_i) + \gamma.$$

Если на первом месте может стоять 0, то $\gamma = 0.5$ и, соответственно,

$$\mathrm{P}(\exists a_i \in \mathrm{Odd}) = \frac{255}{256} \approx 1 - 3.9 \cdot 10^{-3}, \qquad \mathrm{E}(\xi) = 4.$$

Если же 0 стоять на первом месте не может, то $\gamma = 5/9$ и, соответственно,

$$\mathrm{P}(\exists a_i \in \mathrm{Odd}) = 1 - \frac{1}{128}\frac{5}{9} \approx 1 - 4.3 \cdot 10^{-3}, \qquad \mathrm{E}(\xi) = \frac{73}{18} \approx 4.06.$$

T4

Найдём производящую функция для биномиального распределения, вида

$$\mathrm{P}(\xi = k) = C_n^k p^k q^{n-k},$$

что соответствует количеству успехов в схеме Бернулли, где вероятность успеха p.

Коэффициенты в биноме Ньютона выглядят очень похоже на P, так что заметим, что производящая функция вида

$$P(s) = (q + ps)^n,$$

нам подходит. Найдём матожидание и дисперсию, как

$$\mathrm{E}(\xi) = P'_s(s = 1), \qquad \mathrm{D}(\xi) = P''(1) + P'(1) - \mathrm{E}^2(\xi).$$

Производные $P(s)$:

$$P'(s) = np(q+ps)^{n-1}, \quad P'(1) = np, \qquad P''(s) = n(n-1)p^2(q+ps)^{n-1}, \quad P''(1) = n(n-1)p^2,$$

тогда искомые величины:

$$\mathrm{E}(\xi) = np, \qquad \mathrm{D}(\xi) = np(1-p) = npq.$$

T5

Известно, что $\mathrm{E}\,x = 6$, $\mathrm{E}\,y = 19$, $\mathrm{D}\,x = 7$, $\mathrm{D}\,y = 12$, тогда матожидание и дисперсия для $z = 3x - 2y$ равны

$$\mathrm{E}\,z = \mathrm{E}(3x) - \mathrm{E}(2y) = 3\,\mathrm{E}(x) - 2\,\mathrm{E}(y) = -20, \qquad \mathrm{D}\,z = \mathrm{D}(3x) + \mathrm{D}(2y) = 9\,\mathrm{D}(x) + 4\,\mathrm{D}(y) = 111.$$

T6

Для поиска коэффициента корреляции сначала найдём дисперсию и матожидание количества людей, выходящих на 8 этаже, (ξ) и количества людей, выходящих на 8 этаже или выше (η).

Представим величину ξ как сумму четырёх других $\xi = \sum_{i=1}^{4} \xi_i$, где ξ – вероятность выйти на 8 этаже для каждого из четырех людей:

ξ	1	0
P	1/12	11/12

В силу независимости ξ_i верно, что

$$\mathrm{E}(\xi) = \mathrm{E}\left(\sum_{i=1}^{4} \xi_i\right) = \sum_{i=1}^{4} \mathrm{E}(\xi_i) = \frac{4}{12}, \qquad \mathrm{D}(\xi) = \mathrm{D}\left(\sum_{i=1}^{4} \xi_i\right) = \sum_{i=1}^{4} \mathrm{D}(\xi_i) = 4\left(\frac{1}{12} - \left(\frac{1}{12}\right)^2\right) = \frac{11}{36}.$$

Аналогично найдём характеристики η, представив через сумму независимых величин $\eta = \sum_{i=1}^{4} \eta_i$, где η_i – вероятность для каждого человека выйти на этаже, выше восьмого

η	1	0
P	5/12	7/12

Тогда, аналогично, в силу незаивимости η_i, находим

$$\mathrm{E}(\eta) = \sum_{i=1}^{4} \mathrm{E}(\eta_i) = 4 \cdot \frac{5}{12} = \frac{5}{3}, \qquad \mathrm{D}(\eta) = \sum_{i=1}^{4} \mathrm{D}(\eta_i) = 4 \cdot \left(\frac{5}{12} - \frac{25}{144}\right) = \frac{35}{36}.$$

Теперь найдём матожидание $\mathrm{E}(\xi\eta)$, построив таблицу $\mathrm{P}(\xi, \eta)$, где ξ и η принимает значения от 0 до 4. Заметим, что таблица будет верхнетреугольной: если на 8 этаже вышло n людей, то $\eta \geqslant n$. Сформировать вероятность $\mathrm{P}(\xi = \xi_0, \eta = \eta_0)$ можно, выбирая ξ_0 людей из 4 – оставшихся на 8 этаже, выбирая $\eta - \xi$ людей из $4 - \xi$ – оказавшихся на 9 этаже и выше, где вероятность оказаться ниже 8 – (7/12), и вероятность быть на 9 и выше – (5/12), итого находим

$$\mathrm{P}(\xi = \xi_0, \eta = \eta_0) = \binom{4}{\xi_0} \left(\frac{1}{12}\right)^{\xi_0} \binom{4-\xi_0}{\eta_0 - \xi_0} \left(\frac{7}{12}\right)^{4-\eta_0} \left(\frac{5}{12}\right)^{\eta_0 - \xi_0},$$

а искомое матожидание тогда будет равно

$$\mathrm{E}(\xi\eta) = \sum_{\xi_0=0}^{4} \sum_{\eta_0=0}^{4} \xi_0 \eta_0 \, \mathrm{P}(\xi =_{\xi_0}, \eta = \eta_0) = \frac{3}{4},$$

что нетрудно получить прямым вычислением. Конечно, судя по простоте ответа, его можно было получить и более простым путём, но зато мы уверены в результатах.

Наконец, корреляция ξ и η, равна

$$\rho(\xi, \eta) = \frac{\mathrm{cov}(\xi, \eta)}{\sqrt{\mathrm{D}\,\xi}\sqrt{\mathrm{D}\,\eta}} = \frac{\mathrm{E}(\xi\eta) - \mathrm{E}(\xi)\,\mathrm{E}(\eta)}{\sqrt{\mathrm{D}\,\xi}\sqrt{\mathrm{D}\,\eta}} = \sqrt{\frac{7}{55}} \approx 0.36.$$

T7

Так как доска небольшая, то, имея калькулятор, ничего в принципе не мешает просто посчитать количество доступных путей:

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	3	6	10	15	21	28	36	45	55	66	78	91	105	120
1	4	10	20	35	56	84	120	165	220	286	364	455	560	680
1	5	15	35	70	126	210	330	495	715	1001	1365	1820	2380	3060
1	6	21	56	126	252	462	792	1287	2002	3003	4368	6188	8568	11628
1	7	28	84	210	462	924	1716	3003	5005	8008	12376	18564	27132	38760
1	8	36	120	330	792	1716	3432	6435	11440	19448	31824	50388	77520	116280
1	9	45	165	495	1287	3003	6435	12870	24310	43758	75582	125970	203490	319770

Таблица 1: Заполненная количеством доступных путей доска к задаче №T7

Расчёт происходит из предположения о том, что у количество путей $N[i, j]$ равно

$$N[i, j] = N[i-1, j] + N[i, j-1],$$

а левый столбей и верхняя строка «заполнены» единицами – существует единственный способ добраться до этой клетки.

Если нас интересует движение такое, что последние три клетки были сделаны по короткой стороне доски, то вероятность такого маршрута:

$$\mathrm{P}_0 = \frac{11628}{319770} = \frac{2}{55} \approx 3.64 \cdot 10^{-2}.$$

Вообще можно заметить в числах биномиальные коэффициенты – действительно, достигая[4] $[i, j]$ клетки, мы делаем i шагов вправо и j вниз, то есть необходимо в $i+j$ элементах выбрать i элементов (или j), тогда искомая вероятность

$$P_0 = \binom{14+8}{8} \bigg/ \binom{14+5}{5} = \frac{2}{55},$$

что сходится с прямым вычислением.

[4] Формулы удобнее выглядят, когда i и j нумеруются с 0.

T8

Известно следующее совместное распределение:

$$\begin{array}{c|ccc} \xi\backslash\eta & -2 & 0 & 2 \\ \hline -1 & \alpha & \beta & 2/13 \\ 1 & 3/13 & 2/13 & 1/13 \end{array}, \qquad \alpha+\beta = \frac{5}{13}.$$

где также известно, что $7\,\mathrm{D}(\xi) = 19\,\mathrm{D}(\eta)$.

Для начала найдём первые и вторые моменты для ξ и η:

$$\begin{aligned} \mathrm{E}(\xi) &= -2\cdot\left(\alpha + \frac{3}{13}\right) + 2\cdot\frac{3}{13} = -2\alpha, \\ \mathrm{E}(\eta) &= -1\cdot\left(\alpha+\beta+\frac{2}{13}\right) + 1\cdot\frac{6}{13} = -\frac{1}{13}, \\ \mathrm{E}(\xi^2) &= 4\cdot\left(\frac{6}{13}+\alpha\right) = \frac{24}{13} + 4\alpha, \\ \mathrm{E}(\eta^2) &= 1. \end{aligned}$$

Теперь можем перейти к квадратному уравнению

$$19\cdot\left(1 - \frac{1}{13^2}\right) = 7\left(\frac{24}{13} + 4\alpha - 4\alpha^2\right), \quad \Rightarrow \quad 13^2(x^2 - x) + 36 = 0, \quad \Rightarrow \quad x_1 = \frac{4}{13}, \quad x_2 = \frac{9}{13}.$$

Но, так как $\alpha+\beta = 5/13$, а также $\operatorname{sign}\alpha = \operatorname{sign}\beta = 1$, то $\alpha < 5/13$, а значит искомая величина

$$\alpha = \frac{4}{13}.$$

13 Контрольная работа №2

Первая задача

Вероятность выпадения решки равна $p_0 = 0.42$. Монетка подброшена 1000 раз, и решка выпала 360 раз. Сколько раз необходимо подрбросить такую же монетку, чтобы доля выпавших решек отличалась от p_0 менее, чем в первые 1000 бросков с вероятностью $p_1 = 0.95$.

Thr 13.1 (ЦПТ Ляпунова)**.** *Пусть ξ_1, ξ_2, … – независимые и одинакоово распределенные случайные величины с конечной и ненулевой дисперсией: $0 < \mathrm{D}\,\xi_1 < \infty$. Тогда имеет место слабая сходимость*

$$\frac{S_n - n\,\mathrm{E}\,\xi_1}{\sqrt{n\,\mathrm{D}\,\xi_1}} \underset{n\to\infty}{\to} \mathrm{N}_{0,1}, \qquad S_n = \xi_1 + \ldots + \xi_n.$$

последовательности центрированных и нормированных сумм случайных величин к стандартному нормальному распределению.

Точнее S_n стремится к N_{a,σ^2}, где в пределах данной задачи верно, что $A = n\,\mathrm{E}\,\xi_1 = 0.42n$, а $\sigma = \sqrt{n\,\mathrm{D}\,\xi_1} = \sqrt{n0.42(1-0.42)} = \sqrt{npq}$.

По условиям задачи требуется попадание в интервал $[a, b] = [0.36n,\, 0.48n]$. Тогда

$$\int_a^b \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-A)^2}{2\sigma^2}\right) = \Phi\left(\sqrt{n}\frac{A-a}{\sqrt{2pq}}\right) = 0.95, \quad \Rightarrow \quad n = \frac{2X^2pq}{(A-a)^2} = 260,$$

где $X = 1.38$ можно найти по таблице.

Вторая задача

Известно, что ковариационная матрица случайного вектора $(X,\,Y,\,Z)^{\mathrm{T}}$ равна

$$M = \begin{pmatrix} 2 & -1 & \lambda \\ -1 & 2 & 1 \\ \lambda & 1 & 3 \end{pmatrix}.$$

Хочется найти все возможные значения λ, а также λ, соответсвующий минимальной вариации величины $\xi = X + \lambda Y - 2Z$.

Для начала поймём возможные значения λ: матрица неотрицательно определена, а значит, по критерию Сильвестра:

$$\det M = 7 - 2\lambda - 2\lambda^2 > 0, \qquad \Rightarrow \Big/ \lambda_{1,2} = \frac{1}{2}\left(-1 \pm \sqrt{15}\right) \Big/ \Rightarrow \qquad \lambda \in \left[-\frac{1}{2} - \frac{\sqrt{15}}{2};\ -\frac{1}{2} + \frac{\sqrt{15}}{2}\right].$$

Теперь можем найти оптимальное значение λ для $\boldsymbol{\xi} = X + \lambda Y - 2Z = (1,\ \lambda,\ -2)^{\mathrm{T}}$ в базисе (X, Y, Z):

$$\boldsymbol{\xi}^{\mathrm{T}} M \xi = 2\lambda^2 - 10\lambda + 14 = 2\left(\lambda - \frac{5}{2}\right)^2 + \frac{3}{2}, \qquad \Rightarrow \qquad \lambda = \frac{5}{2}.$$

Однако можно заметить, что верхняя граница $(-1 + \sqrt{15})/2 \approx 1.44 < 2.5$, следовательно минимум достигается на правой границе $\lambda_{\text{opt}} = \frac{1}{2}\left(-1 + \sqrt{15}\right)$.

Четвертая задача

Известно, что плотность распределения переменной Y дана

$$f_Y(y) = C \exp\left(-y^2 + 4y - 10\right), \qquad y \in \mathbb{R}.$$

Найдём константу C, а также матожидание и дисперсию Y.

Заметим, что $f_Y(y) = C \exp\left(-(y-2)^2 - 6\right) = \frac{C}{e^6} \exp\left(-(y-2)^2\right)$, – нормальное распределение с $\mathrm{E}\, Y = a = 2$. Осталось найти

$$\int_{-\infty}^{+\infty} \frac{C}{e^6} \exp\left(-(y-2)^2\right)\, dy = \sqrt{\pi} C e^{-6} = 1, \qquad \Rightarrow \qquad C = \frac{e^6}{\sqrt{\pi}}.$$

Итого, распределение перепишется в виде

$$f_\xi(y) = \frac{1}{\sqrt{2\pi}\sqrt{\frac{1}{2}}} \exp\left(-\frac{(y-2)^2}{2(\frac{1}{\sqrt{2}})^2}\right),$$

собственно $\mathrm{D}\, Y = \sigma^2 = 1/2$.

Пятая задача

Случайные величины $X_1, \ldots, X_{100}$ независимы и одинаково распределены, в частности с $N(0, \sigma^2)$, $\sigma = 2$. Найдём распределение вектора $(Y,\ Z)^{\mathrm{T}}$, где $Y = X_{61} + X_{62} + \ldots + X_{100}$, и $Z = X_1 + X_2 + \ldots + X_{80}$.

Lem 13.2. *Если две случайные величины $\xi \in N_{a_1, \sigma_1^2}$, $\eta \in N_{a_2, \sigma_2^2}$ независимы, то $\xi + \eta \in N_{a_1 + a_2, \sigma_1^2 + \sigma_2^2}$.*

Lem 13.3. *Пусть есть две величины $\xi_1 \in N_{a_1, \sigma_1^2}$ и $\xi_2 \in N_{a_2, \sigma_2^2}$ с ковариациоей σ_{12}, тогда $\det \Sigma = \sigma_1^2 \sigma_2^2 (1 - \rho^2)$, где $\rho = \sigma_{12}/(\sigma_1 \sigma_2)$ – коэффицент корреляции. Плотность двумерного нормального распределения в этом случае:*

$$f(x_1, x_2) = \frac{1}{\sqrt{2\pi}^2 \sigma_1 \sigma_2 \sqrt{1 - \rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1 - a_1)^2}{\sigma_1^2} - \rho \frac{2(x_1 - a_1)(x_2 - a_2)}{\sigma_1 \sigma_2} + \frac{(x_2 - a_2)^2}{\sigma_2^2}\right]\right).$$

И это замечательно, ведь по условию X_i и X_j независимы при $i \neq j$, а тогда Y и Z распределены нормально. Получается, остается найти ковариацию $\rho(Y, Z)$, и мы найдём искомое распределение.

Для начала, $\mathrm{D}\, Y = 40\sigma^2$, $\mathrm{D}\, Z = 80\sigma^2$, также по условию $a_1 = a_2 = 0$. Ковариация Y и Z по линейности:

$$\operatorname{cov}\left(\textstyle\sum_{i=61}^{100} X_i,\ \sum_{j=1}^{80} X_j\right) = \sum_{i=61}^{100} \sum_{j=1}^{80} \operatorname{cov}(X_i, X_j) = 20\sigma^2, \qquad \Rightarrow \qquad \rho = \frac{20}{\sqrt{40 \cdot 80}} = \frac{1}{2\sqrt{2}} = 2^{-3/2}.$$

где учтено, что

$$\operatorname{cov}(X_i,\ X_i) = \mathrm{E}\, X_i^2 - \mathrm{E}^2\, X_i = \mathrm{D}\, X_i = \sigma^2.$$

Итого, искомое распределение:

$$f(Y, Z) = \frac{1}{2\pi\sigma^2 \cdot 20\sqrt{7}} \exp\left(-\frac{1}{140\sigma^2}\left[2Y^2 - YZ + Z^2\right]\right).$$