

Теория Вероятностей

Лекции

Савва Чубий, БПИ233

2024–2025

2024-09-06	
Введение	3
Основные понятия	3
Классическое определение вероятности	4
Свойства вероятности	4
2024-09-13	
Геометрическое определение вероятности	4
Свойства $P(A)$	4
Задача	4
Частотное (статистическое) определение	5
Аксиоматическое определение Колмагорова	5
Свойства $P(A)$	5
Условная вероятность	6
2024-09-20	
Независимость в совокупности	7
Теорема умножения вероятностей	7
Биномиальная схема испытаний Бернулли	8
Наиболее вероятное число успехов	8
Формула полной вероятности	9
2024-09-27	
Формула Байеса	10
Задача	10
Случайные величины	11
Дискретные случайные величины	11
Числовые характеристики дискретных случайных величин	12
Математическое ожидание	12
2024-10-04	
Свойства математического ожидания	12
Дисперсия	12
Свойства дисперсии	12
Другие	13

Часто встречающиеся дискретные распределения	13
Распределение Бернулли	13
Биномиальное распределение	13
Распределение Пуассона	14
Геометрическое распределение	15
<hr/>	
2024-10-11	
Непрерывные случайные величины	15
Свойства плотности распределения	16
Числовые характеристики	17
Математическое ожидание	17
Свойства математического ожидания	17
Квантиль	17
Часто встречающиеся непрерывные распределения	18
Равномерное на интервале $(a; b)$	18
<hr/>	
2024-10-18	
Экспоненциальное (показательное) распределение	18
Распределение Гаусса (Нормальное распределение)	19
Стандартное распределение Гаусса	20
Уравнение Лапласа	20
Вероятность попадания в заданный интервал	20
<hr/>	
2024-11-01	
Неравенства Чебышева	21
Частные случаи	21
Пример	21
Случайные векторы	22
Свойства функции распределения	22
Дискретные случайные векторы	22
Непрерывные случайные векторы	23
Плотность распределения	23
Свойства плотности распределения	23
<hr/>	
2024-11-08	
Математическое ожидание	24
Свойства математического ожидания	24
Ковариация	25
Свойства ковариации	25
<hr/>	
2024-11-15	
Свойства коэффициента корреляции	26
Ковариационная матрица	27
Свойства ковариационной матрицы	27
Корреляционной матрица	27
Свойства корреляционной матрицы	27
Формула свертки	28
Пример	28

2024-09-06

Введение

$$\text{Итог} = 0.1 \cdot \text{ИДЗ} + 0.15 \cdot \text{Сем} + 0.25 \cdot \text{КР} + 50 \cdot \text{Экз}$$

Нужно набрать 4 — не 3.5

По ИДЗ бывают защиты

На семинарах могут быть самостоятельные

Кибзун, Горяинова, Наумов «ТВ и МС. Базовый курс с примерами к задачам» 2013 или 2014

КР на тему «случайные события и случайные величины (одномерные)» примерно после 7-ми занятий, в начале 20-ого модуля

Экз на тему «многомерные случайные величины»

Основные понятия

Опр. Теория Вероятностей — раздел математики, изучающий математические модели массовых случайных явлений

При большом кол-ве событий величина $\frac{m}{n} \rightarrow P$ стабилизируется

$\omega_1, \dots, \omega_n$ — элементарные случайные события

Опр. Пространство Элементарных Событий (Ω) — совокупность элементарных случайных событий

Опр. Случайное событие — любое $A \subset \Omega$

Опр. Достоверное событие — событие, которое происходит в опыте всегда. Совпадает с Ω

Опр. Невозможное событие — событие, которое не происходит в опыте никогда. Является \emptyset

Операции над множествами/ событиями:

- Произведение событий $A \cdot B$ — событие из $A \cap B$
- Сумма событий $A + B$ — событие из $A \cup B$
- Разность событий $A \setminus B$
- Противоположное событие $\bar{A} = \Omega \setminus A$

Свойства операций над событиями:

- $A + A = A$
- $A \cdot A = A$
- $A \cdot \Omega = A$
- $A + \Omega = \Omega$
- $A + B = B + A$
- $A \cdot B = B \cdot A$
- $A + (B + C) = (A + B) + C$
- $\overline{A \cdot (B \cdot C)} = (\bar{A} \cdot \bar{B}) \cdot \bar{C}$
- $\overline{\bar{A}} = A$
- $\overline{A + B} = \bar{A} \cdot \bar{B}$

Опр. σ -алгебра событий класс подмножеств в \mathcal{A} на пространстве элементарных событий Ω , если:

1. $\Omega \in \mathcal{A}$

$$2. A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A}$$

$$3. \forall A_1, \dots, A_n, \dots \in \mathcal{A} \Rightarrow \sum_{i=1}^{\infty} A_i \in \mathcal{A} \wedge \prod_{i=1}^{\infty} A_i \in \mathcal{A}$$

Классическое определение вероятности

Опр. Пусть Ω содержит **конечное** число **равновозможных** **взаимоисключающих** исходов, тогда, **вероятность события** A :

$$P(A) = \frac{|A|}{|\Omega|},$$

где $|A|$ – мощность события, количество событий, входящих в A

Свойства вероятности

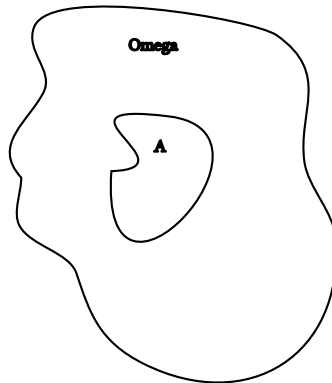
- $P(A) \in [0; 1]$
- $P(\Omega) = 1$
- Если $A \cdot B = \emptyset$, то $P(A + B) = P(A) + P(B)$

2024-09-13

Геометрическое определение вероятности

Рассматриваем подмножества на \mathbb{R}^n , которые имеют конечную меру

Пример эксперимента: попадет ли случайная точка в подмножество



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

Опр. События **несовместны** – $A \cdot B = \emptyset$

Свойства $P(A)$

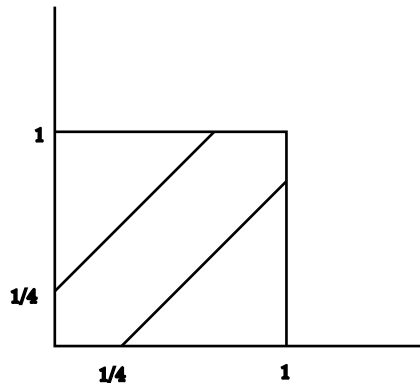
1. $P(A) \geq 0 \forall A \subset \Omega$
2. $P(\Omega) = 1$
3. если A_1 и A_2 несовместны, то $P(A_1 + A_2) = P(A_1) + P(A_2)$

Задача

x – время прихода Джульеты

y – время прихода Ромео

$$|x - y| < 14$$



$$P(\overline{A}) = \frac{\mu(\overline{A})}{\mu(\Omega)} = \frac{9}{16}$$

$$P(A) = 1 - \frac{9}{16} = \frac{7}{16}$$

Частотное (статистическое) определение

Опр. Пусть опыт проведен N раз, и событие произошло m_A раз. Тогда **частота** события A :

$$\nu(A) = \frac{m_A}{N}$$

Опр.

$$P(A) = \lim_{N \rightarrow \infty} \nu(A) = \lim_{N \rightarrow \infty} \frac{m_A}{N}$$

Аксиоматическое определение Колмагорова

Пусть \mathcal{A} — σ -алгебра событий на пространстве Ω . Числовая функция $P : \mathcal{A} \rightarrow \mathbb{R}$ — **вероятность**, если:

1. $\forall A \in \mathcal{A} : P(A) \geq 0$ — аксиома неотрицательности
2. $P(\Omega) = 1$ — условие нормировки
3. если A_1, \dots, A_n, \dots попарно несовместны, то $P\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

Число $P(A)$ называется вероятностью события A

Тройка (Ω, \mathcal{A}, P) — вероятностное пространство

Свойства $P(A)$

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

Док-во

$$\Omega = A + \overline{A}$$

$$A \cdot \overline{A} = \emptyset$$

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

$$2. P(\emptyset) = 1 - P(\Omega) = 0$$

$$3. A \subset B \Rightarrow P(A) \leq P(B)$$

Док-во

$$B = A + (B \setminus A)$$

$$P(B) = P(A + (B \setminus A)) = P(A) + \underbrace{P(B \setminus A)}_{\geq 0}$$

4. $\forall A : 0 \leq P(A) \leq 1$

5. Теорема сложения: $P(A + B) = P(A) + P(B) - P(A \cdot B)$

Док-во

$$A = A\Omega = AB + A\bar{B}$$

$$B = B\Omega = AB + \bar{A}B$$

$$A + B = \underbrace{AB + A\bar{B} + \bar{A}B}_{\text{попарно несовместны}}$$

$$P(A) = P(AB) + P(A\bar{B}) \Rightarrow P(A) - P(AB) = P(A\bar{B})$$

$$P(B) = P(AB) + P(\bar{A}B) \Rightarrow P(B) - P(AB) = P(\bar{A}B)$$

$$P(A + B) = P(AB) + P(A\bar{B}) + P(\bar{A}B) =$$

$$= P(AB) + P(A) - P(AB) + P(B) - P(AB) = P(A) + P(B) - P(AB)$$

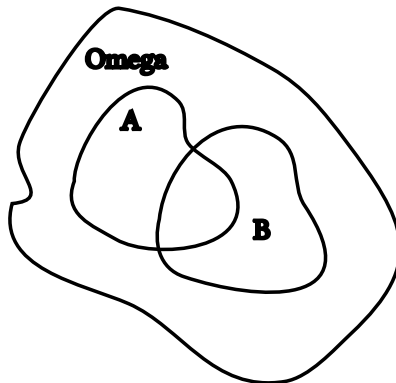
6. Обобщение теоремы сложения:

$$P\left(\underbrace{A_1 + A_2}_A + \underbrace{A_3}_B\right) = P(A) + P(B) =$$

$$= P(A_1) + P(A_2) + P(A_3) - P(A_1A_2) - P(A_1A_3) - P(A_2A_3) + P(A_1A_2A_3)$$

$$P\left(\sum_i A_i\right) = \sum_i P(A_i) - \sum_{i < j} P(A_iA_j) + \sum_{i < j < k} P(A_iA_jA_k) - \dots + (-1)^{n+1} P(A_1 \dots A_n)$$

Условная вероятность



Переходим из Ω в B

Пусть $A, B \in \Omega$ и $P(B) \neq 0$, тогда вероятность A при условии B :

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

Опр. A и B **независимые**, если $P(A|B) = P(A)$

Опр. A и B **независимые**, если $P(AB) = P(A)P(B)$

Любые несовместные события зависимы

2024-09-20

Независимость в совокупности

Опр. События A_1, \dots, A_n **независимы в совокупности**, если

$$\forall 1 \leq i_1 < i_2 < \dots < i_k \leq n : P(A_{i_1} A_{i_2} \dots) = P(A_{i_1}) P(A_{i_2}) \dots$$

- Независимы в совокупности \rightarrow независимы попарно
- Независимы все подмножества \rightarrow независимы совокупно

Пример

Дан тетраэдр. Четыре стороны покрашены в красный, синий, зеленая и все три цвета соответственно.

A_1 — выпала грань с **красным** цветом

A_2 — выпала грань с **синим** цветом

A_3 — выпала грань с **зеленым** цветом

$$P(A_1) = P(A_2) = P(A_3) = \frac{2}{4} = \frac{1}{2}$$

$$P(A_1 A_2) = P(A_1 A_3) = P(A_2 A_3) = \frac{1}{4}$$

$$P(A_1 A_2 A_3) = \frac{1}{4} \neq P(A_1) P(A_2) P(A_3)$$

Теорема умножения вероятностей

Пусть $P(A_1 A_2 \dots A_n) > 0$,

$$P\left(\overbrace{A_1 \dots A_n}^A\right) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 A_2) \dots P(A_n | A_1 \dots A_{n-1})$$

Док-во

Пусть:

$$B_{n-1} = A_1 \dots A_{n-1}$$

$$B_{n-2} = A_1 \dots A_{n-2}$$

\vdots

$$B_1 = A_1$$

Тогда

$$\begin{aligned}
 A &= B_{n-1}A_n \\
 P(A) &= P(B_{n-1}A_n) = \\
 &= P\left(\overbrace{B_{n-1}}^{B_{n-2}A_{n-1}}\right)P(A_n | B_{n-1}) = \\
 &= P(A_n | A_1 \dots A_{n-1})P(B_{n-2})P(A_{n-2} | B_{n-2}) = \dots
 \end{aligned}$$

Пример

Перестановки: МАТАН

$$\begin{aligned}
 P('M' 'A' 'T' 'A' 'H') &= \\
 &= P('M')P('A' | 'M')P('T' | 'M' 'A')P('A' | 'M' 'A' 'T')P('H' | 'M' 'A' 'T' 'A') = \\
 &= \frac{1}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot 1
 \end{aligned}$$

Биномиальная схема испытаний Бернулли

Схема испытаний, которая удовлетворяет условиям:

- Исход двоичен. Происходит A (успех) или \bar{A} (неудача)
- Всех испытания независимы в совокупности
- $p = P(A)$ не изменяется от опыта к опыту

k успехов из n испытаний:

$$P_n(k) = C_n^k p^k (1-p)^{n-k} = C_n^k p^k q^{n-k}$$

Док-во

Если все успехи в начале:

$$P\left(\underbrace{UU\dots U}_k HH\dots H\right) = p^k q^{n-k}$$

Учтем перестановки. Выберем, где места будут успехи (C_n^k способов):

$$P_n(k) = C_n^k p^k q^{n-k}$$

$$P(k_1 \leq k \leq k_2) = \sum_{i=k_1}^{k_2} C_n^i p^i q^{n-i}$$

$$1 = \sum_{k=0}^n P_n(k) = \sum_{k=0}^n C_n^k p^k q^{n-k} = (p+q)^n = 1^n = 1$$

Наиболее вероятное число успехов

По определению:

$$k_0 = \operatorname{argmax}_{1 \leq i \leq n} C_n^i p^i q^{n-i}$$

По удобному:

$$k_0 = \begin{cases} [(n+1)p] & \text{если } (n+1)p \notin \mathbb{Z} \\ (n+1)p \text{ и } (n+1)p - 1 & \text{если } (n+1)p \in \mathbb{Z} \end{cases}$$

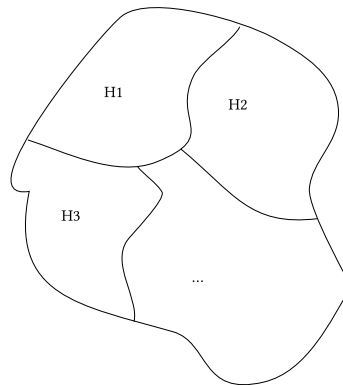
———— Формула полной вероятности ————

Опр. Пусть $H_1, \dots, H_n \in \Omega$. Если

1. $\forall i \neq j: H_i \cdot H_j = \emptyset$
2. $H_1 + \dots + H_n = \Omega$

то H_1, \dots, H_n **полная группа событий (гипотезы)**

Рис. 4. Полная группа событий (гипотезы)



Пусть $A \subset \Omega$, H_1, \dots, H_n — полная группа событий

$$\begin{aligned} P(A) &= P(A \cdot \Omega) = P(A \cdot (H_1 + \dots + H_n)) = \\ &= P(AH_1 + \dots + AH_n) \stackrel{\text{т.к. несовместны}}{\widehat{=}} P(H_1)P(A | H_1) + \dots + P(H_n)P(A | H_n) \end{aligned}$$

Пример

N — всего билетов

m — билетов студент Сидоров выучил

A — Сидорову попался счастливый билет

Иванов заходит первый. Сидоров заходит второй.

H_1 — Иванов вытащил счастливый (для Сидорова)

H_2 — Иванов вытащил **не** счастливый (для Сидорова)

$$P(H_1) = \frac{m}{N}$$

$$P(H_2) = \frac{N - m}{N}$$

$$\begin{aligned} P(A) &= P(H_1)P(A | H_1) + P(H_2)P(A | H_2) = \\ &= \frac{m}{N} \cdot \frac{m-1}{N-1} + \frac{N-m}{N} \cdot \frac{m}{N-1} \end{aligned}$$

Для гипотез:

- Априорные вероятности — знаем ещё до опыта:

$$P(H_1), \dots, P(H_n)$$

- Апостериорные вероятности — вероятности гипотез после эксперимента (когда знаем, что некоторое событие уже произошло):

$$P(H_1 | A), \dots, P(H_n | A)$$

2024-09-27

Формула Байеса

$$P(H_i | A) = \frac{P(AH_i)}{P(A)} = \frac{P(H_i)P(A | H_i)}{\sum_{j=1}^n P(H_j)P(A | H_j)}$$

$$\sum_{i=1}^n P(H_i) = 1$$

$$\sum_{i=1}^n P(H_i | A) = 1$$

Задача

Таблица 1. Условие

Завод	Процент поставленных деталей	Вероятность исправной детали
№ 1	65%	0.9
№ 2	35%	0.8

A — деталь с дефектом оказалась в самолете

H_1 — деталь взяли и 1-ого завода

$$P(H_1) = 0.65$$

$$P(A | H_1) = 0.1$$

H_2 — деталь взяли и 2-ого завода

$$P(H_2) = 0.35$$

$$P(A | H_2) = 0.2$$

$$P(A) = P(H_1)P(A | H_1) + P(H_2)P(A | H_2) = 0.65 \cdot 0.1 + 0.35 \cdot 0.2$$

$$P(H_1 | A) = \frac{P(H_1)P(A | H_1)}{P(A)} = \frac{65}{135} = 0.48$$

$$P(H_2 | A) = \frac{P(H_2)P(A | H_2)}{P(A)} = \frac{70}{135} = 0.52$$

———— Случайные величины ————

Опр. Случайная величина (СВ) — величина, которая после эксперимента принимает заранее неизвестное значение.

Числовая функция $\xi : \Omega \rightarrow \mathbb{R}$, которая удовлетворяет условию измеримости¹:

$$\forall x \in \mathbb{R} : \{\omega : \xi(\omega) \leq x\} \in \mathcal{A}$$

Таблица 2. Пример с кубиком

$$\begin{array}{cccccc} \Omega = \{ & \omega_1, & \omega_2, & \dots, & \omega_6 & \} \\ & \downarrow & \downarrow & & \downarrow & \\ \xi = \{ & 1, & 2, & \dots, & 3 & \} \end{array}$$

Опр. Функция распределения (вероятностей) случайной величины ξ называется функция

$$F_\xi(x) = P(\omega : \xi(\omega) \leq x)$$

Свойства $F(x)$:

$$1. F(+\infty) = 1$$

$$F(-\infty) = 0$$

$$\forall x : 0 \leq F(x) \leq 1$$

$$2. F(x) \text{ не убывает: } x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$$

$$3. F(x) \text{ непрерывна справа:}$$

$$F(x_0) = \lim_{\varepsilon \rightarrow 0+} F(x_0 + \varepsilon),$$

где x_0 — точка разрыва

Если некоторая $F(x)$ удовлетворяет условиям, то она является функцией распределения некоторой величины.

Случайные величины:

- Дискретные
- Непрерывные

———— Дискретные случайные величины ————

Опр. Случайную величину называют **дискретной**, если множество её возможных значений конечно или счетно.

Опр. Ряд распределения для дискретной СВ — табличка из ξ в P :

ξ	x_1	x_2	\dots	x_n
P	p_1	p_2	\dots	p_n

Пример

¹почти всегда выполняется

ξ	-1	0	2
P	0.2	0.3	0.5

$$x < -1 : F(x) = 0$$

$$F(-1) = 0.2$$

$$F(-0.5) = 0.2$$

$$F(0) = 0.2 + 0.3$$

$$F(2) = 1$$

$$x > 2 : F(x) = 1$$

Числовые характеристики дискретных случайных величин

Математическое ожидание

Опр. Математическим ожиданием E (среднее значение) дискретной СВ ξ называется число

$$E\xi = \sum_{i=1}^{\infty} x_i p_i$$

Предполагается, что ряд $\sum_{i=1}^{\infty} |x_i| p_i$ сходится

2024-10-04

Свойства математического ожидания

1. $Ec = c$
2. $E(c\xi) = cE(\xi)$
3. Если $a \leq \xi \leq b$, то $a \leq E\xi \leq b$.
4. $E(\xi_1 + \xi_2) = E(\xi_1) + E(\xi_2)$
5. Пусть $\eta = \varphi(\xi)$, где φ — детерминированная функция, тогда $E\eta = E\varphi(\xi) = \sum_{i=1}^n \varphi(x_i) p_i$

Дисперсия

Опр. Дисперсией случайной величины ξ называется число

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

Свойства дисперсии

1. $Dc = 0$
2. $D(c\xi) = c^2 D(\xi)$
3. $\forall \xi : D(\xi) \geq 0$
4. $D(\xi) = E(\xi^2 - 2\xi E(\xi) + (E\xi)^2) = E(\xi^2) - 2(E\xi)^2 + (E\xi)^2 = E\xi^2 - (E\xi)^2$ — удобная формула для вычислений
5. $D(\xi_1 + \xi_2) = E(\xi_1 + \xi_2)^2 - (E(\xi_1 + \xi_2))^2 = E(\xi_1^2) + 2E(\xi_1 \xi_2) + E(\xi_2^2) - (E\xi_1)^2 - 2E\xi_1 E\xi_2 + (E\xi_2)^2 = D\xi_1 + D\xi_2 + 2 \underbrace{(E(\xi_1 \xi_2) - E\xi_1 E\xi_2)}_{\text{ковариация}} = D\xi_1 + D\xi_2 + 2 \text{cov}(\xi_1, \xi_2)$
6. Если ξ_1 и ξ_2 независимы, то $\text{cov}(\xi_1, \xi_2) = 0 \rightarrow D(\xi_1 + \xi_2) = D\xi_1 + D\xi_2$

· Другие ·

Опр. Среднеквадратическое отклонение: $\sigma_\xi = \sqrt{D\xi}$

Опр. Начальный момент порядка k ($k = 1, 2, \dots$): $\mu_k = E\xi^k$

$$\mu_1 = E\xi$$

Опр. Центральный момент порядка k ($k = 2, 3, \dots$): $\nu_k = E(\xi - E\xi)^k$

$$\nu_2 = D\xi$$

Опр. Центрированная случайная величина: $\xi^0 = \xi - E\xi$

$$E\xi^0 = 0$$

Опр. Нормированная случайная величина: $\xi^* = \frac{\xi^0}{\sigma}$

$$D\xi^* = D\frac{\xi^0}{\sigma} = \frac{1}{\sigma^2}D\xi^0 = 1$$

—— Часто встречающиеся дискретные распределения ——

· Распределение Бернулли ·

ξ имеет распределение $\sim \text{Ber}(p), 0 < p < 1$

ξ	0	1
P	$1-p$	p

ξ^2	0	1
P	$1-p$	p

$$E\xi = 0 + 1 \cdot p = p$$

$$D\xi = E\xi^2 - (E\xi)^2 = p - p^2 = p(1-p) = pq$$

· Биномиальное распределение ·

$\xi \sim \text{Bi}(n, p), 0 < p < 1$

ξ	0	...	k	...	n
P	$C_n^k p^k q^{n-k}$

Матожидание. Способ 1:

$$\begin{aligned} E\xi &= \sum_{k=0}^n k \cdot C_n^k p^k q^{n-k} = \sum_{k=0}^n k \frac{n!}{k!(n-k)!} p^k q^{n-k} = np \sum_{k=1}^n \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} q^{n-k} = \\ &= \{i = k-1\} = np \sum_{i=0}^{n-1} \frac{(n-1)!}{i!(n-1-i)!} p^i q^{n-1-i} = np \cdot 1 = np \end{aligned}$$

Матожидание. Способ 2:

$$\xi = \xi_1 + \xi_2 + \xi_3 + \dots + \xi_n$$

$$\xi_i \sim \text{Ber}(p)$$

$$E\xi = E\left(\sum_{i=1}^n \xi_i\right) = \sum_{i=1}^n (E\xi_i) = np$$

Дисперсия. Способ 1:

$$D\xi = E\xi^2 - (E\xi)^2 = \sum_{k=0}^n k^2 \cdot C_n^k p^k q^{n-k} = \dots$$

Дисперсия. Способ 2:

$$D\xi = D\left(\sum_{i=1}^n \xi_i\right) \underset{\text{т.к. независимы}}{\equiv} \sum_{i=1}^n D(\xi_i) = npq$$

Пример

Бросаем монетку 10 раз.

$$n = 10, p = 0.5 \rightarrow \begin{cases} E\xi = 10 \cdot 0.5 = 5 \\ D\xi = 10 \cdot \frac{1}{2} \cdot \frac{1}{2} = 2.5 \end{cases}$$

Распределение Пуассона

$$\xi \sim \Pi(\lambda), \lambda > 0$$

$$\xi = \{0, 1, 2, \dots\}$$

$$P(\xi = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

Проверим условие нормировки:

$$\sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$$

Матожидание:

$$E\xi = \sum_{k=0}^{\infty} k \frac{e^{-\lambda} \lambda^k}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

Дисперсия:

$$\begin{aligned} D\xi &= E\xi^2 - (E\xi)^2 = \sum_{k=0}^{\infty} k^2 \frac{e^{-\lambda} \lambda^k}{k!} - \lambda^2 = \sum_{k=0}^{\infty} k \frac{e^{-\lambda} \lambda^k}{(k-1)!} - \lambda^2 = \\ &= \lambda \sum_{k=0}^{\infty} (k-1+1) \frac{e^{-\lambda} \lambda^{k-1}}{(k-1)!} - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda \end{aligned}$$

Теорема Пуассона Пусть проводятся испытания по схеме Бернулли, причем $n \rightarrow \infty$, $p \rightarrow 0$, $np \rightarrow \lambda$. Тогда

$$\lim_{n \rightarrow \infty} C_n^k p^k (1-p)^{n-k} = \frac{e^{-\lambda} \lambda^k}{k!}$$

Док-во

$$\begin{aligned}
\lim_{n \rightarrow \infty} C_n^k p^k (1-p)^{n-k} &= \lim_{n \rightarrow \infty} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \\
&= \lim_{n \rightarrow \infty} \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \\
&= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{n^k} \frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^k} = 1 \cdot \frac{e^{-\lambda}}{1} = e^{-\lambda}
\end{aligned}$$

Погрешность при замене Бернулли на Пуассона:

$$\left| C_n^k p^k q^{n-k} - \frac{e^{-np} (np)^k}{k!} \right| \leq np^2$$

Геометрическое распределение

$$\xi \sim G(p), 0 < p < 1$$

$$\xi = \{1, 2, \dots\}$$

$$P(\xi = k) = q^{k-1} p$$

Смысл: Испытание с двумя исходами. Останавливаемся, когда произошел первый успех.

Матожидание:

$$\begin{aligned}
E\xi &= \sum_{k=1}^{\infty} k q^{k-1} p = p \sum_{k=1}^{\infty} k q^{k-1} = p \sum_{k=1}^{\infty} (q^k)' = \\
&= p \left(\sum_{k=1}^{\infty} q^k \right)' = p \left(\frac{q}{1-q} \right)' = \frac{p}{(1-q)^2} = \frac{1}{p}
\end{aligned}$$

Дисперсия:

$$D\xi = \frac{q}{p^2}$$

Пример

Студент знает 80% материала. Его спрашивают, пока не завалят.

$$\xi \sim G(0.2)$$

2024-10-11

Непрерывные случайные величины

Нельзя задать рядом распределения

Можно задать функцией распределения

Опр. Плотность $f_{\xi}(x)$ случайной величины — такая неотрицательная кусочная функция, что

$$\forall x \in R : F_{\xi}(x) = \int_{-\infty}^x f_{\xi}(t) dt$$

Опр. Случайные величины, для которых определена плотность определения, будем называть **непрерывными**².

Канторова лестница

Пример функции, которая непрерывна, но плотности не имеет

$$F(x) = \begin{cases} 0, & x \leq 0 \\ \frac{1}{2}F(3x), & 0 \leq x \leq \frac{1}{3} \\ \frac{1}{2}, & \frac{1}{3} \leq x \leq \frac{2}{3} \\ \frac{1}{2} + \frac{1}{2}F(3x-2), & \frac{2}{3} \leq x \leq 1 \\ 1, & x \geq 1 \end{cases}$$

В точках дифференцируемости функции $F(x)$: $f(x) = F'(x)$

С какой вероятностью будет принято какое-то конкретное значение x :

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{P(x < \xi \leq x + \Delta x)}{\Delta x}$$

$$f(x)\Delta x \underset{\Delta x \rightarrow 0}{\rightleftharpoons} P(x < \xi \leq x + \Delta x)$$

Итого, ответ 0

— Свойства плотности распределения —

- $\forall x : f(x) \geq 0$
- $\int_{-\infty}^{\infty} f(x) dx = F(+\infty) = 1$
- $\int_{x_1}^{x_2} f(x) dx = F(x_2) - F(x_1) = P(x_1 < \xi \leq x_2)$
- Пусть
 - ξ имеет плотность распределения $f_{\xi}(x)$
 - $\eta = \varphi(\xi)$, где φ — монотонная, дифференцируемая, детерминированная функция

Тогда,

$$f_{\eta}(y) = f_{\xi}(\varphi^{-1}(y)) |(\varphi^{-1}(y))'|$$

Док-во

1. Пусть $\varphi(x)$ — монотонно возрастающая

$$F_{\eta}(y) = P(\eta \leq y) = P(\varphi(\xi) \leq y) = P(\xi \leq \varphi^{-1}(y)) = F_{\xi}(\varphi^{-1}(y))$$

²На самом деле есть три вида величин:

1. Дискретные
2. Сингулярные
3. Абсолютно непрерывные

Но мы рассматриваем только два вида

$$f_{\eta}(y) = (F_{\eta}(y))' = f_{\xi}(\varphi^{-1}(y))(\varphi^{-1}(y))'$$

2. Пусть $\varphi(x)$ — монотонно убывающая

$$F_{\eta}(y) = P(\eta \leq y) = P(\varphi(\xi) \leq y) = P(\xi \geq \varphi^{-1}(y)) = 1 - F_{\xi}(\varphi^{-1}(y))$$

$$f_{\eta}(y) = (F_{\eta}(y))' = -f_{\xi}(\varphi^{-1}(y)) \underbrace{(\varphi^{-1}(y))'}_{<0}$$

- Если функция не монотонная, то нужно разделить её на интервалы монотонности и применить прошлый пункт

—— Числовые характеристики ——

· Математическое ожидание ·

Опр. Математическим ожиданием непрерывной случайной величины ξ называется число

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

если интеграл сходится абсолютно:

$$E\xi = \int_{-\infty}^{\infty} |x| f_{\xi}(x) dx.$$

Для бесконечностей:

- Если $f(x) > 0$ только при $x > 0$ и $\int_{-\infty}^{\infty} x f_{\xi}(x) dx$ расходится, то $E\xi = +\infty$
- Если $f(x) > 0$ только при $x < 0$ и $\int_{-\infty}^{\infty} x f_{\xi}(x) dx$ расходится, то $E\xi = -\infty$

Свойства математического ожидания

1. $Ec = c$
2. $E(c\xi) = cE(\xi)$
3. Если $a \leq \xi \leq b$, то $a \leq E\xi \leq b$.
4. $E(\xi_1 + \xi_2) = E(\xi_1) + E(\xi_2)$
5. Пусть $\eta = \varphi(\xi)$, где φ — детерминированная функция, тогда $E\eta = \int_{-\infty}^{\infty} \varphi(x) f_{\xi}(x) dx$

—— Квантиль ——

Опр. Число z_{γ} , $0 < \gamma < 1$ называется **γ -квантилью** непрерывного строго монотонного распределения $F_{\xi}(x)$, если $\underbrace{F_{\xi}(z_{\gamma})}_{=P(\xi \leq z_{\gamma})} = \gamma$

Для непрерывного распределения верно:

$$\int_{-\infty}^{z_{\gamma}} f_{\xi}(x) dx = \gamma$$

Для дискретных величин в качестве квантили берут минимальное подходящее число:

$$z_{\gamma} = \min\{x : F(x) \geq \gamma\}$$

Если $\forall x : f(-x) = f(x)$, то $z_{\gamma} = -z_{1-\gamma}$.

Опр. Квантиль уровня 0.5 называется **медианой**.

Опр. Квантили уровня 0.25 и 0.75 называются **нижним и верхним квартилем**.

—— Часто встречающиеся непрерывные распределения ——

Равномерное на интервале $(a; b)$

$$\xi \sim R(a, b)$$

$$f(x) = \begin{cases} 0, & x \notin (a, b) \\ \frac{1}{b-a}, & x \in (a, b) \end{cases}$$

$$E\xi = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{a+b}{2}$$

$$D\xi = E\xi^2 - (E\xi)^2 = \int_a^b \frac{x^2}{b-a} dx - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

$$F(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0, & x \leq a \\ \int_{-\infty}^a 0 dt + \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}, & x \in (a, b) \\ \int_{-\infty}^a 0 dt + \underbrace{\int_a^x \frac{1}{b-a} dt}_{=1} + \underbrace{\int_b^x 0 dt}_{=0} = 1, & x \geq b \end{cases}$$

2024-10-18

Пусть есть генератор случайной величины $\xi \sim R(0; 1)$.

Хотим получить случайную величину $\eta \sim F_\eta(y)$

$$\eta = F_\eta^{-1}(\xi)$$

Обратная функция всегда существует т.к. F возрастает

Экспоненциальное (показательное) распределение

$$\xi \sim E(\lambda), \lambda > 0$$

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

$$\begin{aligned} E\xi &= \int_0^\infty x \lambda e^{-\lambda x} dx = \{\text{по частям}\} = \\ &= \underbrace{-e(-\lambda x)}_0 \Big|_0^\infty + \frac{1}{\lambda} \underbrace{\int_0^\infty \lambda e^{-\lambda x} dx}_{=1 \text{ (из усл. нормировки)}} = \frac{1}{\lambda} \end{aligned}$$

$$D\xi = E\xi^2 - \frac{1}{\lambda^2} = \int_0^\infty x^2 \lambda e^{-\lambda x} dx - \frac{1}{\lambda^2} = \dots = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

$$F(x) = \begin{cases} 0, & x \leq 0 \\ \int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda t} \Big|_0^x = 1 - e^{-\lambda x}, & x \geq 0 \end{cases}$$

Характеристическое свойство экспоненциального распределения

Пусть:

$$\xi \sim E(\lambda)$$

Тогда:

$$\forall t > 0 \forall \tau > 0 : P(\xi > t + \tau \mid \xi > t) = P(\xi > \tau)$$

Док-во

$$\begin{aligned} P(\xi > t + \tau \mid \xi > t) &= \frac{P(\xi > t + \tau)}{P(\xi > t)} = \frac{1 - F(t + \tau)}{1 - F(t)} = \\ &= \frac{e^{-\lambda(t+\tau)}}{e^{-\lambda t}} = e^{-\lambda \tau} = 1 - F(\tau) = P(\xi > \tau) \end{aligned}$$

Из непрерывных только экспоненциальное обладает этим свойством. Из дискретных — только геометрическое.

Пример применения: теория массового обслуживания (например, обслуживание клиентов, обработка интернет-запросов)

Распределение Гаусса (Нормальное распределение)

$$\xi \sim N(m, \sigma^2)$$

m — математическое ожидание

σ^2 — дисперсия

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$$

Симметрична относительно прямой $x = m$

$$f_{\max} = f(m) = \frac{1}{\sqrt{2\pi}\sigma}$$

$$\lim_{x \rightarrow \pm\infty} f(x) = 0$$

При увеличении σ график становится шире, но ниже.

Интеграл от $f(x)$ не берется, $F(x)$ записать нельзя, поэтому используют таблички.

$$\begin{aligned} E\xi &= \int_{-\infty}^{+\infty} x f(x) dx = \int_{-\infty}^{+\infty} \frac{x}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx = \int_{-\infty}^{+\infty} \frac{x-m+m}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx = \\ &= \underbrace{\int_{-\infty}^{+\infty} \frac{x-m}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx}_{\text{нечетная}} + m \underbrace{\int_{-\infty}^{+\infty} f(x) dx}_{=1} = m \end{aligned}$$

$$D\xi = \int_{-\infty}^{+\infty} \frac{(x-m)^2}{\sqrt{1\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) = \left\{y = \frac{x-m}{\sigma}\right\} =$$

$$= 2\sigma^2 \int_0^{+\infty} \frac{y^2}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy = 2\sigma^2 \left(\underbrace{\frac{y}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right)}_{=0} \Big|_0^{+\infty} - \underbrace{\int_0^{\infty} f(y) dy}_{=\frac{1}{2}} \right) = \sigma^2$$

Стандартное распределение Гаусса

Чтобы использовать таблички, используем **стандартное** гауссовское распределение.

$$\xi \sim N(0; 1)$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

Рис. 5. Купюра с изображением гауссовского распределения



Уравнение Лапласа

$$\Phi_0(x) = \int_0^x \frac{\exp\left(-\frac{t^2}{2}\right)}{\sqrt{2\pi}} dt$$

Очень быстро стремится к нулю (для числа пять почти равна нулю (до 7-ого знака))

$$\Phi_0(+\infty) = \frac{1}{2}$$

$$\Phi_0(-x) = -\Phi_0(x)$$

$$-\frac{1}{2} \leq \Phi_0(x) \leq \frac{1}{2}$$

$$\Phi(x) := F(x) = \Phi_0(x) + \frac{1}{2}$$

Вероятность попадания в заданный интервал

Хотим посчитать вероятность попадания ξ в интервал (α, β) .

Общий случай:

$$P(\alpha < \xi < \beta) = \left\{ y = \frac{x-m}{\sigma}; dy = \frac{1}{\sigma} dx \right\} = \int_{\frac{\alpha-m}{\sigma}}^{\frac{\beta-m}{\sigma}} \frac{\exp\left(-\frac{y^2}{2}\right)}{\sqrt{2\pi}} = \Phi_0\left(\frac{\beta-m}{\sigma}\right) - \Phi_0\left(\frac{\alpha-m}{\sigma}\right)$$

Если интервал симметричен относительно m

$$P(|\xi - m| < \delta) = \Phi_0\left(\frac{\delta}{\sigma}\right) - \Phi_0\left(-\frac{\delta}{\sigma}\right) = 2\Phi_0\left(\frac{\delta}{\sigma}\right)$$

Правило трех сигм:

$$P(|\xi - m| < 3\sigma) = 2\Phi_0\left(\frac{3\sigma}{\sigma}\right) = 2\Phi_0(3) \approx 0.997$$

Т.е. почти все значения лежат в промежутке $(-3\sigma, 3\sigma)$.

2024-11-01

Неравенства Чебышева

Пусть

$$E|\xi|^r < \infty,$$

Тогда

$$\forall \varepsilon > 0 : P(|\xi| \geq \varepsilon) \leq \frac{E|\xi|^r}{\varepsilon^r}.$$

Часто дает очень грубую оценку

Док-во

$$E|r| = \int_{-\infty}^{+\infty} |x|^2 f(x) dx \geq \int_{|x| \geq \varepsilon} |x|^2 f(x) dx \geq \varepsilon^r \int_{|x| \geq \varepsilon} f(x) dx = \varepsilon^r P(|\xi| \geq \varepsilon)$$

Частные случаи

- Неравенство Маркова: $r = 1$ и $P(\xi \geq 0) = 1$

$$P(\xi \geq \varepsilon) \leq \frac{E(\xi)}{\varepsilon}$$

- $r = 2$:

$$P(|\xi| \geq \varepsilon) \leq \frac{E\xi^2}{\varepsilon^2}$$

- $r = 2$ и $\eta = E\xi$:

$$P(|\xi - E\xi| \geq \varepsilon) \leq \frac{D}{\varepsilon^2}$$

Пример

ξ — расход электроэнергии $E\xi = 4000 \frac{\text{Кв}}{\text{ч}}$

Оценить, что в какой-то день $P(\xi \geq 10000)$

$$r = 1$$

$$P(\xi \geq 0) = 1$$

$$P(\xi \geq 10000) = \frac{4000}{10000} = 0.4$$

————— Случайные векторы —————

Опр. Вектор $\xi = (\xi_1, \dots, \xi_n)$, где ξ_1, \dots, ξ_n — случайные величины, называется **случайным вектором**.

Случайные вектора нужны, так как случайные величины обычно полезно рассматривать в совокупности.

Опр. **Функцией распределения случайного вектора** $\xi = (\xi_1, \dots, \xi_n)$ называется функция

$$F_\xi(x_1, \dots, x_n) = P(\xi_1 \leq x_1, \dots, \xi_n \leq x_n) := P(\xi_1 \leq x_1 \wedge \dots \wedge \xi_n \leq x_n)$$

Пусть $n = 2 : F_\xi(xy) = P(\xi_1 \leq x, \xi_2 \leq y)$

————— Свойства функции распределения —————

На примере $n = 2$

1. $\forall (x, y) \in \mathbb{R}^2 : 0 \leq F(x, y) \leq 1$
2. $F(-\infty, -\infty) = F(-\infty, y) = F(x, -\infty) = 0$
3. $F(+\infty, +\infty) = 1$
4. $F_\xi(+\infty, y) = F_{\xi_2}(y)$
 $F_\xi(x, +\infty) = F_{\xi_1}(x)$
5. $P(a_1 \leq \xi_1 \leq a_2, b_1 \leq \xi_2 \leq b_2) = F(a_2, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, b_1)$
6. $F(x, y)$ монотонно не убывает по каждому аргументу

Док-во

$$F(x + \Delta x, y) = P(\xi_1 \leq x + \Delta x, \xi_2 \leq y) = \underbrace{P(\xi_1 \leq x, \xi_2 \leq y)}_{=F(x,y)} + \underbrace{P(x \leq \xi_1 \leq x + \Delta x, \xi_2 \leq y)}_{\geq 0}$$

Опр. **Частное/ маргинальное распределение** — распределение одной из компонент вектора

Опр. Компоненты ξ_1, ξ_2 случайного вектора $\xi = (\xi_1, \xi_2)$ называются **независимыми**, если

$$F_\xi(x, y) = F_{\xi_1}(x) \cdot F_{\xi_2}(y)$$

————— Дискретные случайные векторы —————

Опр. Случайный вектор $\xi = (\xi_1, \xi_2)$ называется **дискретным**, если ξ_1 и ξ_2 — дискретные случайные величины.

Табличный способ задания

	y_1	\dots	y_k
x_1	p_{11}	\dots	p_{1k}
\vdots	\vdots	\ddots	\vdots
x_n	p_{n1}	\dots	p_{nk}

$$P_{ij} = P(\xi_1 = x_i, \xi_2 = y_j)$$

$$P_{i\cdot} := \sum_{j=1}^k p_{ij}$$

$$P(\xi_1 = x_1) = P_{1\cdot}$$

$$P(\xi_2 = y_1) = P_{\cdot 1}$$

Опр. Компоненты ξ_1, ξ_2 дискретного вектора $\xi = (\xi_1, \xi_2)$ **независимы**, если $\forall i, j : P_{ij} = p_{i\cdot} p_{\cdot j}$

Непрерывные случайные векторы

Плотность распределения

Опр. Неотрицательная кусочно непрерывная функция $f_\xi(x, y)$, такая что

$$F_\xi(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(t_1, t_2) dt_2 dt_1$$

называется **плотностью распределения** $\xi = (\xi_1, \xi_2)$.

В точках, где $F_\xi(x, y)$ дифференцируема:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

Свойства плотности распределения

- $\forall x, y : f(x, y) \geq 0$
- $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = F(+\infty, +\infty) = 1$
- $\int_{a_1}^{a_2} \int_{b_1}^{b_2} f(x, y) dx dy = F(a_2, b_2) - F(a_2, b_1) - (F(a_1, b_2) - F(a_1, b_1)) = P(a_1 \leq \xi_1 \leq a_2, b_1 \leq \xi_2 \leq b_2)$
- $P(\xi \in D) = \iint_D f(x, y) dx dy$

$$F_{\xi_1}(x) = F_\xi(x, +\infty) = \int_{-\infty}^x \int_{-\infty}^{+\infty} f(t_1, t_2) dt_2 dt_1$$

$$f_{\xi_1}(x) = \frac{d}{dx} F_{\xi_1}(x) = \frac{d}{dx} \int_{-\infty}^x \int_{-\infty}^{+\infty} f(t_1, t_2) dt_2 dt_1 = \int_{-\infty}^{+\infty} f(x, t_2) dt_2 = \int_{-\infty}^{+\infty} f(x, y) dy$$

2024-11-08

Пусть $\xi = (\xi_1, \xi_2)$ — непрерывный случайный вектор. Тогда

ξ_1 и ξ_2 независимы $\Leftrightarrow f_\xi(x, y) = f_{\xi_1}(x) f_{\xi_2}(y)$

Док-во

• (\rightarrow)

$$f_\xi(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = \frac{\partial^2 F_{\xi_1}(x) F_{\xi_2}(y)}{\partial x \partial y} = \frac{dF_{\xi_1}(x)}{dx} \frac{dF_{\xi_2}(y)}{dy} = f_{\xi_1}(x) f_{\xi_2}(y)$$

• (\leftarrow)

$$F_{\xi}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{\xi}(t, s) ds dt = \int_{-\infty}^x \int_{-\infty}^y f_{\xi_1}(t) f_{\xi_2}(s) ds dt = F_{\xi_1}(x) F_{\xi_2}(y)$$

Опр. Случайный вектор $\xi = (\xi_1, \dots, \xi_n)$ имеет равномерное распределение в области $D \in \mathbb{R}^n$, если

$$f_{\xi}(x_1, \dots, x_n) = \begin{cases} c, & \text{if } (x_1, \dots, x_n) \in D \\ 0, & \text{oth} \end{cases}$$

При $n = 2 : c = \frac{1}{S_D}$

Пример 1

Пусть $\xi = (\xi_1, \xi_2)$ распределен равномерно в прямоугольнике с углами в $(0, 0), (1, 1)$. Хотим проверить [не]зависимость компонент.

$$f_{\xi}(x, y) = \begin{cases} 1, & \text{if } x \in (0, 1) \wedge y \in (0, 1) \\ 0, & \text{oth} \end{cases}$$

$$f_{\xi_1}(x) = \int_{-\infty}^{+\infty} f_{\xi}(x, y) dy = \begin{cases} \int_0^1 1 dy = 1, & \text{if } x \in (0, 1) \\ 0, & \text{if } x \notin (0, 1) \end{cases}$$

$$f_{\xi_2}(y) = \int_{-\infty}^{+\infty} f_{\xi}(x, y) dx = \begin{cases} \int_0^1 1 dx = 1, & \text{if } y \in (0, 1) \\ 0, & \text{if } y \notin (0, 1) \end{cases}$$

Т.к. $f_{\xi}(x, y) = f_{\xi_1}(x) \cdot f_{\xi_2}(y)$, то компоненты независимы

Пример 2

Пусть $\xi = (\xi_1, \xi_2)$ распределен равномерно в круге $c = (0, 0), r = r$.

$$f_{\xi}(x, y) = \begin{cases} \frac{1}{\pi r^2}, & \text{if } x^2 + y^2 \leq R^2 \\ 0, & \text{oth} \end{cases}$$

$$f_{\xi_1}(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \frac{1}{\pi r^2} dy = \begin{cases} \frac{2\sqrt{R^2-x^2}}{\pi r^2}, & \text{if } |x| \leq r \\ 0, & \text{oth} \end{cases}$$

$$f_{\xi_2}(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_{-\sqrt{R^2-y^2}}^{\sqrt{R^2-y^2}} \frac{1}{\pi r^2} dx = \begin{cases} \frac{2\sqrt{R^2-y^2}}{\pi r^2}, & \text{if } |y| \leq r \\ 0, & \text{oth} \end{cases}$$

Т.к. $f_{\xi}(x, y) \neq f_{\xi_1}(x) f_{\xi_2}(y)$, то ξ_1 и ξ_2 — зависимы

Математическое ожидание

Опр. Математическим ожиданием вектора $\xi = (\xi_1, \xi_2)$ называется вектор $E\xi = (m_1, \dots, m_n)$, где $m_i = E\xi_i$.

Свойства математического ожидания

- $E(\xi_1 + \xi_2) = E\xi_1 + E\xi_2$

Док-во

Для дискретного случая (для непрерывного аналогично):

$$\begin{aligned} E\left(\underbrace{\xi_1 + \xi_2}_{\eta}\right) &= \sum_i \sum_j (x_i + y_j) p_{ij} = \sum_i \sum_j x_i p_{ij} + \sum_i \sum_j y_j p_{ij} = \\ &= \sum_i x_i p_{i\cdot} + \sum_j y_j p_{\cdot j} = E\xi_1 + E\xi_2 \end{aligned}$$

- Если ξ_1 и ξ_2 независимы, то $E(\xi_1 \xi_2) = E\xi_1 \cdot E\xi_2$

Док-во

Для непрерывного случая (для дискретного аналогично):

$$\begin{aligned} E\xi_1 \xi_2 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f(x, y) dx dy \underset{\text{независимы}}{=} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f_{\xi_1}(x) f_{\xi_2}(y) dx dy = \\ &= \int_{-\infty}^{+\infty} x f_{\xi_1}(x) \cdot \int_{-\infty}^{+\infty} y f_{\xi_2}(y) dy = E\xi_1 E\xi_2 \end{aligned}$$

— Ковариация —

Опр. Ковариация³ ξ_1 и ξ_2 ($\text{cov}(\xi_1, \xi_2)$ или $k_{\xi_1 \xi_2}$):

$$k_{\xi_1 \xi_2} = E(\xi_1 - E\xi_1)(\xi_2 - E\xi_2) = E\xi_1^0 \xi_2^0$$

Опр. Коэффициентом корреляции ξ_1 и ξ_2 называется

$$\rho_{\xi_1 \xi_2} = \frac{k_{\xi_1 \xi_2}}{\sqrt{D\xi_1 D\xi_2}} = k_{\xi_1^* \xi_2^*}$$

Опр. Величины ξ_1 и ξ_2 называются **некоррелированными**, если $\rho_{\xi_1 \xi_2} = 0$

Опр. Величины ξ_1 и ξ_2 называются **положительно коррелированными**, если $\rho_{\xi_1 \xi_2} > 0$

Опр. Величины ξ_1 и ξ_2 называются **отрицательно коррелированными**, если $\rho_{\xi_1 \xi_2} < 0$

· Свойства ковариации ·

- $\text{cov}(\xi, \xi) = D\xi$
 - $D(\xi + \eta) = D\xi + D\eta + 2 \text{cov}(\xi, \eta)$
 - $\text{cov}(\xi, \eta) = \text{cov}(\eta, \xi)$
 - $\text{cov}(\xi, \eta) = E(\xi - E\xi)(\eta - E\eta) = \dots = E\xi\eta - E\xi E\eta$
 - $\text{cov}(a\xi + b, c\eta + d) = ac \text{cov}(\xi, \eta)$
 - $|\rho_{\xi\eta}| = \frac{|\text{cov}(\eta, \xi)|}{\sigma_\xi \sigma_\eta} \leq 1$ — коэффициент корреляции
- $$|\text{cov}(\eta, \xi)| \leq \sigma_\xi \sigma_\eta$$

Док-во

³от «совместная изменяемость»

$$\begin{aligned}
 0 \leq D(\xi^* + \eta^*) &= D\xi^* + D\eta^* + 2 \operatorname{cov}(\xi^*, \eta^*) = 1 + 1 + 2\rho_{\xi\eta} \\
 &\Rightarrow \rho_{\xi\eta} \geq -1 \\
 0 \leq D(\xi^* - \eta^*) &= D\xi^* + D\eta^* - 2 \operatorname{cov}(\xi^*, \eta^*) = 1 + 1 - 2\rho_{\xi\eta} \\
 &\Rightarrow \rho_{\xi\eta} \leq 1 \\
 -1 \leq \rho_{\xi\eta} &\leq 1
 \end{aligned}$$

- Если ξ и η независимы и с конечными дисперсиями, то $\operatorname{cov}(\xi, \eta) = 0$

Док-во

$$\begin{aligned}
 \operatorname{cov}(\xi, \eta) &= E\xi^0\eta^0 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - m_\xi)(y - m_\eta) f(x, y) dx dy = \\
 &\stackrel{\text{независимы}}{=} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - m_\xi) f_\xi(x) (y - m_\eta) f_\eta(y) dx dy = \\
 &= \int_{-\infty}^{+\infty} (x - m_\xi) f_\xi dx \cdot \int_{-\infty}^{+\infty} (y - m_\eta) f_\eta dy = 0 \cdot 0 = 0
 \end{aligned}$$

2024-11-15

Свойства коэффициента корреляции

- $\rho_{\xi\eta} \leq 1$
- $\rho_{\xi\xi} = 1$
- Если $\eta = a\xi + b, a \neq 0$, то

$$\rho_{\xi\eta} = \begin{cases} 1, & \text{if } a > 0 \\ -1, & \text{if } a < 0 \end{cases}$$

Док-во

$$\begin{aligned}
 \rho_{\xi\eta} &= \frac{\operatorname{cov}(\xi, \eta)}{\sqrt{D\xi D\eta}} = \frac{E\xi\eta - E\xi E\eta}{\sqrt{D\xi D\eta}} = \frac{E\xi(a\xi + b) - E\xi E(a\xi + b)}{\sqrt{D\xi D(a\xi + b)}} = \\
 &= \frac{aE\xi^2 + bE\xi - a(E\xi)^2 - bE\xi}{\sqrt{a^2(D\xi)^2}} = \frac{a(E\xi^2 - (E\xi)^2)}{|a|D\xi} = \frac{aD\xi}{|a|D\xi}
 \end{aligned}$$

- Пусть $|\rho_{\xi\eta}| = 1 \Rightarrow \eta = a\xi + b$

Док-во

- Пусть $\rho_{\xi\eta} = 1$:

$$\begin{aligned}
 D(\xi^* - \eta^*) &= D\xi^* + D\eta^* + 2 \operatorname{cov}(\xi^*, -\eta^*) = \\
 &= D\xi^* + D\eta^* - 2 \operatorname{cov}(\xi^*, \eta^*) = 2 - 2\rho_{\xi\eta} = 2(1 - \rho_{\xi\eta}) = 0
 \end{aligned}$$

$$D(\xi^* - \eta^*) = 0$$

$$\xi^* - \eta^* = c$$

$$\frac{\xi - m_\xi}{\sigma_\xi} - \frac{\eta - m_\eta}{\sigma_\eta} = c$$

► Пусть $\rho_{\xi\eta} = -1$:

$$D(\xi^* + \eta^*) = D\xi^* + D\eta^* + 2 \operatorname{cov}(\xi^*, \eta^*) = 2(1 + \rho_{\xi\eta}) = 0$$

Пример: зависимы, но не коррелированы

$$x \sim R(-a, a)$$

$$n = \xi^2$$

$$\operatorname{cov}(\xi, \eta) = E\xi\eta - \underbrace{E\xi E\eta}_{=0} = E\xi^3 = \int_{-a}^a \xi^3 \cdot \frac{1}{2a} dx = 0$$

Так получается, потому что ковариация «отлавливает» только линейные зависимости

— Ковариационная матрица —

Опр. Ковариационной матрицей вектора $\xi = (\xi_1, \dots, \xi_n)$ называется матрица $K_\xi = (k_{ij})$, где $k_{ij} = \operatorname{cov}(\xi_i, \xi_j)$

Свойства ковариационной матрицы

- $k_{ij} = k_{ji}$
- $k_{ii} = D\xi_i$
- K_ξ — неотрицательно определенная:

$$\forall \lambda_1, \dots, \lambda_n \in \mathbb{R} : \sum_i \sum_j \lambda_i \lambda_j k_{ij} \geq 0$$

Док-во

$$0 \leq E \left(\sum_{i=1}^n \lambda_i \xi_i^0 \right)^2 = E \left(\sum_i \sum_j \lambda_i \lambda_j \xi_i^0 \xi_j^0 \right) = \lambda_i \lambda_j \sum_i \sum_j E(\xi_i^0 \xi_j^0) = \lambda_i \lambda_j \sum_i \sum_j k_{ij}$$

— Корреляционной матрица —

Опр. Корреляционной матрицей вектора $\xi = (\xi_1, \dots, \xi_n)$ называется матрица $R_\xi = (\rho_{ij})$, где $\rho_{ij} = \rho(\xi_i, \xi_j)$

Свойства корреляционной матрицы

- $\rho_{ij} = \rho_{ji}$
- $\rho_{ii} = 1$
- R_ξ — неотрицательно определенная

По ковариационной матрице можно построить корреляционную, а наоборот — не можем

—— Формула свертки ——

Пусть ξ_1 и ξ_2 — независимы, $\xi_1 \sim f_{\xi_1}(x)$, $\xi_2 \sim f_{\xi_2}(y)$, $\eta = \xi_1 + \xi_2$. $f_\eta(z) = ?$

$$F_{\eta(z)} = P(\eta \leq z) = P(\xi_1 + \xi_2 \leq z) = \iint_{x+y \leq z} f_{\xi_1}(x) f_{\xi_2}(y) dy dx = \int_{-\infty}^{+\infty} \int_{-\infty}^{z-x} f_{\xi_1}(x) f_{\xi_2}(y) dy dx$$

$$f_\eta(z) = \frac{d}{dz} F_\eta(z) = \int_{-\infty}^{+\infty} f_{\xi_1}(x) f_{\xi_2}(z-x) dx$$

· Пример ·

$$\xi_1 \sim N(0, 1); \xi_2 \sim N(0, 1)$$

ξ_1, ξ_2 — независимы

$$\eta = \xi_1 + \xi_2$$

$$\begin{aligned} f_\eta(z) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-x)^2}{2}} dx = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2} - \left(\frac{z^2}{2} - \frac{2xz}{2} + \frac{x^2}{2}\right)} dx = \\ &= \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}\sqrt{2}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sqrt{\frac{1}{2}}} e^{-\left(x - \frac{z}{2}\right)^2} dx = \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}\sqrt{2}} \end{aligned}$$

Снова получили гауссово распределение

Пусть ξ_1, \dots, ξ_n — независимые случайные величины такие, что $\xi_i \sim N(m_i, \sigma_i^2)$, то

$$(\eta = \xi_1 + \dots + \xi_n) \sim N(m_\eta, D_\eta),$$

где

$$m_\eta = \sum m_i$$

$$D_\eta = \sum D_i$$