

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

ОСНОВИ СТАТИСТИЧНОЇ ОБРОБКИ ДАНИХ: КОНСПЕКТ ЛЕКЦІЙ

ЙМОВІРНІСТЬ І СТАТИСТИКА

*Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського
як навчальний посібник для здобувачів ступеня бакалавра
за освітньо-професійною програмою “Мікро- та нанoeлектроніка”
за спеціальністю 153 “Мікро-та наносистемна техніка”
галузі знань 15 “Автоматизація та приладобудування”*

Київ
КПІ ім. Ігоря Сікорського
2021

Основи статистичної обробки даних: Ймовірність і статистика [Електронний ресурс] : навч. посіб. для студ. спеціальності 153 “Мікро-та наносистемна техніка” / КПІ ім. Ігоря Сікорського ; уклад.: Д.В.Корольок. – Електронні текстові дані (1 файл: 1,98 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2021. – 78 с.

Гриф надано Методичною радою КПІ ім. Ігоря Сікорського (протокол № 8 від 28.06.2021 р.)
за поданням Вченої ради Факультету електроніки (протокол № 06 від 22.06.2021 р.)

Електронне мережне навчальне видання

ОСНОВИ СТАТИСТИЧНОЇ ОБРОБКИ ДАНИХ КОНСПЕКТ ЛЕКЦІЙ

ЙМОВІРНІСТЬ І СТАТИСТИКА

Укладач: *Корольок Дмитро Володимирович, д.ф.-м.н., с.н.с.*

Відповідальний
редактор *Орлов А.Т., канд.т.н., професор*

Рецензент *Клесов І.І., доктор фіз.-мат. наук, професор*

Навчальний посібник присвячений основам теорії ймовірностей і математичної статистики у мікроелектроніці і містить матеріали, які читалися студентам бакалаврам.

Підготовка студентів за спеціальністю 153 "Мікро-та наносистемна техніка" передбачає знайомство майбутніх спеціалістів з потужними методами та техніками, які широко і успішно застосовуються як в експериментальних вимірюваннях, так і в математичному моделюванні реальних фізичних процесів.

Навчальний посібник розглядає базові поняття як теорії ймовірностей, так і теорії випадкових процесів, а також відповідні інверсні задачі математичної статистики – оцінки невідомих параметрів випадкових вибірок і перевірки статистичних гіпотез.

© КПІ ім. Ігоря Сікорського, 2021

Випадкові події і процеси

В інженерному досвіді при вивченні систем і процесів часто доводиться зустрічатися з особливого типу явищами, які прийнято називати випадковими.

Для них характерна велика, порівняно з іншими, ступінь невизначеності, непередбачуваності. Випадкове явище - це таке явище, яке при неодноразовому відтворенні одного і того ж досвіду (випробування, експерименту) протікає щораз дещо по-іншому.

Наведемо кілька прикладів випадкових явищ.

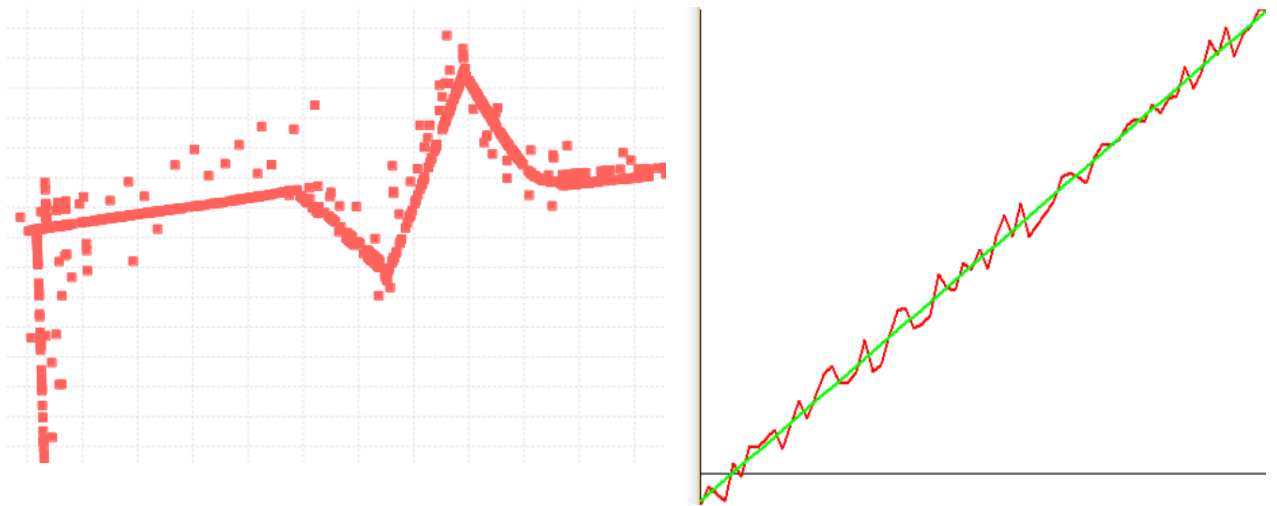
Прогноз погоди



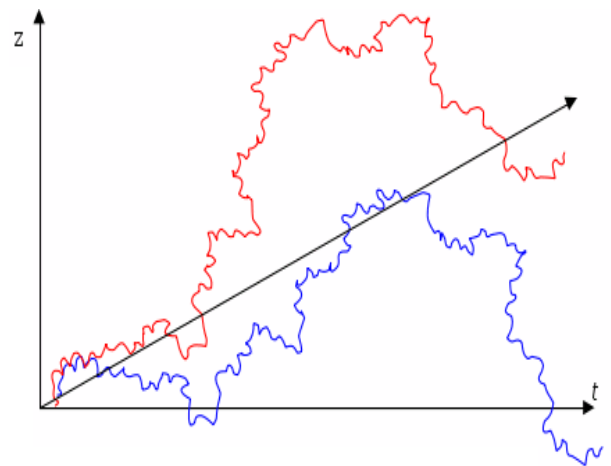
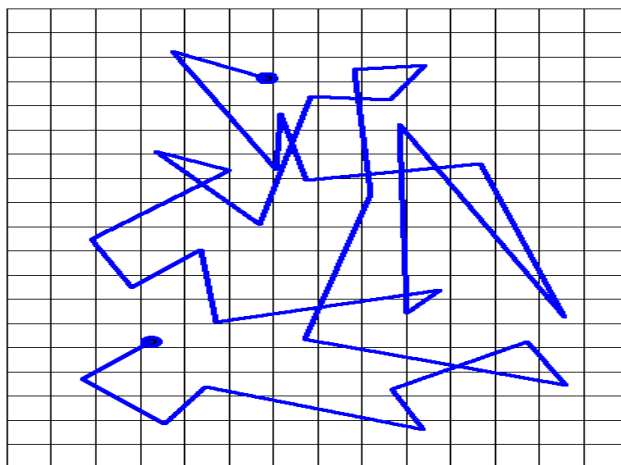
0:00		Ясно	+20	747		4	36	+16
3:00		Ясно	+17	747		0	42	+16
6:00		Ясно	+14	747		2	58	+12
9:00		Ясно	+23	748		3	31	+20
12:00		Ясно	+28	750		3	29	+26
15:00		Ясно	+29	749		3	22	+28
18:00		Ясно	+27	749		2	32	+26

Іншим прикладом є вимірювання параметру з випадковою помилкою

$$X = X_t + W_{t,\omega}$$



Броунівський рух



Природа випадкових подій

- 1) ∞ можливих факторів (параметрів)
- 2) наявність прихованих параметрів
- 3) флуктуації роботи вимірювальних пристроїв
- 4) складність систем
- 5)

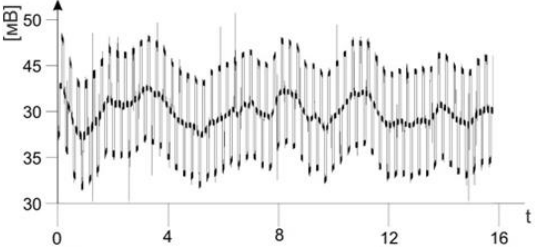
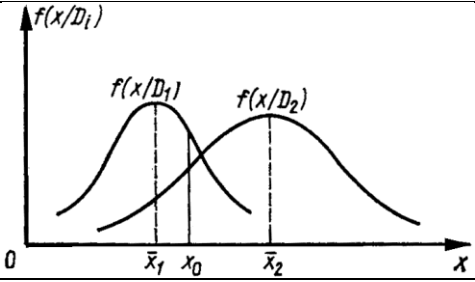
Предмет теорії ймовірностей

- { Теорія ймовірностей
- { Теорія випадкових процесів
- { Математична статистика

Основні поняття

Випадкова подія та її ймовірність

Експеримент з випадковим результатом

Підкидання монети, або гра «орел і решка»	
Передача сигналу по каналу зв'язку в присутності перешкод	
Термін служби технічного пристрою $X_i = \bar{X}_i + A_i(\omega)$	
Бозони ймовірність того, що в квантовомеханічній багаточастинковій системі існує бозон у одночастинковому квантовому стані $n\rangle$ із енергією ε_n обчислюється за формулою	Статистика Бозе-Ейнштейна $f(\varepsilon_n) = \frac{1}{e^{(\varepsilon_n - \mu)/k_B T} - 1}$ де μ — хімічний потенціал, k_B — стала Больцмана, T — температура.
Ферміони ймовірність того, що в квантовомеханічній багаточастинковій системі існує ферміон у одночастинковому квантовому стані $n\rangle$ із енергією ε_n обчислюється за формулою	Статистика Фермі-Дірака $f(\varepsilon_n) = \frac{1}{e^{(\varepsilon_n - \mu)/k_B T} + 1}$ де μ — хімічний потенціал, k_B — стала Больцмана, T — температура.
Канонічний ансамбль - статистичний ансамбль, який описує закриту термодинамічну систему, яка може обмінюватися енергією, але не частинками із середовищем. Ймовірність реалізації даного конкретного стану з енергією E_n залежить лише від значення енергії й задається розподілом	Розподіл Гібса $w_n = \frac{1}{Z} e^{-E_n/k_B T}$

Гібса, де Z - певна стала нормування, яка вираховується із цієї умови, щоб сума ймовірностей реалізації усіх мікроскопічних станів дорівнювала 1.	$Z = \sum_n e^{-E_n/k_B T}$ Z називається статистичною сумою.
---	--

Фундаментальні факти математичного аналізу

Фундаментальні константи

$$\pi = 3,14159265358... \quad \pi := Lc/D$$

$$e = 2,71828182845... \quad e := \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$$

Факторіал $n! := n(n-1) \cdot \dots \cdot 2 \cdot 1$; $0! := 1$

Комбінація, або сполука, або біноміальний коефіцієнт

$$C_n^k := \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Границя числової послідовності

$$\lim_{n \rightarrow \infty} a_n = a : \forall \varepsilon > 0 \exists N: |a_n - a| < \varepsilon \quad \forall n > N$$

Функції

$$f(x)$$

Приклади: $x^2, x^3, (x+a)^n = \sum_{k=0}^n \binom{n}{k} x^k a^{n-k}, e^x, \log_e x = \ln x$
 есс.

Границя функції

$$\lim_{x \rightarrow x_0} f(x) = a : \forall \varepsilon > 0 \exists \delta > 0: |f(x) - a| < \varepsilon \quad \forall x: |x - x_0| < \delta$$

Неперервна функція в точці x_0

$$\lim_{x \rightarrow x_0} f(x) = f(x_0)$$

Похідна

$$f'(x_0) := \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

Базові поняття теорії ймовірностей

Елементарні події $\omega \in \Omega$, випадкові події $A \subseteq \Omega$

A: Скінчений простір елементарних подій (скінченна кількість результатів)

Пр. 1 Кидання монет. $\Omega = \{T, C\}$

Пр.2 Кидання монети 2 рази. $\Omega = \{TT, TC, CT, CC\}$

Пр.3 Кидання монети n разів. $\Omega = \{(i_1, i_2, \dots, i_n), 1 \leq i_k \leq 6, 1 \leq k \leq n\}$

A: Нескінченний простір елементарних подій (нескінченна кількість результатів)

Пр.4 Кидання монети до першого виходу T. $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots, \omega_\infty\}$,

де $\omega_n = \underbrace{CC \dots C}_{(n-1) \text{ volte}} T$ подія, що T з'являється в n-му киданні, ω_∞ подія, що T не випаде ніколи.

Пр. 5 Задача про зустріч. 2 людини погоджуються зустрітися в інтервалі часу $[0, T]$. Нехай x - час прибуття персонажу A, y - час прибуття персонажу B, тоді

$$\Omega = \{(x, y) : 0 \leq x \leq T, \quad 0 \leq y \leq T\}$$

Пр. 6 Ми спостерігаємо частинку, яка здійснює броунівський рух. Простір елементарних подій Ω - сукупність усіх можливих траєкторій руху частинки. Такі події розглядаються як набори елементарних подій.

Операції з множинами (подіями):

Сума (об'єднання) $A \cup B$ Перетин $A \cap B$

Різниця $A \setminus B$ Тотальна сукупність, порожнеча $\Omega, \emptyset$

Додаткова множина $\bar{A} = \Omega \setminus A$

Деякі правила

Комутативність $A \cup B = B \cup A$; $A \cap B = B \cap A$

Дистрибутивність $A \cup (B \cap C) = (A \cap B) \cup (A \cap C)$;

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup \bar{A} = \Omega ; A \cap \bar{A} = \emptyset$$

Таблиця інтерпретацій

Позначення	Мова теорії множин	Мова теорії ймовірностей
Ω	Універсальна множина	Простір елементарних подій (результати експериментів)
ω	Елемент Ω	Елементарна подія
A	Множина елементів ω	Подія A (якщо $\omega \in A$, кажуть, що відбулася подія A)
Ω	Множина всіх ω	Детермінована подія
$\emptyset$	Пуста множина	Неможлива подія
$A \subset B$	A є підмножиною B	Виникнення події A обов'язково передбачає настання події B
$A \cup B$	Набір точок, які належать A або B	Подія “відбулося A або B ”
$A \cap B$	Набір точок, які належать A і B	Подія “відбулося як A , так і B ”
$A \cap B = \emptyset$	Множини A і B не перетинаються	A і B - це несумісні події
$A \setminus B$	Різниця множин A і B	Подія “відбулося A , і не відбулося B ”

Класичне визначення ймовірності

A – подія, яка може статися в випадковому експерименті, який можна відтворити кілька разів; n – кількість повторних експериментів; $k_n(A)$ – кількість експериментів, в яких відбулася подія A .

$\nu_n(A) := \frac{k_n(A)}{n}$ відносна частота A в серії n експериментів.

Властивості:

- 1) $0 \leq \nu_n(\Omega) \leq 1$
- 2) $\nu_n(\Omega) = 1$, $\nu_n(\emptyset) = 0$
- 3) Для несумісних подій A і B має місце: $\nu_n(A \cup B) = \nu_n(A) + \nu_n(B)$.

Визначення ймовірності $P(A) = \lim_{n \rightarrow \infty} \nu_n(A)$

Приклад 1. Кидання монети. 2 результати: (T, C)

$$v_n(T) = \frac{k_n(T)}{n} \rightarrow \frac{1}{2}, \quad n \rightarrow \infty.$$

Приклад 2. Кидання кубика 2 рази.

Ймовірність $P(A)$, $A = \{\text{сума чисел парна}\}$

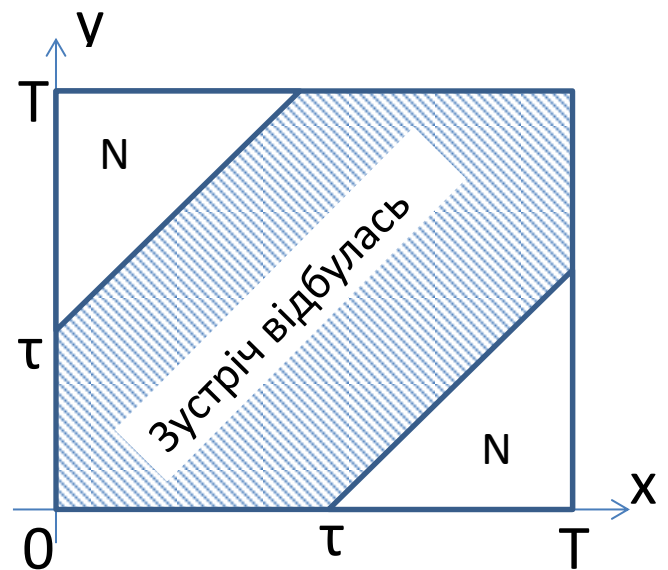
$\Omega = \{(1:1), (1:2), (1:3), \dots, (6:6)\}$ всього $6 \cdot 6 = 36$ елементарних результатів, з яких половина (18) мають парну суму чисел.

Частина, що стосується A , є $18/36 = 1/2$

(Неперервний випадок)

Приклад 3. (пр. 5 вище) Двоє людей погоджуються зустрітися в інтервалі часу $[0, T]$, кожен чекає час $\tau < T$ а потім він йде. Імовірність їх зустрічі обчислюється наступним чином:

$$\begin{aligned} & \frac{\text{Міра(зустріч відбулась)}}{\text{Міра(всі випадки)}} \\ &= \frac{T^2 - (T - \tau)^2}{T^2} = 1 - \left(1 - \frac{\tau}{T}\right)^2 \end{aligned}$$



Отже, з одного боку,

$$P(A) = \lim_{n \rightarrow \infty} v_n(A) \quad (1)$$

А з іншого боку

$$P(A) = \frac{\text{Міра(Сприятливих елементарних результатів)}}{\text{Міра(всіх можливих результатів)}} \quad (2)$$

Зауваження: Подвійна природа імовірності:

З одного боку, це частота, тобто границя частоти випадіння події (1)

З іншого боку, це міра простору елементарних результатів (2).

Незалежні події A, B :

$$P(A \cap B) = P(A)P(B)$$

Умовна ймовірність $P(A|B)$

Приклад 1. Кидання костей 2 рази.

$A = \{\text{перший вихід } 1\}$ і відомо, що відбулося $B = \{\text{сума чисел} < 4\}$

$$A = \{(1:1), (1:2), (1:3), (1:4), (1:5), (1:6)\}$$

$$B = \{(1:1), (1:2), (2:1)\}$$

$$P(A) = 6/36 = 1/6; \quad P(B) = 3/36 = 1/12$$

$$P(A \cap B) = 2/36 = 1/18$$

$$P(A|B) = \frac{2}{3} = \frac{P(A \cap B)}{P(B)}$$

Приклад 2. На площі $\Omega = [0,1]^2$ випадково кидається точка.

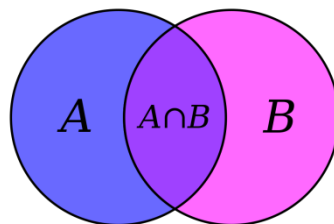
Для підмножини $A \subset \Omega$, $P(A) = m(A)$: площа фігури A . Отже, за визначенням ймовірності,

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

і це – визначення умовної ймовірності.

Деякі важливі факти:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (\text{формула додавання})$$



$$P(A \cap B) = P(B) P(A|B) = P(A) P(B|A) \quad (\text{формула множення})$$

Визначення: $H_1, H_2, \dots, H_n$ – повна група подій, якщо $H_1 \cup H_2 \cup \dots \cup H_n = \Omega$,

$H_i \cap H_j = \emptyset$ для всіх i, j

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

Формула Байеса

$$P(H_i|B) = \frac{P(H_i \cap B)}{P(B)} = \frac{P(H_i) P(B|H_i)}{\sum_{k=1}^n P(H_k) P(B|H_k)}$$

Випадкові величини та функції розподілу

Випадкова величина - одне з основних понять в теорії ймовірностей

Приклад 1

Підкидання монети 2 рази (Орел, Решка). $\Omega = \{OO, OP, PO, PP\}$

$\xi = \{ \text{кількість випадінь орла} \}$. Тоді таблиця функції $\xi = \xi(\omega)$ має наступну форму:

ω	OO	OP	PO	PP
$\xi(\omega)$	2	1	1	0

Приклад 2

Кидання кістки. $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$,

де $\omega_i = \{ \text{випало } i \text{ очок} \}$

Нехай $\xi = \{ \text{кількість очок, що випали} \}$. Тоді

$\xi(\omega_i) = i, i = 1, 2, \dots, 6$.

Приклад 3

Кидання монети до 1-го випадіння «Орел».

$$\Omega = \left\{ O, PO, PPO, \dots, \underbrace{PP \dots P}_{n-1} O, \dots \right\} = \{\omega_1, \omega_2, \omega_3, \dots, \omega_n, \dots\}$$

Нехай $\xi = \{ \text{кількість здійснених кидань монети} \}$.

Тоді $\xi(\omega) = i$, для $\omega = \omega_i, i = 1, 2, \dots$

Означення.

Випадкова величина - це числова функція

$\xi(\omega): \Omega \rightarrow \mathbb{R}$ з властивістю вимірюваності.

$(\Omega, \mathcal{F})$ $\mathcal{F}$ – сімейство підмножин Ω , замкнене відносно операцій об'єднання $A \cup B$, перетину $A \cap B$, різниці $A \setminus B$ доповнення $\bar{A} = \Omega \setminus A$, а також неперервності: замкненість відносно переходу до границі.

Перехід до границі

Монотонно зростаючі події $A_1 \subset A_2 \subset \dots \subset A_n \subset \dots$

Тоді $\lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n = A_{\infty}$

Монотонно спадаючі події $A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$

Тоді $\lim_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n = A_{\infty}$

Замкненість відносно переходу до границі означає, що

$$P\left(\lim_{n \rightarrow \infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n)$$

Випадкові величини | $\rightarrow$ дискретні (див. приклади вище)

| $\rightarrow$ неперервні (див. приклади вище)

Функція розподілу

(дискретний випадок)

$$F_{\xi}(x_n) = P\{\omega: \xi(\omega) = x_n\}$$

(неперервний випадок)

$$F_{\xi}(x) = P\{\omega: \xi(\omega) < x\}$$

Приклад 4.

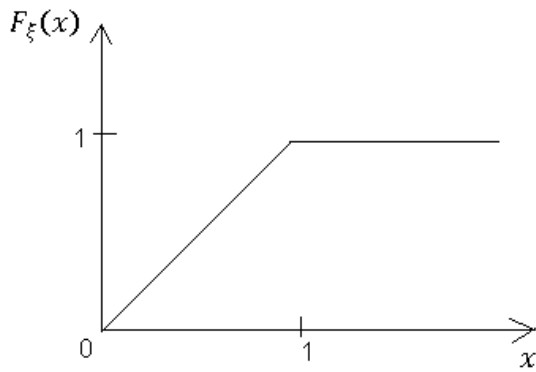
Рівномірний розподіл на $[0,1]$: кидається точка випадковим чином, але так, що всі позиції точки "однаково можливі".

Тоді випадкова величина має вигляд

$$\{\omega: \xi(\omega) < x\} = \begin{cases} \emptyset, & x \leq 0, \\ [0, x), & 0 < x \leq 1, \\ \Omega = [0,1], & x > 1 \end{cases}$$

Міра відрізка $P\{[0, x]\} = x$

$$F_{\xi}(x) = P\{\omega: \xi(\omega) < x\} = \begin{cases} 0, & x \leq 0, \\ x, & 0 < x \leq 1, \\ 1, & x > 1 \end{cases}$$



Приклад 5. Задача про зустріч.

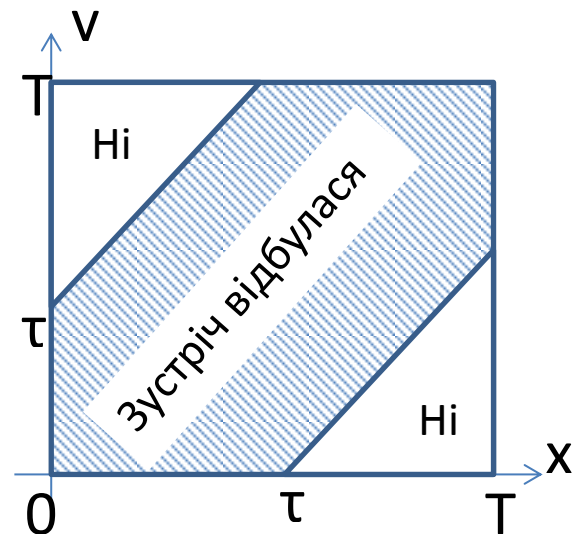
$$\Omega = \{(x, y) : 0 \leq x \leq T, 0 \leq y \leq T\}$$

Умови зустрічі:

$$C = \{(x, y) : |x - y| \leq \tau\}$$

Ймовірність зустрічі

$$\begin{aligned} P(C) &= \frac{m(C)}{T^2} \\ &= \frac{T^2 - (T - \tau)^2}{T^2} = 1 - \left(1 - \frac{\tau}{T}\right)^2 \end{aligned}$$



Позначимо $\xi = \xi(\omega)$ час очікування зустрічі.

Звідси для $\omega = (x, y)$ маємо $\xi(\omega) = |x - y|$

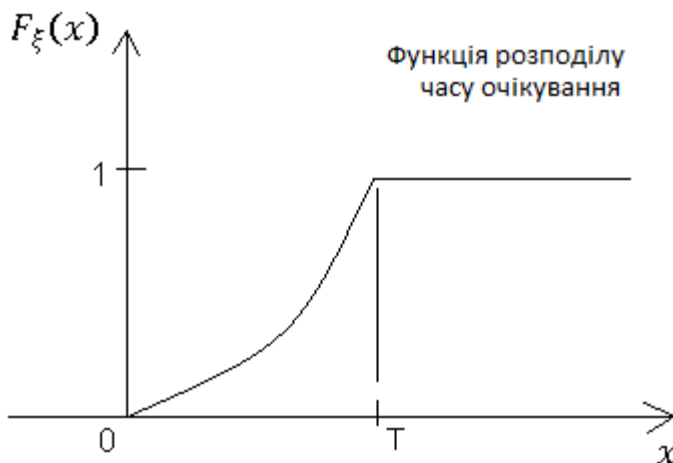
$$\text{Так що } \{\omega: \xi(\omega) < t\} = \begin{cases} \emptyset, & x \leq 0, \\ \{(x, y): |x - y| < t\}, & 0 \leq x, y \leq T, \\ [0, T] * [0, 1], & x > T \end{cases}$$

міра зустрічі

$$P\{[0, x]\} = x,$$

так що функція розподілу ймовірності зустрічі має вигляд:

$$F_{\xi}(x) = P\{\omega: \xi(\omega) < x\} = \begin{cases} 0, & x \leq 0, \\ \frac{T^2 - (T - x)^2}{T^2} = 1 - \left(1 - \frac{x}{T}\right)^2, & 0 < x \leq T, \\ 1, & x > T \end{cases}$$



Властивості ФУНКЦІЇ РОЗПОДІЛУ

$F(-\infty) = 0$, $F(+\infty) = 1$, $F(x)$ – монотонно зростаюча функція.

ВАЖЛИВІ типи функцій розподілу

Біноміальний розподіл : імовірнісна модель n незалежних тестів з 2 можливими виходами: успіх і невдача, p – ймовірність успіху.

Позначимо $\xi = \xi(\omega)$ число успіхів в n незалежних випробуваннях.

$$P_n(k) = P\{\omega: \xi(\omega) = k\} = C_n^k p^k q^{n-k}, q := 1 - p, \quad (k = 0, 1, \dots, n).$$

Геометричний розподіл: імовірнісна модель незалежних тестів з 2-ма можливими результатами: успіх і невдача до першої появи успіху.

$$P\{\omega: \xi(\omega) = n\} = q^{n-1}p \quad (n = 0, 1, \dots)$$

Розподіл Пуассона: імовірнісна модель “класичного процесу очікування”:

$$P\{\omega: \xi(\omega) = n\} = \frac{\lambda^n e^{-\lambda}}{n!}$$

Щільність розподілу (неперервний випадок)

Нехай $\xi = \xi(\omega)$ - випадкова величина з функцією розподілу $F(x)$.

Функція $p(u)$ називається ймовірнісною щільністю, якщо для всіх x має місце наступне:

$$F(x) = \int_{-\infty}^x p(u) du$$

Важливо ! якщо $p(u)$ є диференційованою функцією, і $F'(x) = p(x)$

Зокрема,

$$P\{a \leq \xi < b\} = F(b) - F(a)$$

З іншої сторони,

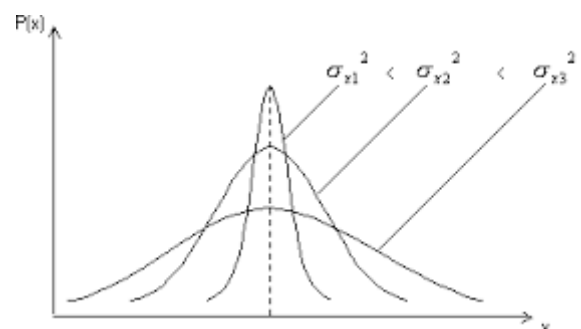
$$P\{a \leq \xi < b\} = \int_a^b p(u) du$$

Функція щільності **рівномірного розподілу** на $[0,1]$ має такий вигляд:

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

Нормальний (гаусовський) розподіл

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



Експоненціальний розподіл

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

має властивість **відсутності післядії** :

$$P\{t < \xi < s \mid \xi > t\} = P\{\xi < s - t\} \quad \text{by } t < s.$$

Це важлива властивість показникового розподілу, яку серед неперервних розподілів має тільки показниковий розподіл.

Позначимо

$$Q(x) := P\{\xi \geq x\} = 1 - F(x)$$

Тоді

$$\begin{aligned} P\{\xi < t + x \mid \xi \geq t\} &= \frac{P\{t \leq \xi < t + x\}}{P\{\xi \geq t\}} = \frac{Q(t) - Q(t + x)}{Q(t)} \\ &= 1 - Q(x) \end{aligned}$$

Звідки

$$Q(t + x) = Q(t)Q(x).$$

Математичне сподівання випадкової величини

А: Дискретний випадок.

Нехай випадкова величина ξ має числену (скінчену або нескінчену) множину значень $\{a_1, a_2, \dots, a_n, \dots\}$ з розподілом $p_k = P\{\xi = a_k\}$.

Позначимо $\{x_1, x_2, \dots, x_n, \dots\}$ спостереження випадкової величини $\xi = \xi(\omega)$. Тоді середнє значення спостережуваних величин представляється наступним чином:

$$\frac{1}{N} \sum_{k=1}^N x_k = \sum_{i=1}^r a_i \frac{k_i}{N},$$

де k_i - кількість випадків події $\{\xi = a_i\}$ у нашій серії експериментів. Отже, таким чином,

$$\frac{1}{N} \sum_{k=1}^N x_k = \sum_{i=1}^r a_i \nu_N(\{\xi = a_i\}),$$

де $\nu_N(A)$ - це частота події А, яка, як ми знаємо, $\rightarrow P(A)$.

Підставляючи замість частот ймовірності, маємо:

$$\sum_{i=1}^r a_i p_i = E\xi$$

МАТЕМАТИЧНЕ СПОДІВАННЯ, АБО ЙМОВІРНОСНЕ СЕРЕДНЄ ЗНАЧЕННЯ

В: Неперервний випадок.

Нехай $\xi = \xi(\omega)$ має неперервний розподіл, з функцією щільності $p(u)$:

$$F(x) = \int_{-\infty}^x p(u) du$$

де $F(x)$ - функція розподілу випадкової величини ξ

Тоді

$$\int_{-\infty}^{\infty} up(u) du = E\xi$$

МАТЕМАТИЧНЕ СПОДІВАННЯ, АБО ЙМОВІРНОСНЕ СЕРЕДНЄ ЗНАЧЕННЯ

Приклад 1

Кидання грального кубика.

$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, де $\omega_i = \{\text{випало } i \text{ пунктів}\}$
 $= \{\text{кількість очок, що випали}\}$.

Нехай $\xi(\omega_i) = i$, $i = 1, 2, \dots, 6$. Тоді $E\xi = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3,5$

Приклад 2 Геометричний розподіл

Кидання монети до 1-го випадіння орла.

$\Omega = \left\{O, PO, PPO, \dots, \underbrace{PP \dots P}_{n-1} O, \dots\right\} = \{\omega_1, \omega_2, \omega_3, \dots, \omega_n, \dots\}$

Нехай $\xi = \{\text{кількість здійснених кидань}\}$.

Так що $\xi(\omega) = i$, при $\omega = \omega_i$, $i = 1, 2, \dots$

Що має геометричний розподіл ймовірностей

$$p_1 = P\left\{\xi(\omega_1) = \frac{1}{2}\right\}, \quad p_2 = P\left\{\xi(\omega_2) = \frac{1}{2} \cdot \frac{1}{2}\right\}, \dots,$$
$$p_n = P\left\{\xi(\omega_n) = \frac{1}{2} \cdot \frac{1}{2^{n-1}}\right\}$$

Тоді $E\xi = \sum_{n=1}^{\infty} n \cdot p_n = \sum_{n=1}^{\infty} n \cdot \frac{1}{2^n} = \frac{1}{2} \cdot 2^2 = 2$

Приклад 3 : Біноміальний розподіл - імовірнісна модель незалежних тестів з 2 можливими виходами: успіх (У) й невдача (Н).

Позначимо $\xi = \xi(\omega)$ кількість успіхів у n випробуваннях.

$$P_n(k) = P\{\omega: \xi(\omega) = k\} = C_n^k p^k q^{n-k} \quad (k = 0, 1, \dots, n).$$

Тоді

$$E\xi = \sum_{k=0}^n k C_n^k p^k q^{n-k} = np \sum_{k=1}^n \frac{k}{n} C_n^k p^{k-1} q^{n-k} =$$
$$= np \sum_{k=1}^n C_{n-1}^{k-1} p^{k-1} q^{n-k} = np \sum_{r=0}^{n-1} C_{n-1}^r p^r q^{n-r-1} = np.$$

Приклад 3 : Розподіл Пуассона

$$P\{\omega: \xi(\omega) = k\} = \frac{\lambda^k e^{-\lambda}}{k!} \quad k = 0, 1, 2, \dots$$

За визначенням маємо

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

Визначення: ξ інтегроване, якщо існує його математичне сподівання.

Властивості математичного сподівання

1. If $P\{\xi = c\} = 1$, тоді $E\xi = c$.
2. If $\xi \geq 0$, тоді $E\xi \geq 0$.
3. If $|\xi| \leq C$, тоді $|E\xi| \leq C$.
4. If ξ інтегроване, тоді $a\xi$ також інтегроване і $Ea\xi = aE\xi$.
5. If ξ, η are інтегроване, тоді $\xi + \eta$ також інтегроване і $E(\xi + \eta) = E\xi + E\eta$.
6. $E(a\xi + b\eta) = aE\xi + bE\eta$.
7. If ξ, η – інтегровані і $\xi \leq \eta$, тоді $E\xi \leq E\eta$.
8. R.V. ξ , $|\xi|$ є разом інтегрованими і $E\xi \leq E|\xi|$.
9. $\xi_n, n \geq 0$ – послідовність невід'ємних випадкових величин :
 $\xi_n \leq \xi_{n+1}$ і $\xi_n \rightarrow \xi$.

Тоді, якщо $\sup_n E\xi_n < \infty$, то ξ - інтегроване, і

$$E\xi = \lim_{n \rightarrow \infty} E\xi_n.$$

10. If ξ – невід'ємна випадкова величина і $a > 0$. Тоді

$$P\{\xi \geq a\} \leq \frac{1}{a} E\xi \quad (\text{Нерівність Чебишева})$$

Декілька прикладів математичного сподівання для неперервних випадкових величин

Приклад 4 : Рівномірний розподіл

Рівномірний розподіл на $[0,1]$: кидання точки випадковим чином на відрізок, $[0,1]$ таким чином, що всі положення точки "однаково можливі"

$$\{\omega: \xi(\omega) < x\} = \begin{cases} \emptyset, & x \leq 0, \\ [0, x), & 0 < x \leq 1, \\ \Omega = [0,1], & x > 1 \end{cases}$$

$$P\{[0, x]\} = x,$$

$$F_{\xi}(x) = P\{\omega: \xi(\omega) < x\} = \begin{cases} 0, & x \leq 0, \\ x, & 0 < x \leq 1, \\ 1, & x > 1 \end{cases}$$

$$E\xi = \int_0^1 x dx = \frac{x^2}{2} \Big|_0^1 = \frac{1}{2}$$

Приклад 5 : Експоненціальний розподіл

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

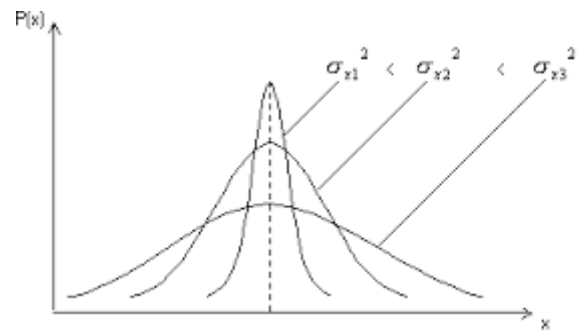
$$E\xi = \int_0^{\infty} x \lambda e^{-\lambda x} dx = - \int_0^{\infty} x d(e^{-\lambda x}) = - x e^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx = \frac{1}{\lambda}.$$

(в теорії надійності, λ є інтенсивністю технічних несправностей,

$E\xi$ – математичне сподівання (середнє значення часу безвідмовної роботи (MTBF - mean time between failures)).

Приклад 6 : Нормальний (гаусовський) розподіл

$$p(x) = \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^x e^{-\frac{(u-a)^2}{2\sigma^2}} dx$$



$$\begin{aligned} E\xi &= \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{+\infty} x e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{+\infty} (x - a + a) e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \\ &= \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{+\infty} (x - a) e^{-\frac{(x-a)^2}{2\sigma^2}} dx + a \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{+\infty} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = a. \end{aligned}$$

Деякі числові характеристики, що характеризують розподіл випадкової величини.

Центрировані моменти k-го класу:

$$m_k(\xi) = \sum_{i=1}^N (x_i - E\xi)^k p_i$$
$$m_k(\xi) = \int_{-\infty}^{+\infty} (u - E\xi)^k p(u) du$$

Дісперсія (Середньоквадратичне відхилення) : k=2

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

Властивості дисперсії:

- (Нерівність Чебишева)

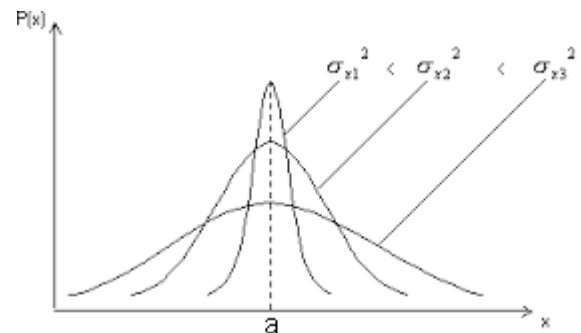
$$P\{|\xi - E\xi| > \varepsilon\} \leq \frac{D\xi}{\varepsilon^2}$$

- (Конструктивний розрахунок)

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

Зокрема, у разі **НОРМАЛЬНОГО (ГАУСОВСЬКОГО) РОЗПОДІЛУ**

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



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

$$\sigma = \sqrt{D\xi} = \sqrt{E(\xi - E\xi)^2} = \sqrt{E\xi^2 - (E\xi)^2}$$

Коефіцієнт варіації

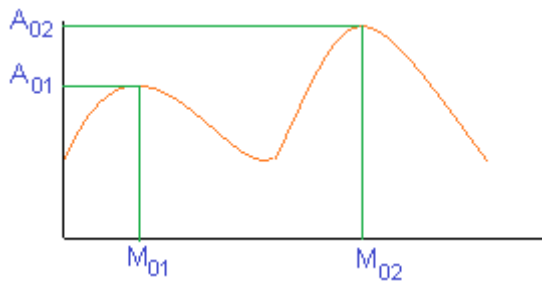
$$V_{\xi} = \frac{\sigma}{E\xi} = \frac{\sqrt{D\xi}}{E\xi}$$

Зазвичай використовується для оцінки мінливості випадкових величин ξ .

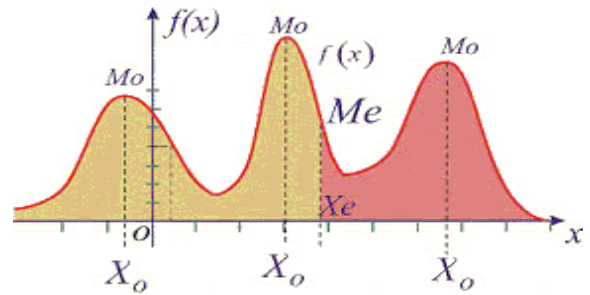
Мода розподілу

Значення випадкової величини ξ , для яких ймовірність максимальна (або локально максимальна), називається **модою** і позначається $M_0(\xi)$.

Відповідне значення ймовірності називаються **амплітудою** та позначаються $A_0(\xi)$.



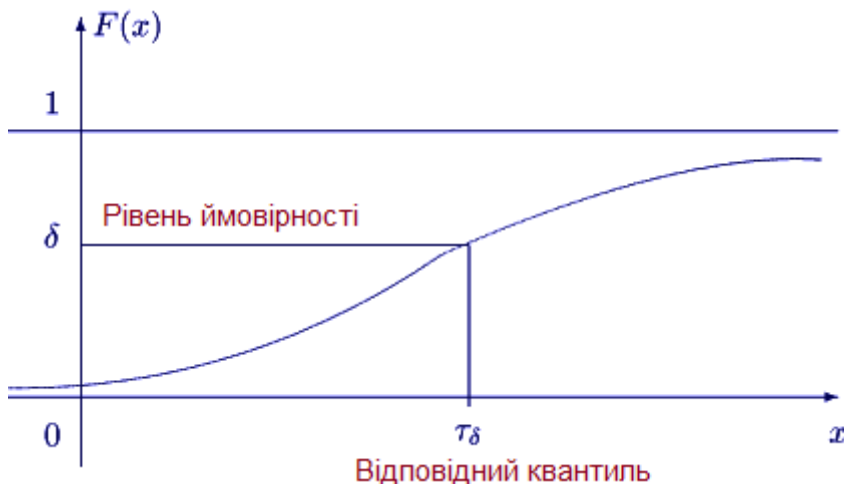
Бімодальний розподіл



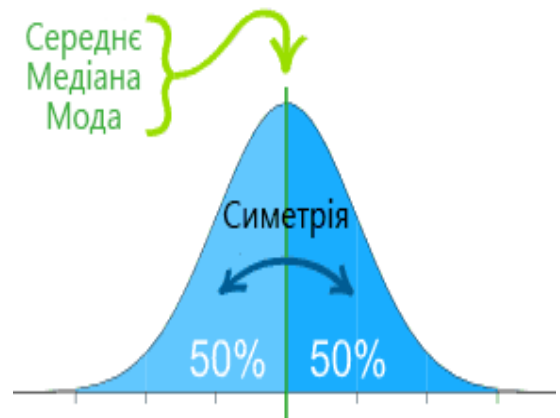
Тримодальний розподіл

Квантіль, що відповідає рівню ймовірності p ,

є таким значенням випадкової величини, що $\xi = x: F(x) = p$.



Медіана розподілу називається квантилем, який відповідає значенню ймовірності $p = \frac{1}{2}$.



Нерівності Чебишева та деякі інші

Нерівність Чебишева – 1

Нехай ξ – невід’ємна випадкова величина і $a > 0$. Тоді

$$P\{\xi \geq a\} \leq \frac{1}{a} E\xi$$

Доведення: розглянемо випадкову величину, яка приймає два значення:

$$\eta := \begin{cases} 0, & \text{якщо } \xi < a, \\ a, & \text{якщо } \xi \geq a, \end{cases}$$

Очевидно, що $\xi \geq \eta$, тому $E\xi \geq E\eta$. Математичне сподівання η :

$$E\eta = a P\{\eta = a\} = a P\{\xi \geq a\} \leq E\xi.$$

Нерівність Чебишева – 2

$$P\{|\xi - E\xi| > \varepsilon\} \leq \frac{D\xi}{\varepsilon^2}, \quad D\xi = E(\xi - E\xi)^2 = E\xi^2 - (E\xi)^2$$

Доведення: враховуючи нерівність Чебишева – 1, маємо

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

Нерівності для дисперсії

- Якщо $D\xi$ існує, то для всіх c має місце нерівність

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

і отже, $E(\xi - c)^2$ приймає найменше значення при $c = E\xi$.

- Якщо $D\xi$ існує, то для всіх a і b має місце нерівність

$$D(a\xi + b) \leq a^2 D\xi$$

Нерівність Гьольдера

Нехай $p > 1$, $q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$ і випадкові величини ξ, η такі, що

$E|\xi|^p < +\infty$, $E|\eta|^q < +\infty$. Тоді

$$E|\xi\eta| \leq (E|\xi|^p)^{\frac{1}{p}} (E|\eta|^q)^{\frac{1}{q}}$$

Доведення.

Розглянемо криву ОВА, рівняння якої

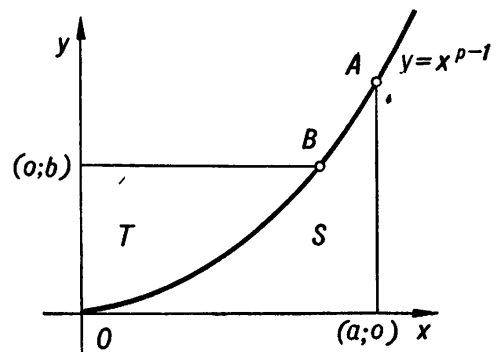
$$y = x^{p-1}, \text{ або } x = y^{q-1}$$

Якщо $b < a^{p-1}$, загальна площа фігур

$$S = \int_0^a x^{p-1} dx = \frac{1}{p} a^p, \quad T = \int_0^b y^{q-1} dy = \frac{1}{q} b^q$$

Не перевищує площу квадрата ab :

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$



Положимо в цій нерівності

$$a = \frac{|\xi|}{(E|\xi|^p)^{\frac{1}{p}}}, \quad b = \frac{|\eta|}{(E|\eta|^q)^{\frac{1}{q}}}$$

і візьмемо

математичне сподівання їх добутку

$$E(ab) = E\left(\frac{|\xi|}{(E|\xi|^p)^{\frac{1}{p}}} \cdot \frac{|\eta|}{(E|\eta|^q)^{\frac{1}{q}}}\right) \leq \frac{1}{p} + \frac{1}{q} = 1, \text{ що доводить нерівність.}$$

Нерівність Коші-Буняковського

Якщо для випадкових величин ξ, η їх другі моменти скінчені

$$E\xi^2 < \infty, \quad E\eta^2 < \infty,$$

то справедлива нерівність

$$\boxed{(E\xi\eta)^2 \leq E\xi^2 E\eta^2}$$

Доведення. Є частинним випадком нерівності Гьольдера для $p = q = 2$.

Нерівність Шварца

Якщо для випадкових величин ξ, η їх другі моменти скінчені

$$E\xi^2 < \infty, \quad E\eta^2 < \infty,$$

то справедлива нерівність

$$\boxed{\sqrt{E(\xi + \eta)^2} \leq \sqrt{E\xi^2} + \sqrt{E\eta^2}}$$

Нерівність Йєнсена

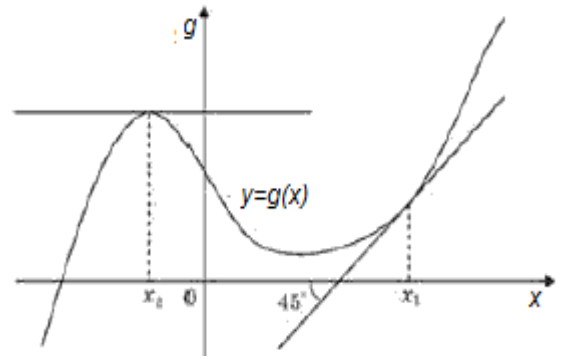
Нехай випадкова величина ξ належить області значень $\mathfrak{X}$, а $g(x)$ – опукла функція на $\mathfrak{X}$, тобто $g''(x) \geq 0$ для $x \in \mathfrak{X}$. Тоді

$$\boxed{Eg(\xi) \geq g(E\xi)}$$

Доведення.

Нехай $C = (x_0, g(x_0))$ – точка, яка належить графіку функції $g(x)$.

Розглянемо пряму l , яка проходить через точку C . Така пряма l зветься опорною, якщо графік $g(x)$ цілком знаходиться над l або на l .



В силу опуклості функції $g(x)$ існує коефіцієнт нахилу $k(x_0)$ прямої l такий, що для всіх $x \in \mathfrak{H}$ має місце нерівність

$$g(x) \geq g(x_0) + k(x_0)(x - x_0).$$

Положимо в цій нерівності $x = \xi$, $x_0 = E\xi$ і візьмемо математичне сподівання від обох частин. Використовуючи властивість мажорації, маємо

$$Eg(\xi) \geq g(E\xi).$$

Нерівність Ляпунова

Нехай $0 < s < t$. Тоді

$$\boxed{(E|\xi|^s)^{\frac{1}{s}} \leq (E|\xi|^t)^{\frac{1}{t}}}$$

Доведення. Покладемо $r = \frac{t}{s}$, $g(x) = |x|^r$, $r \geq 1$, $\eta = |\xi|^s$.

Враховуючи нерівність Йєнсена, маємо нерівність Ляпунова.

Деякі застосування нерівності Чебишева

Нерівність Чебишева дає оцінку ймовірності «значних» відхилень випадкової величини відносно свого математичного сподівання. Запишем нерівність Чебишева у такій формі

$$P\{|\xi - E\xi| > k\sqrt{D\xi}\} \leq \frac{D\xi}{k^2 D\xi} = \frac{1}{k^2}, \quad k > 0.$$

Тоді

$$P\{|\xi - E\xi| \leq k\sqrt{D\xi}\} = 1 - P\{|\xi - E\xi| > k\sqrt{D\xi}\} \geq 1 - \frac{1}{k^2}.$$

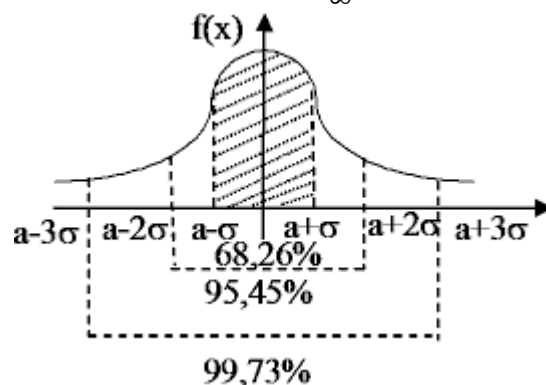
Нехай $k = 3$. Тоді

$$P\{|\xi - E\xi| \leq 3\sqrt{D\xi}\} \geq 1 - \frac{1}{9} = \frac{8}{9}.$$

ξ - нормальний розподіл $N(a, \sigma^2)$. Тоді величина $\frac{\xi - a}{\sigma}$ має нормальний розподіл $N(0, 1)$ і в силу нерівності Чебишева маємо

$$P\{|\xi - a| > 3\sigma\} = P\left\{\left|\frac{\xi - a}{\sigma}\right| > 3\right\} = 2\Phi(3) \approx 0,997.$$

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$



Незалежні випадкові величини

Ймовірносний простір $(\Omega, \mathfrak{F}, P)$

Події та їх сукупності, про які мова йтиме далі, належать до $\mathfrak{F}$.

Події $A_1, A_2, \dots, A_n$ називаються незалежними (або незалежними

в сукупності), якщо для всіх цілих $1 \leq m \leq n$, $1 \leq i_1 < i_2 < \dots < i_m \leq n$ виконується рівність

$$P\left\{\bigcap_{k=1}^m A_{i_k}\right\} = \prod_{k=1}^m P\{A_{i_k}\}$$

Послідовність подій $\{A_n, n \geq 1\}$ називається послідовністю незалежних подій, якщо довільна скінченна сукупність подій цієї послідовності є сукупність незалежних подій.

Позначимо через $\mathfrak{B}_n$ мінімальну σ -алгебру, яка включає σ -алгебри $\mathfrak{F}_m$ $m \geq n$.

σ -алгебра $\bigcap_{m=1}^{\infty} \mathfrak{B}_m$ і події з неї називаються **асимптотичними або залишковими** відносно послідовності σ -алгебр $\{\mathfrak{F}_n, n \geq 1\}$.

При розгляді послідовностей подій часто використовується верхня границя подій. **Верхня границя послідовності подій** $\{A_n, n \geq 1\}$ є подія

$$\overline{\lim}_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

подія $\overline{\lim}_{n \rightarrow \infty} A_n$ відбувається тоді і тільки тоді, коли відбувається нескінченно багато з подій $A_n, n \geq 1$.

Лема Бореля—Кантеллі. Нехай $\{A_n, n \geq 1\}$ — послідовність подій, для якої

$$\sum_{n=1}^{\infty} P(A_n) < \infty.$$

Тоді

$$P\left(\overline{\lim}_{n \rightarrow \infty} A_n\right) = 0.$$

Якщо для послідовності незалежних подій $\{A_n, n \geq 1\}$, ряд

$$\sum_{n=1}^{\infty} P(A_n) = +\infty.$$

Тоді

$$P\left(\overline{\lim}_{n \rightarrow \infty} A_n\right) = 1.$$

Закон 0 і 1 Колмогорова. Нехай $\{\mathfrak{F}_n, n \geq 1\}$ — послідовність незалежних σ -алгебр

і $\mathfrak{F}$ є залишкова σ -алгебра відносно цієї послідовності. Для всякої події $A \in \mathfrak{F}$ імовірність $P(A)$ дорівнює або 0 або 1.

Граничні теореми для біноміального розподілу

Схема Бернуллі:

n незалежних випробувань, два результати: успіх з ймовірністю p і невдача з ймовірністю q , ($p + q = 1$).

Позначимо γ_n число успіхів в n випробуваннях. Тоді

$$B_p(m, n) = P\{\gamma_n = m\} = C_n^m p^m q^{n-m},$$

$$m = 0, 1, \dots, n,$$

$$E\gamma_n = np, \quad D\gamma_n = npq.$$

Задача: Яка буде гранична поведінка ймовірності $B_p(m, n)$ при $n \rightarrow \infty$?

Локальна гранична теорема.

Позначимо $a_n = np$, $b_n = npq$, $x_{n,m} = (m - a_n)b_n^{-\frac{1}{2}}$

Тоді, якщо при $n \rightarrow \infty$ також $b_n \rightarrow \infty$, причому $|x_{n,m}| \leq c$ для деякої константи c .

Тоді

$$\lim_{n \rightarrow \infty} \frac{B_p(m, n)}{\frac{1}{\sqrt{2\pi} b_n} \exp\left\{-\frac{1}{2} x_{n,m}^2\right\}} = 1.$$

Наслідок.

Інтегральна гранична теорема.

В умовах попередньої теореми нехай $-\infty \leq a < b \leq \infty$.

Тоді

$$P\left\{a < \frac{\gamma_n - np}{\sqrt{npq}} < b\right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}x^2} dx.$$

Це означає, що розподіл випадкової величини $\frac{\gamma_n - np}{\sqrt{npq}}$ прямує до нормального розподілу зі середнім 0 і дисперсією 1.

Формула Пуассона (закон рідкісних подій)

Якщо $a_n = E\gamma_n = np \leq c$, де c – довільна константа, то для всіх m

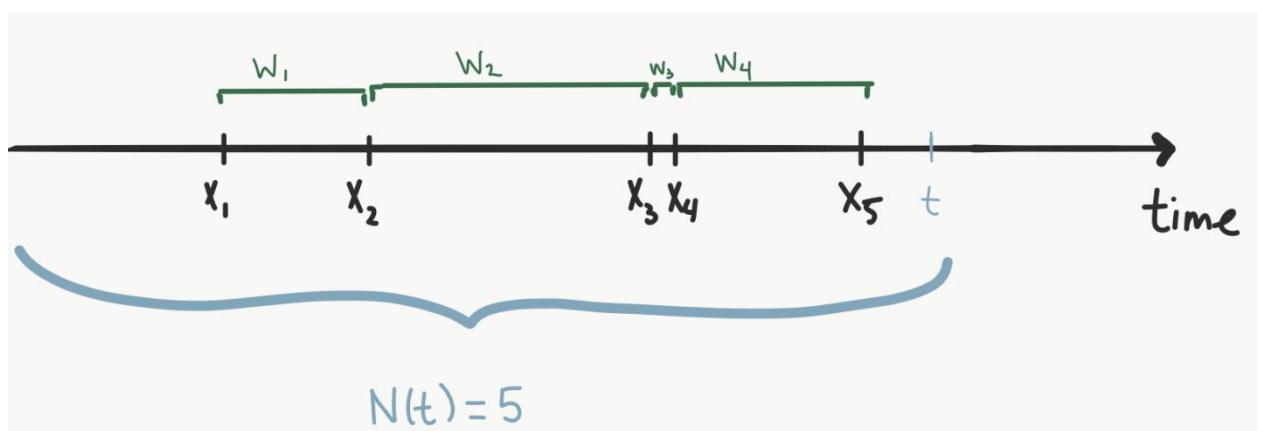
$$B_p(m, n) \sim \frac{a_n^m}{m!} e^{-a_n}. \quad (*)$$

Доведення

$$\begin{aligned} B_p(m, n) &= \frac{n(n-1) \dots (n-m+1)}{m!} p^{n-m} q^{n-m} = \\ &= \frac{n(n-1) \dots (n-m+1)}{n^m m!} a_n^m \left(1 - \frac{a_n}{n}\right)^{n-m} q^{n-m} \sim \frac{a_n^m}{m!} e^{-a_n}. \end{aligned}$$

Формула (*) має назву формули Пуассона. З неї випливає, що при великій кількості n випробувань, число появ події, що має ймовірність порядку $\frac{1}{n}$, наближено розподілено згідно із законом Пуассона. Формула Пуассона називається ще **законом рідкісних подій**. Зазвичай цим законом підпорядковується число появ деякої події, залежить від великого числа незалежних чинників, наприклад, число частинок, зареєстрованих лічильником космічних частинок, число катастроф, число викликів, які надійшли на телефонну станцію, і т. п.

Процес Пуассона



Розглянемо події, які можуть відбуватися в кожен момент безперервно мінливого часу. До таких подій можна віднести реєстрацію частки лічильником, надходження телефонного виклику на телефонну станцію, випуск ядром частинки при радіоактивному розпаді тощо. Якщо

прийняти, що за кінцевий час відбувається лише скінчене число подій, то при досить загальних припущеннях число появ події за будь-який час t має розподіл Пуассона.

Нехай s – число появ події на проміжку $[0, t]$. Будемо вважати, що $\xi(t)$ визначена для всіх t і є невід'ємною цілочисленною випадковою величиною. Тоді для $t_1 < t_2$ число появ подій на $(t_1, t_2]$ дорівнює $\xi(t_2) - \xi(t_1)$. Основне припущення щодо сукупності розглянутих подій узагальнює припущення про їх незалежність:

1) число подій, що з'явилися на проміжках, що не перетинаються, є незалежними випадковими величинами.

Це означає, що при $0 < t_1 < t_2 < \dots < t_n$ випадкові величини

$$\xi(t_1), \xi(t_2) - \xi(t_1), \dots, \xi(t_n) - \xi(t_{n-1}) \quad (1)$$

незалежні між собою.

Інші дві умови забезпечують однорідність процесу появи подій в часі (ця умова аналогічна припущенню про те, що в дискретної послідовності події мають однакову ймовірність) і

неможливість появи одночасно декількох подій (це властивість називається ординарністю потоку подій). Отримаємо ці умови в математичній формі.

Однорідність процесу означає, що число появ подій в будь-який проміжок часу залежить лише від тривалості цього проміжку, тобто розподіл величини $\xi(t + h) - \xi(t)$ залежить від h , але не залежить від t .

Умова ординарності потоку:

$$\lim_{h \rightarrow 0} \frac{1}{h} P\{\xi(t + h) - \xi(t) > 1\} = 0. \quad (2)$$

Далі, для однорідного ординарного потоку подій існує таке $\lambda > 0$, що

$$\lim_{h \rightarrow 0} \frac{1}{h} P\{\xi(t + h) - \xi(t) > 0\} = \lambda. \quad (3)$$

Теорема. Нехай для сімейства ціло-чисельних невід'ємних випадкових величин $\xi(t)$ виконана умова 1), розподіл величини $\xi(t + h) - \xi(t)$ не залежить від t і виконані умови 2) і 3). Тоді

$$P\{\xi(t) = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, 2, \dots \quad (4)$$

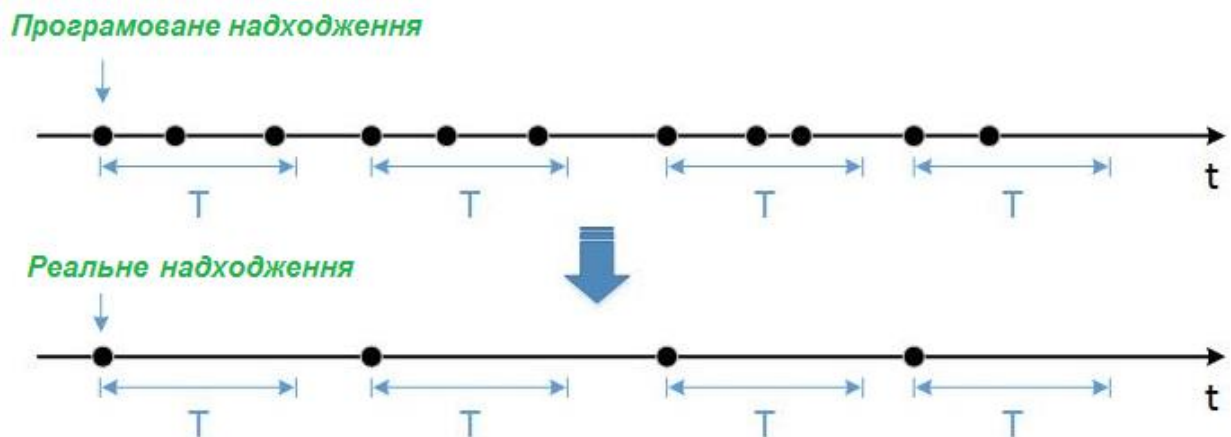
При вивченні потоку подій введено поняття сімейства випадкових величин $\xi(t)$, визначених для кожного $t \in [0, \infty)$.

Іншими словами, розглянута функція на деякій множині дійсних чисел, значеннями якої є випадкові величини. Такого роду функції називаються випадковими функціями. Як і в розглянутому нами випадку,

аргумент функції часто інтерпретується як час. Тоді така випадкова функція називається випадковим процесом.

Випадковий процес, розглянутий в цьому параграфі, називається процесом Пуассона. Умова 1), якому він задовольняє, - це умова незалежності приростів процесу $\xi(t)$. Будь-який процес, для якого властивість 1) виконана, називається процесом з незалежними приростами. Теорема, доведена в цьому параграфі, дає умови, при яких процес з незалежними приростами є процесом Пуассона.

Ілюстрація процесів Пуассона: надходження заявок



Послідовності випадкових величин

Нерівність Колмогорова

При вивченні поведінки сум незалежних випадкових величин часто доводиться оцінювати ймовірності відхилення максимальної суми.

Найбільш проста, і в той же час, загальна оцінка такої ймовірності отримана А.М. Колмогоровим.

Теорема. Нехай $\xi_1, \dots, \xi_n$ – незалежні випадкові величини, для яких $E\xi_k = 0$, $k = 1, 2, \dots, n$, $D\xi_k < \infty$, $k = 1, 2, \dots, n$,
 Покладемо $\zeta_k = \xi_1 + \dots + \xi_k$. Тоді для $a > 0$

$$P\left\{\sup_{0 \leq k \leq n} |\zeta_k| \geq a\right\} \leq \frac{1}{a^2} \sum_{k=1}^n D\xi_k.$$

Ця нерівність називається нерівністю Колмогорова.

Наслідок.

Нехай ξ_k , $k = 1, 2, \dots$ - нескінченна послідовність незалежних випадкових величин, $E\xi_k = 0$, $D\xi_k < \infty$, а також $\sum_{k=1}^{\infty} D\xi_k < \infty$.
 Покладемо $\zeta_k = \xi_1 + \dots + \xi_k$. Тоді

$$P\left\{\sup_{0 \leq k \leq n} |\zeta_k| > a\right\} \leq \frac{1}{a^2} \sum_{k=1}^{\infty} D\xi_k.$$

Ймовірнісні збіжності

Визначення 1. Послідовність випадкових величин

ξ_n збігається **за ймовірністю** до випадкової величини ξ , якщо для будь-якого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\{|\xi_n - \xi| > \varepsilon\} = 0.$$
$$\xi_n \xrightarrow{P} \xi.$$

Визначення 2. Послідовність випадкових величин

ξ_n збігається **у середньому** до випадкової величини ξ , якщо для всіх n існують $E\xi_n < \infty$, $E\xi < \infty$ і має місце

$$\lim_{n \rightarrow \infty} E|\xi_n - \xi| = 0.$$

Очевидно, в такому разі

$$\lim_{n \rightarrow \infty} E\xi_n = E\xi.$$

Порівняння

Якщо $\xi_n \rightarrow \xi$ у середньому, то $\xi_n \rightarrow \xi$ за ймовірністю.

Це випливає з нерівності Чебишева

$$P\{|\xi_n - \xi| > \varepsilon\} \leq \frac{1}{\varepsilon} E|\xi_n - \xi|.$$

Визначення 3. Послідовність випадкових величин

ξ_n збігається **у середньому квадратичному** до випадкової величини ξ , якщо для всіх n існують $E\xi_n^2$, $E\xi^2$ і має місце

$$\lim_{n \rightarrow \infty} E|\xi_n - \xi|^2 = 0.$$

Порівняння

Якщо $\xi_n \rightarrow \xi$ у середньому квадратичному, то також $\xi_n \rightarrow \xi$ у середньому,

а значить $\xi_n \rightarrow \xi$ за ймовірністю $\xi_n \xrightarrow{P} \xi$.

Це випливає з нерівності Коші-Буняковського

$$E|\xi_n - \xi| \leq (E|\xi_n - \xi|^2)^{\frac{1}{2}}.$$

Закон великих чисел

Нехай ξ_n – послідовність випадкових величин, для яких існують $E\xi_n$.

Законом великих чисел називаються теореми, які стверджують, що різниця

$$\frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n E\xi_k$$

збігається до нуля за ймовірністю.

Теорема Чебишева. Нехай ξ_n – послідовність незалежних випадкових величин і для всіх n існують $D\xi_n < \infty$, причому $D\xi_n \leq c$ для всіх n . Тоді

$$\frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n E\xi_k \xrightarrow{P} 0.$$

Наслідок.

Нехай $\xi_1, \dots, \xi_n$ – послідовність незалежних випадкових величин, така що

$$E\xi_n = a, \quad D\xi_n \leq c, \quad n = 1, 2, \dots$$

Тоді для кожного $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{\xi_1 + \dots + \xi_n}{n} - a \right| > \varepsilon \right\} = 0.$$

Цей окремий випадок теореми Чебишева дає обґрунтування правила середнього арифметичного в теорії обробки результатів вимірювань. Припустимо, що треба виміряти деяку фізичну величину a . Повторивши вимірювання n разів в однакових умовах, спостерігач отримує результати вимірювань $\xi_1, \dots, \xi_n$. Як наближене значення a приймається середнє арифметичне результатів вимірювань

$$\frac{\xi_1 + \dots + \xi_n}{n}$$

Якщо спостереження позбавлені систематичної помилки, тобто $E\xi_n = a$, то згідно сформульованому вище наслідку,

$$\frac{\xi_1 + \dots + \xi_n}{n} - a \xrightarrow{P} 0.$$

Визначення 4. Послідовність випадкових величин

ξ_n збігається з ймовірністю 1 до випадкової величини ξ , якщо має місце

$$P\{\lim_{n \rightarrow \infty} \xi_n = \xi\} = P\left\{\bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \left\{|\xi_n - \xi| \leq \frac{1}{k}\right\}\right\} = 1.$$

Теорема 2.

Для того, щоб послідовність ξ_n збігалася з ймовірністю 1 до випадкової величини ξ , необхідно і достатньо, щоб для всіх $\varepsilon > 0$

$$\lim_{N \rightarrow \infty} P\left\{\bigcap_{n=N}^{\infty} \{|\xi_n - \xi| \leq \varepsilon\}\right\} = 1.$$

з ймовірністю 1 до величини ξ ,

Наслідок. Теорема 2 еквівалентна наступній умові:

для всіх $\varepsilon > 0$

$$\lim_{N \rightarrow \infty} P\left\{\bigcup_{n=N}^{\infty} \{|\xi_n - \xi| > \varepsilon\}\right\} = 0.$$

Послідовності випадкових величин – 2

Посилений закон великих чисел

Розглянемо послідовність випадкових величин ξ_n зі скінченими математичними сподіваннями $E\xi_n < \infty$.

Теореми, які стверджують, що різниця

$$\frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n E\xi_k$$

збігається до нуля з ймовірністю 1, називаються **посиленим законом великих чисел** (порівняймо з законом великих

чисел, що стверджують, що зазначена різниця сходиться до нуля лише за ймовірністю). Нижче наведено дві теореми про посилений закон великих чисел.

Теорема 1. Нехай ξ_n – послідовність незалежних випадкових величин, для якої $E\xi_n < \infty$, $D\xi_n < \infty$ визначені.

Якщо

$$\sum_{k=1}^{\infty} \frac{1}{k^2} D\xi_k < \infty,$$

то

$$P\left\{\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (\xi_k - E\xi_k) = 0\right\} = 1.$$

Наслідок (теорема Бореля).

Розглянемо випробування Бернуллі (біноміальний розподіл).

Позначимо γ_n число успіхів (з ймовірністю p) в n випробуваннях.

Тоді

$$P\left\{\frac{\gamma_n}{n} \rightarrow p\right\} = 1.$$

Теорема 2. Нехай ξ_n – послідовність незалежних, однаково розподілених випадкових величин, зі скінченим математичним сподіванням $E\xi_n = a$.

Тоді

$$P\left\{\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \xi_k = a\right\} = 1.$$

Центральна гранична теорема

Найпростіший варіант теореми про збіжність до нормального розподілу. Цей параграф присвячений одному з найбільш чудових результатів теорії ймовірностей: при широких припущеннях суми великої кількості незалежних малих випадкових доданків має розподіл, близький до нормального

(Гауссовского). Значення цього результату виходить далеко за рамки теорії ймовірностей. Він є теоретичною основою застосування нормального розподілу при вирішенні багатьох практичних задач. Завжди, коли можна вважати, що розглянута величина є сумою малих незалежних доданків, її розподіл мало відрізняється від нормального розподілу. Такими є, наприклад, помилки вимірювань у вимірювальних приладах, тривалість справної роботи приладу, коли вихід його з ладу є наслідком постійно діючих втомних змін, швидкості руху молекул газу, що отримуються внаслідок неодноразових зіткнень з сусідніми молекулами. Наведемо спочатку найпростіший варіант центральної граничної теореми, що відноситься до сум незалежних однаково розподілених доданків зі скінченною дисперсією.

Теорема 1. Нехай $\xi_1, \xi_2, \dots, \xi_n$ – послідовність незалежних випадкових величин, зі скінченною дисперсією $D\xi_n = \sigma^2 > 0$.

Позначимо

$$s_n = \sum_{k=1}^n \xi_k.$$

Тоді при $n \rightarrow \infty$, для будь-якого x

$$P\left\{\frac{s_n - Es_n}{\sqrt{Ds_n}} < x\right\} \rightarrow \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}y^2} dy.$$

Наслідок (інтегральна гранична теорема Муавра-Лапласа).

Розглянемо випробування Бернуллі (біноміальний розподіл).

Позначимо γ_n число успіхів (з ймовірністю p) в n випробуваннях.

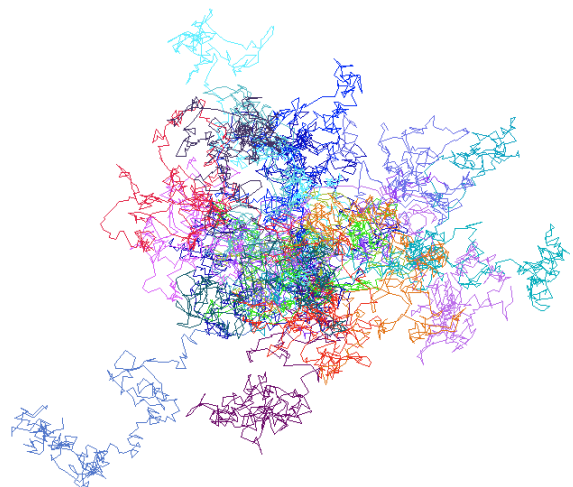
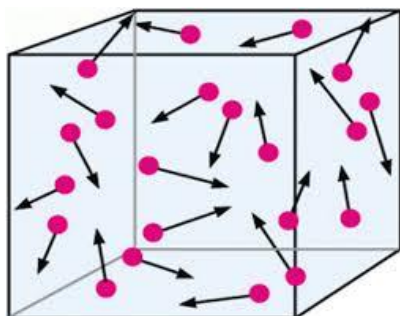
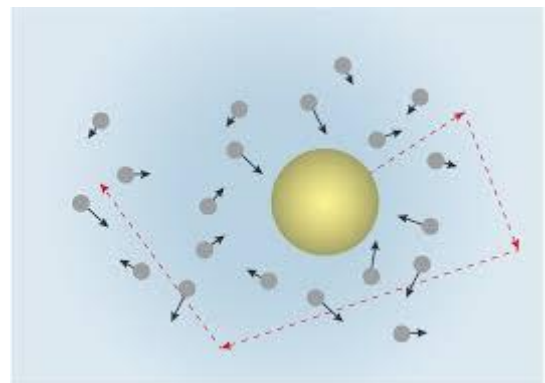
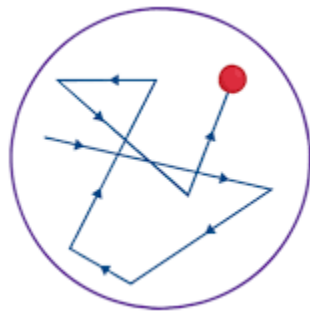
Тоді при $a < b$

$$\lim_{n \rightarrow \infty} P\left\{a < \frac{\gamma_n - np}{\sqrt{npq}} < b\right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}y^2} dy.$$

Деякі основні поняття про ПРОЦЕСИ ДИФУЗІЙНОГО ТИПУ

Броунівський рух — невпорядкований, хаотичний рух частинки під дією нерівномірних ударів молекул речовини з різних боків у розчинах. Названий на честь ботаніка Роберта Броуна, який спостерігав це явище під мікроскопом у 1827 р.

Відкриття й пояснення броунівського руху мало велике значення для фізики, оскільки було свідченням теплового руху молекул. Броун 1827 року відкрив хаотичний рух спор плауна у воді. Рух завислих частинок відбувався внаслідок руху молекул. Молекули рідини зіштовхуються з завислими у ній частинками, а отже передають їм імпульс. Таким же чином рухаються частинки фарби у воді, пилинки в променях світла (хоча на рух пилинок також впливають і мікропотоки в повітрі) тощо.



Досліди свідчать, що інтенсивність броунівського руху тим більша, чим вища температура рідини, що ще раз підтверджує безпосередній зв'язок броунівського руху з тепловим рухом молекул. Перша кількісна теорія броунівського руху з'явилася у 1905. Її автором був Альберт Ейнштейн. Він записав рівняння, яке враховувало хаотичність сили, що діє на броунівську частинку, й, розв'язавши його, отримав співвідношення

$$\langle x^2 \rangle_t = \frac{kT}{3\eta\pi a} t$$

де $\langle x^2 \rangle$ — середнє значення квадрата зміщення броунівської частинки вздовж осі X за час t , T — абсолютна температура рідини, a — радіус частинки, η — динамічна в'язкість а k — універсальна фізична константа, стала Больцмана.

1. Рівняння Ланжевена руху частинок у рідкому середовищі

Математична модель стаціонарної дифузії Гауса, визначена процесом дифузійного типу, мотивована важливим особливим випадком руху частинок у віскозному (в'язкому) рідкому середовищі, відома у фізиці як рівняння Ланжевена.

Розглянемо рівняння:

$$m\dot{v}(t) = -\beta v(t) + \text{"шум"} , \quad t \geq 0 ,$$

де m - маса частинок, $v(t)$ - швидкість частинок, β - в'язкість середовища, а компонент шуму буде додатково уточнений.

Тут складова "сили" $m\dot{v}(t)$, яка уповільнює частинку, пропорційна "швидкості" $v(t)$, а наявність "шуму" пов'язано з хаотичними зіткненнями частинки з молекулами середовища (внаслідок теплового руху останніх). У цьому формулюванні як "шум" використовується "випадковий процес"

$$\mu_t = \dot{W}_t ,$$

де $W = \{W_t , t \geq 0\}$ - стохастичний процес броунівського руху.

Відповідно до теореми Пейлі-Вінера-Зигмунда, випадкова функція W ніде не диференційована майже напевно.

Для того, щоб надати значення рівнянню:

$$m\dot{v}(t) = -\beta v(t) + \dot{W}(t), \quad t \geq 0, \quad (1)$$

відоме у фізиці як рівняння Ланжевена, слід надати імовірнісне значення похідного процесу $\dot{W}(t), t \geq 0$. Це робиться шляхом опрацювання концепції стохастичного інтегралу Іто.

Стохастичний інтеграл

Покладемо

$$\int_0^t \sigma dW_s := \sigma_t, \quad E\sigma_t = 0, \quad E\sigma_t^2 = \sigma^2 t. \quad (2)$$

Рівняння Ланжевена можна формально переписати як

$$V(t) - V(0) = -a \int_0^t V(s) ds + \sigma W_t \quad (3)$$

Лема.

$$V(t) = V(0)e^{-at} + \int_0^t e^{-a(t-s)} \sigma dW_s \quad (4)$$

Властивості розв'язків рівняння (4):

Теорема 1. За вимірюваністю вихідного значення $V(0)$ та його гауссовості

$$V(0) \sim N(0, \sigma_0), \quad \sigma_0 = \sigma^2 / 2a$$

то розв'язок рівняння Ланжевена - це процес Орнштейна-Уленбека, тобто процес Гауса $V = V(t), t \geq 0$, який має математичне сподівання 0 і функцію коваріації

$$\operatorname{cov}(V(s), V(t)) = \frac{\sigma^2}{2\alpha} e^{-\alpha|s-t|} \quad s, t \geq 0. \quad (5)$$

Тут за визначенням $\operatorname{cov}(\alpha, \beta) := E[\alpha - E\alpha][\beta - E\beta]$.

Стохастичне диференціальне рівняння:

$$dX_t = a_t(X_t)dt + \sigma_t(X_t)dW_t \quad (6)$$

Або еквівалентно,

$$X_t = X_0 + \int_0^t a_s(X_s) ds + \underbrace{\int_0^t \sigma_s(X_s) dW_s}_{\text{стохастичний інтеграл}} \quad (7)$$

Процес Іто

$$d\xi_t = a_t(\omega)dt + \sigma_t(\omega)dW_t \quad \int_0^t (dW_t)^2 = \int_0^t dt = t \quad (8)$$

Тут стохастичний інтеграл визначається наступним чином:

$$\int_0^t \varphi_s(\omega) dW_s := \Phi_t \sim N(0, \sigma_t^2(\varphi)), \quad \sigma_t^2(\varphi) := \int_0^t \varphi_s^2 ds \quad (9)$$

Процес дифузійного типу визначається рівнянням (8) з визначеннями (9), для якого:

$$a_s(\omega) = A_s(\xi), \quad \sigma_s(\omega) = B_s(\xi).$$

Формула Іто заміни змінної

$$df(\xi_t) = Lf(x)|_{\xi} dt + f'_x(\xi_t) dW_t, \quad (10)$$

де

$$Lf(x) := a(\omega)f'(x) + \frac{1}{2}\sigma^2(\omega)f''(x) \quad (11)$$

Теорема 2. Розглянемо векторне стохастичне диференціальне рівняння

$$dX_t = [a_0(t) + a_1(t)X_t]dt + b(t)dW_t, \quad x_0 = \eta, \quad (12)$$

Тоді за початкових умов та коефіцієнта інтегрованості коефіцієнтів маємо

$$X_t = \Phi_t[\eta + \int_0^t \Phi_s^{-1} a_0(s)ds] + \int_0^t \Phi_s^{-1} b(s)dW_s, \quad (13)$$

Де Φ_t - фундаментальна матриця

$$\Phi_t = \mathbb{I} + \int_0^t a_1(s) \Phi_s ds$$

Класифікація неперервних процесів

I. Неперервний процес Вінера W_t з незалежними приростами та $EW_t^2 = t$

II. Процес Орнштейна-Уленбека

$$dX_t = -aX_t dt + \sigma dW_t$$

III. Процес Іто

$$d\xi_t = a_t(\omega)dt + \sigma_t(\omega)dW_t$$

IV. Процес дифузійного типу

$$d\xi_t = a_t(\xi_t)dt + \sigma_t(\xi_t)dW_t$$

Обмеження на параметри:

$$\int_0^t a_t^2(\xi)dt < \infty \text{ з імовірністю } 1$$

$$\int_0^t \sigma_t^2(\xi)dt < \infty \text{ з імовірністю } 1$$

Задачі статистики:

Оцінка параметрів & Перевірка гіпотез

$\zeta = (\xi_1, \xi_2, \dots, \xi_n)$ - випадкова вибірка спостережень (приймає значення у просторі вибірових значень).

Задача оцінки параметра θ полягає в наступному:

побудувати наближені формули :

$$\theta \approx h(\zeta) , \quad (1)$$

де $h(x): R \rightarrow \theta$ називається статистикою; $h(\zeta)$ називається **оцінкою параметра θ** у значенні (1).

Тепер назвемо $\check{\theta}$ **оцінкою $\check{\theta} \approx h(\zeta)$ параметра θ** .

Квадратна відстань між $\check{\theta}$, θ - це також випадкова величина

$$|\check{\theta} - \theta|^2$$

її дисперсія

$$\delta^2 = E|\check{\theta} - \theta|^2$$

називається середньоквадратичною помилкою оцінки $\check{\theta}$.

Тобто $\delta^2 \in$ мірою точності оцінки $\check{\theta}$.

Визначення 1. Якщо $E\check{\theta} = \theta$, оцінка $\check{\theta}$ називається незміщеною.

Визначення 2. Послідовність оцінок $\check{\theta}_k$, $k \geq 1$ називається **спроможною**, якщо $\lim_{k \rightarrow \infty} \check{\theta}_k = \theta$. (в імовірнісному сенсі).

Приклад 8.1. Оцінка ймовірності події

$$P(A) \approx \frac{v_n(A)}{n} = \check{p} \quad (2)$$

Нехай $P(A) = p$, тоді $P(\bar{A}) = q = 1 - p$

So $E\check{p} = p$, $E|p - \check{p}|^2 = D\check{p} = \frac{pq}{n}$. Тож ця оцінка є незміщеною + спроможною.

Приклад 8.2. Теорема Мура-Лапласа

Нехай $\frac{v_n(A)}{n}$ є відносною частотою події A у серії незалежних експериментів з $P(A) = p$ і $P(\bar{A}) = q = 1 - p$.

Тоді має місце збіжність при $n \rightarrow \infty$:

$$P \left\{ a < \sqrt{\frac{n}{pq}} \left(\frac{v_n(A)}{n} - p \right) < b \right\} \rightarrow p(x) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{u^2}{2}} du$$

Іншими словами, розподіл випадкових величин

$$\zeta_n = \sqrt{\frac{n}{pq}} \left(\frac{v_n(A)}{n} - p \right)$$

збігається, при $n \rightarrow \infty$, до стандартного нормального розподілу $N(0,1)$.

Тож можна оцінити розподіл помилки оцінки

$$P(A) \approx \frac{v_n(A)}{n} = \check{p}$$

наступним чином:

$$P \left\{ |\check{p} - p| \geq c \sqrt{\frac{p(1-p)}{n}} \right\} \approx 1 - \Phi(c),$$

де

$$\Phi(c) = \frac{1}{\sqrt{2\pi}} \int_0^c e^{-\frac{u^2}{2}} du$$

Тож якщо ймовірність $\alpha = 1 - \Phi$ досить малий, тоді має місце двостороння оцінка

$$p - c \sqrt{\frac{p(1-p)}{n}} < \check{p} < p + c \sqrt{\frac{p(1-p)}{n}}$$

Це приклад так званого **інтервала довіри**.

3- σ -rule for normal (gaussian) розподіл

Для нормального розподілу мають місце такі міркування:

$x_1, x_2, \dots, x_n$ - спостереження випадкових величин $N(0, \sigma^2)$

Середнє значення вибірки

$$\bar{x} = \frac{x_1 + \dots + x_n}{n}$$

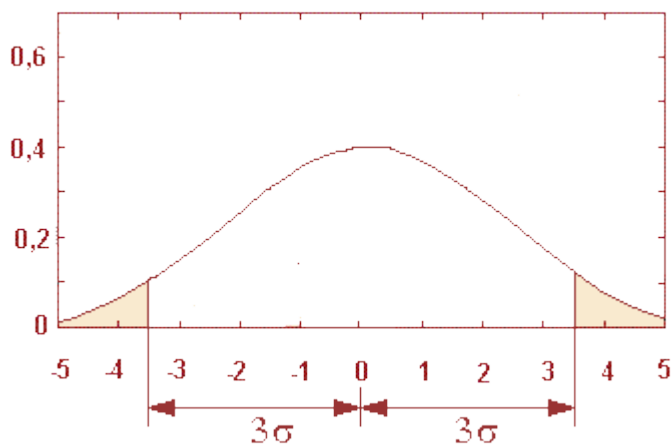
Тоді

$$P\{-\sigma < \bar{x} < \sigma\} = \Phi\left(\frac{1}{\sqrt{2}}\right) \approx 0,683;$$

$$P\{-2\sigma < \bar{x} < 2\sigma\} = \Phi(\sqrt{2}) \approx 0,954;$$

$$P\{-3\sigma < \bar{x} < 3\sigma\} = \Phi\left(\frac{3}{\sqrt{2}}\right) \approx 0,997;$$

Висновок: майже напевнено, що значення відхилення похибки випадкових величин від їх математичного сподівання не відхиляться на величину, більшу за 3σ .

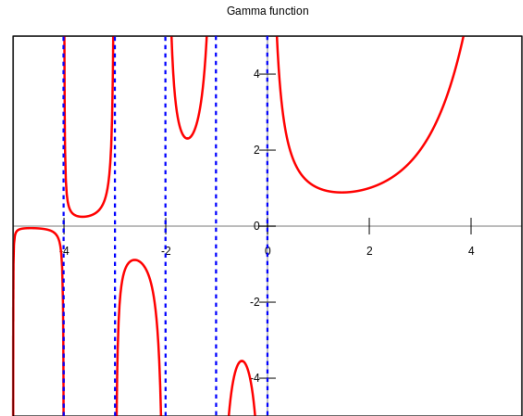


Деякі важливі ймовірнісні розподіли

Визначення. Г-функція

$$\Gamma(n) = (n-1)! \\ (\text{дискретний випадок})$$

$$\Gamma(t) = \int_0^{\infty} x^{t-1} e^{-x} dx \\ (\text{неперервний випадок})$$



Розподіл Стюдента (t-розподіл)

Припустимо, що випадкові величини $\xi, \xi_1, \xi_2, \dots, \xi_n$ є незалежними, нормально розподіленими стандартними випадковими величинами $N(0,1)$. Тоді розподілом Стюдента називається функція розподілу випадкові величини :

$$\frac{\xi}{\sqrt{\frac{1}{n} \sum_{k=1}^n \xi_k^2}}$$

Введемо позначення $\eta := \sqrt{\frac{1}{n} \sum_{k=1}^n \xi_k^2}$

Випадкові величини η мають таку ймовірнісну щільність:

$$f_{\eta}(x) = \begin{cases} \frac{n^{\frac{n}{2}}}{2^{\frac{n}{2}-1} \Gamma(\frac{n}{2})} x^{n-1} e^{-\frac{nx^2}{2}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Ми можемо використати цей факт для отримання ймовірнісної щільності розподілу Стюдента

$$f_t(z) = \frac{n^{\frac{n}{2}}}{2^{\frac{n}{2}-1} \Gamma(\frac{n}{2})} \int_0^{\infty} y^n e^{-\frac{y}{2}(z^2+n)} dy$$

$$= \frac{1}{\sqrt{n\pi}} \cdot \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \cdot \frac{1}{(1 + \frac{z^2}{n})^{\frac{n+1}{2}}}$$

Фішера – Снедекора (F-розподіл)

Припустимо, що випадкові величини

$\xi_1, \xi_2, \dots, \xi_m$ і $\eta_1, \eta_2, \dots, \eta_n$ є незалежними, нормально розподіленими стандартними $N(0,1)$.

Визначення Розподіл Фішера – Снедекора (F-розподіл) з (m,n) ступенями свободи називається розподіл таких випадкових величин:

$$\gamma = \frac{\frac{1}{m} \sum_{k=1}^m \xi_k^2}{\frac{1}{n} \sum_{k=1}^n \eta_k^2}$$

Прямі розрахунки дають наступну формулу для ймовірнісної щільності F-розподілу:

$$f_\gamma(x) = \begin{cases} (m/n)^{\frac{m}{2}} \cdot \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} \cdot x^{\frac{m}{2}-1} (1 + \frac{m}{n}x)^{-\frac{m+n}{2}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Хі-квадрат (χ^2 розподіл)

Визначення χ^2 -розподіл з n ступенями свободи називається розподіл таких випадкових величин:

$$\chi^2 = \xi_1^2 + \dots + \xi_n^2,$$

де $\xi_1, \xi_2, \dots, \xi_n$ є незалежними, нормально розподіленими стандартними $N(0,1)$.

Безпосередні розрахунки дають таку формулу функції ймовірнісної щільності:

$$f_{\chi^2}(x) = \begin{cases} \frac{1}{\Gamma(\frac{n}{2}) 2^{\frac{n}{2}}} \cdot x^{\frac{n}{2}-1} e^{-\frac{1}{2}x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

В статистичній механіці значення $\chi^2 = \xi_1^2 + \xi_2^2 + \xi_3^2$ з'являється як квадрат швидкості частинки, якщо припустити, що швидкісні проекції на осі мають нормальний розподіл. Відповідний розподіл називається розподілом Максвелла.

Статистика оцінки параметрів та довірчі інтервали

Розглянемо випадкову вибірку спостережень $\xi_1, \xi_2, \dots, \xi_n$ -
Середнє значення вибірки дорівнює

$$\bar{\xi} = \frac{\xi_1 + \dots + \xi_n}{n} \quad (3)$$

Дисперсія вибірки

$$\bar{\sigma}^2 = \frac{1}{n} \sum_{k=1}^n (\xi_k - \bar{\xi})^2 \quad (4)$$

Теорема 1.

Нехай $\xi_1, \xi_2, \dots, \xi_n$ - послідовність незалежних, однаково розподілених випадкових величин зі стандартним розподілом $N(0,1)$.

Тоді

- 1) середнє значення (3) і дисперсія вибірки (4) є незалежними випадковими величинами;
- 2) випадкові величини $(n-1)\bar{\sigma}^2$ мають χ^2 -розподіл з $(n-1)$ степенями свободи.

Теорема 2.

Нехай $\xi_1, \xi_2, \dots, \xi_n$ - послідовність незалежних, однаково розподілених випадкових величин з нормальним розподілом $N(a, \sigma^2)$.

Тоді випадкова величина

$$\tau = \frac{\sqrt{n}(\bar{\xi} - a)}{\bar{\sigma}} = \frac{\sqrt{n}(\bar{\xi} - a)}{\sqrt{\frac{1}{n-1} \sum_{k=1}^n (\xi_k - \bar{\xi})^2}}$$

має t -розподіл (Ст'юдента) з $(n-1)$ степенями свободи.

Нагадування

$\zeta = (\xi_1, \xi_2, \dots, \xi_n)$ - випадкова вибірка спостережень

Оцінка параметра $\theta \approx h(\zeta) = \check{\theta}$

Мірою точності оцінки $\check{\theta}$ є величина

$$\delta^2 = E|\check{\theta} - \theta|^2.$$

Якщо $E\check{\theta} = \theta$, оцінка $\check{\theta}$ називається незміщеною.

Якщо $\lim_{k \rightarrow \infty} \check{\theta}_k = \theta$, оцінка $\check{\theta}$ називається спроможною.

Оцінка $\check{\theta}_1$ називається ефективною, якщо для будь-якої

іншої оцінки $\check{\theta}_2$ має місце нерівність

$$E|\check{\theta}_1 - \theta|^2 \leq E|\check{\theta}_2 - \theta|^2.$$

Достатні статистики

Нехай ζ - випадковий вектор зі значеннями в R^n , щільність розподілу якого $f(x, \theta)$ залежить від невідомого параметра θ , $\theta \in \Theta$, де Θ – підмножина в R^s . У ряді задач параметр θ можна вважати випадковим вектором, які приймають в розглянутих реалізаціях вектора ζ фіксоване значення.

Інформація, яку несе в собі реалізація вектора ζ про значення параметра θ , міститься в умовному розподілі вектора θ при даному ζ . Припустимо, що для вектора θ існує щільність розподілу $g(\theta)$.

Щільність $g(\theta)$ називають апіорною щільністю розподілу параметра θ . Умовна щільність розподілу параметра θ при гіпотезі $\zeta = x$ задається формулою Байєса

$$f(\theta|x) = \frac{f(x, \theta)g(\theta)}{\int_{\Theta} f(x, \theta)g(\theta)d\theta}$$

Зауважимо, що якщо щільність $f(x, \theta)$ розподілу вектора має вигляд

$$f(x, \theta) = \varphi(T(x), \theta)h(x), \quad (1)$$

де функції $h(x)$ і $T(x)$ не залежить від θ , а $T(x)$ – деяка вимірنا скалярна або векторна функція, то

$$f(\theta|x) = \frac{\varphi(T(x), \theta)g(\theta)}{\int_{\theta} \varphi(T(x), \theta)g(\theta)d\theta},$$

тобто умовний розподіл параметра θ при будь апіорної щільності $g(\theta)$ залежить тільки від θ і $T(x)$.

Неважко помітити, що якщо «апостеріорний» розподіл $f(\theta|x)$ параметра θ залежить тільки від θ і деякої статистики $T(x)$: $f(\theta|x) = K(T(x), \theta)$ для деякої апіорної щільності $g(\theta)$ має вигляд (1). Дійсно, в цьому випадку

$$f(\theta|x) = \frac{1}{g(\theta)} K(T(x), \theta) \int_{\theta} f(x, \theta)g(\theta)d\theta = \varphi(T(x), \theta)h(x),$$

де

$$\varphi(T, \theta) = K(T, \theta) / g(\theta); \quad h(x) = \int_{\theta} f(x, \theta)g(\theta)d\theta.$$

Означення. Якщо сімейство щільностей розподілу $f(x, \theta)$, $\theta \in \theta$, допускає зображення (1), то вектор $T(\zeta)$ називають достатньою статистикою для даного сімейства розподілів (або для параметра θ).

Поняття достатньої статистики можна пояснити так: статистика $T(\zeta)$ достатня для сімейства щільності $f(x, \theta)$, $\theta \in \theta$ вектора ζ тоді і тільки тоді, коли умовний розподіл вектора ζ при заданому $T(\zeta)$ не залежить від θ .

Формулу (1) називають **факторізаційним поданням** щільності розподілу.

Відзначимо наступну важливу властивість достатньої статистики.

Нехай $h(\zeta)$ - незміщена оцінка параметра θ і $T = T(\zeta)$ – достатня статистика. Тоді існує незміщена оцінка параметра θ , що є функцією від достатньої статистики та має не меншу ефективність, ніж $h(\zeta)$.

Теорема 1. Нехай T - достатня статистика для сімейства щільностей $f(x, \theta)$, $\theta \in \theta$, $h(\zeta)$ – незміщена оцінка параметра θ . Покладемо $E[h(\zeta)|T] = K(T)$.

Тоді $K(T)$ – незміщена оцінка параметра θ з ефективністю
 $E[K(T) - \theta]^2 \leq E[h(\zeta) - \theta]^2$
 причому знак рівності тут має місце тільки тоді, коли
 $h(\zeta) = K(T)$ з ймовірністю 1.

Наслідок. Якщо для параметра θ існує оцінка мінімальної дисперсії, то вона є функцією від достатньої статистики.

Метод максимальної правдоподібності

Нехай спостерігається випадковий вектор ξ , щільність розподілу якого $f(x, \theta)$ залежить від параметра θ , що приймає значення з деякої множини $\Theta \in R^s$.

Метод оцінки параметра θ , який отримав назву методу максимальної правдоподібності, заснований на такому інтуїтивному уявленні: в більшій частині випадків у експерименті спостерігається те значення вектора ξ , при якому щільність ймовірності $f(x, \theta)$ близька до максимального значення.

Таким чином, в якості оцінки $\check{\theta}$ параметра θ природно прийняти розв'язок рівняння

$$f(\xi, \check{\theta}) = \max_{\theta} f(\xi, \theta) \quad (1)$$

якщо це рівняння має сенс і його розв'язок існує.

Розв'язок рівняння (1) називаються **оцінками максимальної правдоподібності**.

У багатьох випадках замість функції $f(x, \theta)$ простіше розглядати функцію $\ln f(x, \theta)$, яка досягає максимуму в тих же точках, що і $f(x, \theta)$. Якщо функція $f(x, \theta)$ неперервно диференційовна по θ , то для розв'язання рівняння (1) можна спочатку знайти стаціонарні точки функції $f(x, \theta)$, а потім порівняти значення функції $f(\xi, \theta)$ в цих та граничних точках множини $\Theta \in R^s$ і вибрати точку максимуму $f(x, \theta)$. Для відшукування стаціонарних точок мають місце

рівняння

$$\frac{\partial \ln f(x, \theta)}{\partial \theta_i} = 0, \quad i = 1, 2, \dots, s.$$

Ці рівняння називають **рівняннями максимальної правдоподібності**.

Метод максимальної правдоподібності в деяких випадках має оптимальні властивості. У разі великого числа n незалежних спостережень при достатній регулярності щільності розподілу він асимптотично оптимальний, але практичні застосування методу максимальної правдоподібності часто призводять до вельми складних обчислень. І до того ж при невеликих значеннях n не завжди ясно, наскільки цей метод доцільний.

Відзначимо кілька властивостей оцінок максимальної Правдоподібності

а) Якщо для сімейства щільностей $f(x, \theta)$ існує достатня статистика, то розв'язок рівняння правдоподібності (коли воно існує і єдино) є функцією від достатньої статистики.

Дійсно, якщо $f(x, \theta) = h(x)g(T(x), \theta)$, то рівняння (1) набирає вигляду

$$g(t, \check{\theta}) = \max_{\theta} g(t, \theta)$$

Звідки одержуємо $\check{\theta} = \varphi(t) = \varphi(T(\zeta))$.

б) Якщо виконані умови регулярності (існують похідні $g'_{\theta}(t, \theta)$, $g''_{\theta\theta}(t, \theta)$), причому

$$\begin{aligned} E_{\theta} \left| \frac{\partial}{\partial \theta} \ln f(\zeta, \theta) \right| &< \infty, \\ E_{\theta} \left| \frac{\partial^2}{\partial \theta^2} \ln f(\zeta, \theta) \right| &< \infty, \\ E_{\theta} \left| \frac{\partial}{\partial \theta} \ln f(\zeta, \theta) \right|^2 &< \infty, \end{aligned}$$

і існує ефективна оцінка $h(\zeta)$ одновимірного параметра θ , то $h(\zeta)$ є також оцінкою максимальної правдоподібності.

в) Нехай функція $h(\theta)$, $\theta \in \Theta$ взаємно однозначно відображає Θ в Q , де Q – деяка множина в R^s . Якщо $\check{\theta}$ – оцінка максимальної правдоподібності параметра θ сімейства щільності

$$f(x, \theta), \theta \in \Theta$$

то $h(\check{\theta})$ є оцінкою максимальної правдоподібності параметра $h(\theta)$.

Асимптотичні властивості оцінок максимальної правдоподібності.

Припустимо, що оцінка параметра θ розподілу $g(x, \theta)$ проводиться за вибіркою обсягу n . Розглянемо асимптотичну поведінку оцінок $\check{\theta} = \check{\theta}_n$ параметра θ , отриманих за методом максимальної правдоподібності. Покажемо, що при досить широких припущеннях оцінки максимальної правдоподібності є консистентними (спроможними), а при кілька більш жорстких – асимптотично ефективними й асимптотично нормальними. Зробимо кілька попередніх зауважень про функції правдоподібності. Якщо $\xi_1, \xi_2, \dots, \xi_n$ – компоненти вибіркового вектора ζ , то функція правдоподібності має вигляд

$$\Lambda(\theta) = \prod_{k=1}^n g(\xi_k, \theta).$$

Припустимо, що область значень Θ параметра θ компактна і справжнє значення θ параметра належить множині Θ . Розглянемо відношення

$$\frac{\Lambda(\lambda)}{\Lambda(\theta)}, \lambda \in \Theta$$

Оскільки функція $\ln x$, $x > 0$, строго опукла, то з нерівності Йенсена випливає така нерівність

$$E_{\theta} \ln \frac{\Lambda(\lambda)}{\Lambda(\theta)} \leq \ln E_{\theta} \frac{\Lambda(\lambda)}{\Lambda(\theta)} = 0,$$

оскільки

$$E_{\theta} \frac{\Lambda(\lambda)}{\Lambda(\theta)} = \int \prod_{k=1}^n g(x_k, \lambda) dx_1 \dots dx_n = 1.$$

При цьому, якщо $\Lambda(\lambda) \neq \Lambda(\theta)$ з позитивною ймовірністю, то

$$E_{\theta} \frac{\Lambda(\lambda)}{\Lambda(\theta)} < 0 \quad \text{і} \quad E_{\theta} \ln \Lambda(\lambda) < E_{\theta} \ln \Lambda(\theta)$$

з ймовірністю 1, якщо

$$l(\lambda) = E_{\theta} \ln g(x, \lambda) g(x, \theta)$$

Перейдемо до доведення спроможності оцінок методу максимальної правдоподібності. Покладемо

$$I(\theta, V_{\lambda}) = E_{\theta} \left\{ \inf_{\lambda' \in V_{\lambda}} \ln \frac{g(\zeta, \theta)}{g(\zeta, \lambda')} \right\},$$

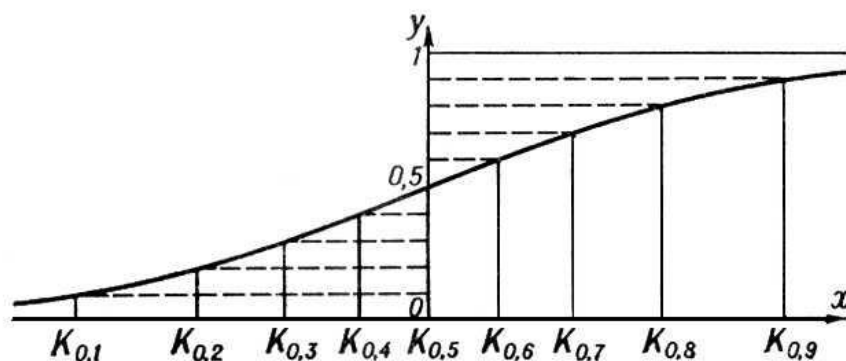
де V_{λ} - окіл точки $\lambda \in \Theta$. Припустимо, що виконані умови:

- 1) якщо $\theta \neq \lambda$, то $g(x, \theta) \neq g(x, \lambda)$ на множині позитивної лебегової міри;
- 2) функція $g(x, \theta)$ неперервна по θ ;
- 3) для кожної точки $\lambda \in \Theta$ існує окіл V_{λ} такий, що $I(\theta, V_{\lambda}) > -\infty$.

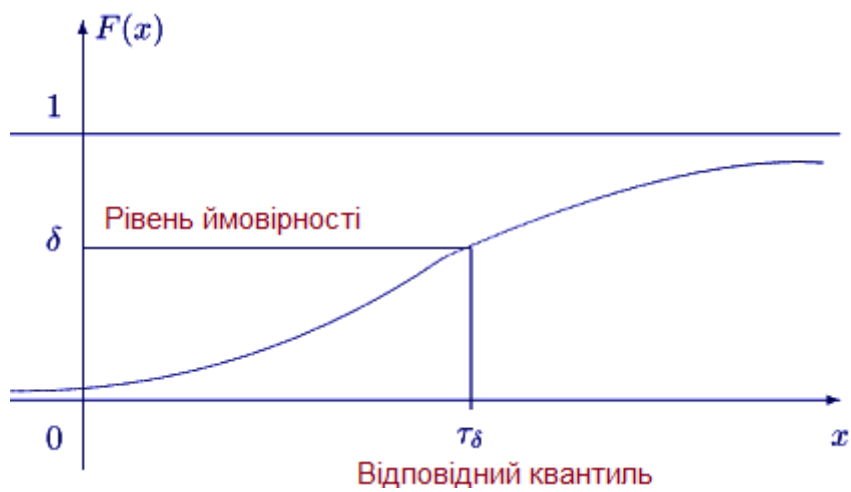
Теорема 1. Якщо виконані умови 1), 2) і 3), то оцінки максимальної правдоподібності сильно спроможні.

Нагадування:

Квантиль в математичній статистиці - значення, яке задана випадкова величина не перевищує з фіксованою ймовірністю.



Квантиль, що відповідає рівню ймовірності δ , є таким значенням випадкової величини ξ , що $\xi = \tau_{\delta}$: $F(\tau_{\delta}) = \delta$.



$$F(\tau_\delta) = \delta \quad \rightarrow \quad \tau_\delta = F^{-1}(\delta)$$

Теорема Мура-Лапласа (правило трьох сігм як приклад довірчих інтервалів)

Нехай $\frac{v_n(A)}{n}$ є відносною частотою події A у серії незалежних експериментів з $P(A) = p$ і $P(\bar{A}) = q = 1 - p$.

Тоді має місце збіжність при $n \rightarrow \infty$:

$$P\left\{a < \sqrt{\frac{n}{pq}}\left(\frac{v_n(A)}{n} - p\right) < b\right\} \rightarrow p([a, b]) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{u^2}{2}} du$$

Тобто, розподіл випадкових величин

$$\zeta_n = \sqrt{\frac{n}{pq}}\left(\frac{v_n(A)}{n} - p\right)$$

збігається, при $n \rightarrow \infty$, до стандартного нормального розподілу $N(0,1)$.

Значить розподіл помилки оцінки

$$P(A) \approx \frac{v_n(A)}{n} = \check{p}$$

можна оцінити наступним чином:

$$P\left\{|\check{p} - p| \geq c \sqrt{\frac{p(1-p)}{n}}\right\} \approx 1 - \Phi(c),$$

де

$$\Phi(c) = \frac{1}{\sqrt{2\pi}} \int_0^c e^{-\frac{u^2}{2}} du$$

Середнє значення вибірки

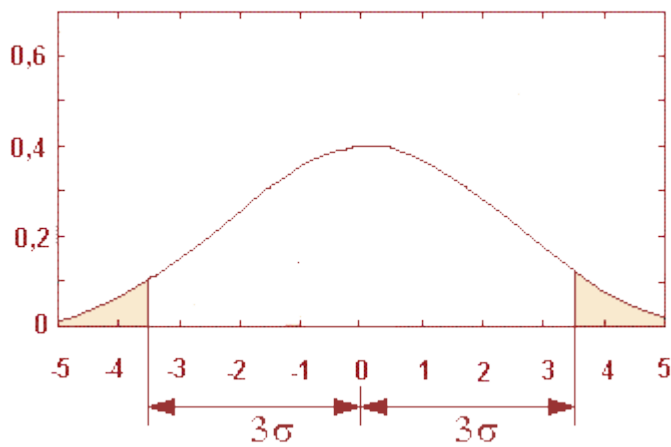
$$\bar{x} = \frac{x_1 + \dots + x_n}{n}$$

Тоді

$$P\{-\sigma < \bar{x} < \sigma\} = \Phi\left(\frac{1}{\sqrt{2}}\right) \approx 0,683;$$

$$P\{-2\sigma < \bar{x} < 2\sigma\} = \Phi(\sqrt{2}) \approx 0,954;$$

$$P\{-3\sigma < \bar{x} < 3\sigma\} = \Phi\left(\frac{3}{\sqrt{2}}\right) \approx 0,997;$$



Нагадування

$\zeta = (\xi_1, \xi_2, \dots, \xi_n)$ - випадкова вибірка спостережень

Оцінка параметра $\theta \approx h(\zeta) = \check{\theta}$

Мірою точності оцінки $\check{\theta}$ є величина

$$\delta^2 = E|\check{\theta} - \theta|^2.$$

Якщо $E\check{\theta} = \theta$, оцінка $\check{\theta}$ називається незміщеною.

Якщо $\lim_{k \rightarrow \infty} \check{\theta}_k = \theta$, оцінка $\check{\theta}$ називається спроможною.

Оцінка $\check{\theta}_1$ називається ефективною, якщо для будь-якої іншої оцінки $\check{\theta}_2$ має місце нерівність

$$E|\check{\theta}_1 - \theta|^2 \leq E|\check{\theta}_2 - \theta|^2.$$

Довірчі інтервали

Довірчі інтервали є випадковими інтервалами, всередині яких з ймовірністю γ , близькою до 1, міститься точне значення оцінюваного параметра. Величина γ називається довірчою ймовірністю.

При заданому γ довжина довірчого інтервалу характеризує точність, з якою локалізовано значення параметра. Тому природно прагнення вибрати серед різних довірчих інтервалів при заданій довірчій ймовірності інтервал найкоротшою довжини. Такий вибір забезпечує найбільш точну локалізацію параметра.

Нехай ζ - вибірковий вектор зі значеннями R^n , P_θ - розподіл вектора ζ , залежне від невідомого параметра θ , $\check{\theta}(\zeta)$ - деяка оцінка параметра θ .

Одновимірний параметр.

Розглянемо дві дійсні функції $\underline{\theta}(x)$, $\bar{\theta}(x)$ зі значеннями в R^1 , які не залежать від невідомих параметрів розподілу вектора ζ і такі, що $\underline{\theta}(x) < \bar{\theta}(x)$, $x \in R^n$.

Якщо

$$P_\theta\{\underline{\theta}(x) < \bar{\theta}(x)\} \geq \gamma, \quad \theta \in \Theta,$$

то випадковий інтервал $(\underline{\theta}(x), \bar{\theta}(x))$ називають γ -довірчим інтервалом, число γ ($0 < \gamma < 1$) – довірчим рівнем, $\underline{\theta}(x)$ і $\bar{\theta}(x)$ - нижньою і верхньою довірчими межами, відповідно.

У ряді випадків представляє інтерес тільки нижня або тільки верхня довірча межа для θ . Вони відповідають вибору

$\underline{\theta}(x) = -\infty$ і $\bar{\theta}(x) = +\infty$ відповідно. Такі інтервали називають верхнім (відповідно нижнім) довірчим інтервалом. Довірчі інтервали мають наступний сенс.

Будемо користуватися таким статистичним правилом:

якщо результати вибірки описуються вектором ζ , то невідоме значення параметра ϑ лежить всередині інтервалу $(\underline{\theta}(\zeta), \bar{\theta}(\zeta))$.

Це правило в деяких випадках виявляється помилковим.

Однак, якщо його застосовувати багаторазово в різних, не залежних одне від іншого ситуаціях (при яких, все ж P_0 - розподіл вибіркового вектора ζ), то запропоноване правило виявиться помилковим при γ , близькому до 1, тільки в малому числі випадків, а саме, приблизно в $100(1 - \gamma)\%$ загального числа випадків застосування цього правила.

Довірчі інтервали часто можна будувати таким

чином. Припустимо для простоти, що розподіл статистики

$\check{\theta}(\zeta)$ неперервний. Для кожного θ знайдемо такі величини $\theta_1 = \theta_1(\theta)$ і $\theta_2 = \theta_2(\theta)$, що $P_\theta(\theta_1 < \check{\theta}(\zeta) < \theta_2) = \gamma$. Щоб ця процедура була однозначною, часто вибирають θ_1 і θ_2 , виходячи з вимог

$$P_\theta(\check{\theta}(\zeta) \leq \theta_1) = P_\theta(\check{\theta}(\zeta) \geq \theta_2) = \frac{1}{2}(1 - \gamma).$$

Позначимо через $\mathcal{D}$ підмножину $\Theta \times \Theta$:

$$\mathcal{D} = \{(\theta, \theta') : \theta_1(\theta) < \theta' < \theta_2(\theta)\}.$$

При будь-якому $\theta \in \Theta$ подія $\{(\theta, \check{\theta}(\zeta)) \in \mathcal{D}\}$

має P_θ -ймовірність γ . Розглянемо тепер при фіксованому θ' множину $\mathcal{D}(\theta') = \{\theta : (\theta, \theta') \in \mathcal{D}\}$.

Множина $\mathcal{D}(\check{\theta}(\zeta))$ є випадковою, і подія $\theta \in \mathcal{D}(\check{\theta}(\zeta))$ відбувається

тоді і тільки тоді, коли $\check{\theta}(\zeta) \in (\theta_1(\theta), \theta_2(\theta))$ і, отже, при кожному $\theta_1(\theta)$ має ймовірність γ . Якщо множина $\mathcal{D}(\theta')$ є інтервалом ($\theta \in \Theta$),

то $\mathcal{D}(\check{\theta}(\zeta))$ є γ -довірчим інтервалом для параметра θ .

Ясно, що існують різні способи побудови довірчих інтервалів.

Наприклад, в попередній побудові вони

залежали від вибору оцінки $\check{\theta}(\zeta)$. Природно бажання серед різних довірчих інтервалів, які відповідають цьому довірчому рівню, вибрати якомога коротші.

Припустимо, що розглядаються незміщені і приблизно

нормальні оцінки θ . Тоді введені вище інтервали $\mathcal{D}(\check{\theta}(\zeta))$ будуть

тим коротше, чим менше дисперсія оцінки.

Таким чином, ефективні і асимптотично ефективні оцінки призводять до найкоротшим або асимптотично найкоротшим довірчим інтервалом.

У разі, коли вибірка побудована по великому числу незалежних однаково розподілених випадкових величин, для наближеної побудови довірчих інтервалів можна припустити наступний спосіб. Відомо, що в разі регулярності щільності розподілу $f(x, \theta)$ вибіркового вектора ζ величина

$\frac{\partial}{\partial \theta} \ln f(\zeta, \theta)$ розподілена асимптотично нормально.

Випадкова величина

$$\eta = \frac{\frac{\partial}{\partial \theta} \ln f(\zeta, \theta)}{\sqrt{D\left(\frac{\partial}{\partial \theta} \ln f(\zeta, \theta)\right)}},$$

то величина η асимптотично нормальна параметрами $(0, 1)$.

Нехай c_γ визначається з умови $P(|\xi| < c_\gamma) = \gamma$,

де ξ - стандартна гаусовська величина.

Якщо величина $\eta = \eta(\theta)$ є монотонною функцією від θ , то вирішуючи нерівність $|\eta| \leq c_\gamma$ щодо θ , ми для θ отримуємо γ -довірчий інтервал.

Нормальна лінійна регресія

У практичних застосуваннях методів математичної статистики часто зустрічається завдання, найпростіший варіант якої зараз буде розглянуто.

Передбачається, що між величинами $y, x_1, \dots, x_d$ існує лінійна залежність виду

$$y = \alpha_0 + \beta_1 x_1 + \dots + \beta_d x_d \quad (1)$$

яка називається **лінійною регресією**.

Тут $x_1, \dots, x_d$ – деякі випадкові змінні, а коефіцієнти $\alpha_0, \beta_1, \dots, \beta_d$ називаються **коефіцієнтами регресії**, які невідомі і повинні бути визначені експериментально.

Здійснюється n незалежних експериментів. У кожному з них величини $x_1, \dots, x_d$ приймають певні відомі значення $x_{1k}, \dots, x_{dk}$, $k = 1, \dots, n$ які, взагалі кажучи, різні в різних експериментах.

У кожному експерименті вимірюється нормальна випадкова величина ξ_k , $k = 1, \dots, n$ із середнім значенням

$$E\xi_k = \alpha_0 + \beta_1 x_{1k} + \dots + \beta_d x_{dk}, \quad k = 1, \dots, n \quad (2)$$

з постійною, але невідомою дисперсією σ^2 .

Таким чином, розглянута вибірка $\zeta = (\xi_1, \xi_2, \dots, \xi_n)$, що утворена послідовністю незалежних нормальних випадкових величин $\xi_1, \xi_2, \dots, \xi_n$ з дисперсією σ^2 і різним середнім, що має вигляд (2).

Оцінюються коефіцієнти $\alpha_0, \beta_1, \dots, \beta_d$.

Розглянемо спочатку випадок $d = 1$. В такому разі регресія називається простою. Запишемо рівняння регресії (1) у формі

$$y = \alpha + \beta(x - \bar{x}), \quad \bar{x} = \frac{1}{n} \sum_{k=1}^n x_k$$

Функція правдоподібності вибірки має вигляд

$$f(\zeta, \alpha, \beta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{k=1}^n (\xi_k - \alpha - \beta(x_k - \bar{x}))^2 \right\}.$$

Для оцінки трьох невідомих параметрів α, β, σ^2 скористаємося методом максимальної правдоподібності.

Щоб відшукати максимум функції правдоподібності, мінімізуємо квадратическую форму

$$Q(\alpha, \beta) = \sum_{k=1}^n (\xi_k - \alpha - \beta(x_k - \bar{x}))^2,$$

а потім знаходимо максимум по функції θ

$$\varphi(\theta) = (2\pi\theta)^{-\frac{n}{2}} \exp \left\{ -\frac{Q^*}{2\theta} \right\}, \quad Q^* = \min_{\alpha, \beta} Q(\alpha, \beta).$$

Оскільки (покладемо $\bar{\xi} = \frac{1}{n} \sum_{k=1}^n \xi_k$)

$$\begin{aligned}
Q(\alpha, \beta) &= \sum_k [(\bar{\xi} - \alpha) + (\xi_k - \bar{\xi}) - \beta(x_k - \bar{x})]^2 = \\
&= n[(\bar{\xi} - \alpha)^2 + \beta^2 S_1^2 + S_2^2 - 2\beta R_{12}] = \\
&= n\left[(\bar{\xi} - \alpha)^2 + S_1\beta - \frac{R_{12}}{S_1} + S_2^2 - \frac{R_{12}^2}{S_1^2}\right],
\end{aligned}$$

де

$$\begin{aligned}
S_1^2 &= \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2, \quad S_2^2 = \frac{1}{n} \sum_{k=1}^n (\xi_k - \bar{\xi})^2, \\
R_{12} &= \frac{1}{n} \sum_{k=1}^n (\xi_k - \bar{\xi})(x_k - \bar{x}),
\end{aligned}$$

то мінімум Q^* функції $Q(\alpha, \beta)$ досягається при умовах

$$\alpha = \check{\alpha}, \quad \beta = \check{\beta}, \quad \check{\alpha} = \bar{\xi}, \quad \check{\beta} = \frac{R_{12}}{S_1^2}$$

і дорівнює

$$\begin{aligned}
Q^* &= \min_{\alpha, \beta} Q(\alpha, \beta) = \sum_k [\xi_k - \check{\alpha} - \check{\beta}(x_k - \bar{x})]^2 = nS_2^2(1 - \check{\rho}^2), \\
\check{\rho} &= \frac{R_{12}}{S_1 S_2} = \frac{\sum_{k=1}^n (\xi_k - \bar{\xi})(x_k - \bar{x})}{\sqrt{\sum_{k=1}^n (\xi_k - \bar{\xi})^2 \sum_{k=1}^n (x_k - \bar{x})^2}}.
\end{aligned}$$

Для визначення значення $\check{\theta}$, при якому функція $\varphi(\theta)$

досягає максимуму, маємо рівняння

$$(\ln \varphi(\check{\theta}))' = 0, \quad \text{або} \quad \frac{n}{2\check{\theta}} - \frac{Q^*}{2\check{\theta}^2} = 0$$

звідки $\check{\theta} = \check{\sigma}^2 = \frac{1}{n} Q^*$.

Таким чином, в якості оцінки величини σ^2 знаходимо величину

$$\check{\sigma}^2 = S_2^2(1 - \check{\rho}^2).$$

Отже, отримано оцінки максимальної правдоподібності $\check{\alpha}$, $\check{\beta}$, $\check{\sigma}^2$ параметрів α, β, σ^2 :

$$\check{\alpha} = \bar{\xi} = \frac{1}{n} \sum_{k=1}^n \xi_k; \quad \check{\beta} = \frac{\sum_{k=1}^n (\xi_k - \bar{\xi})(x_k - \bar{x})}{\sum_{k=1}^n (x_k - \bar{x})^2}; \quad (3)$$

$$\check{\sigma}^2 = \sum_k [\xi_k - \check{\alpha} - \check{\beta}(x_k - \bar{x})]^2. \quad (4)$$

Зазначимо довірчі інтервали для цих параметрів. Зауважимо, що

$$E\check{\alpha} = \frac{1}{n} \sum_{k=1}^n E\xi_k = \frac{1}{n} \sum_{k=1}^n (\alpha + \beta(x_k - \bar{x})) = \alpha .$$

Оскільки $E(\xi_k - \bar{\xi}) = \beta(x_k - \bar{x})$, то

$$E\check{\beta} = \frac{1}{n} \sum_{k=1}^n E\xi_k = \frac{1}{S_1^2} \sum_{k=1}^n (x_k - \bar{x}) E(\xi_k - \bar{\xi}) = \beta .$$

Таким чином, $\check{\alpha}$ і $\check{\beta}$ є незміщеними оцінками параметрів α і β .

Враховуючи незалежність величин ξ_k , маємо:

$$D\check{\alpha} = \frac{1}{n^2} \sum_{k=1}^n D\xi_k = \frac{\sigma^2}{n}$$

З огляду на те, що

$$\check{\beta} = \frac{1}{nS_1^2} \sum_{k=1}^n (x_k - \bar{x})(\xi_k - \bar{\xi}) = \frac{1}{nS_1^2} \sum_{k=1}^n (x_k - \bar{x})\xi_k ,$$

отримаємо

$$D\check{\beta} = \frac{1}{n^2 S_1^4} \sum_{k=1}^n (x_k - \bar{x})^2 \sigma^2 = \frac{\sigma}{nS_1^2} .$$

Оскільки величини $\eta_k = \xi_k - \alpha - \beta(x_k - \bar{x})$, $k = 1, \dots, n$ – незалежні і нормально розподілені з параметрами $(0, \sigma^2)$, то квадратична форма $\frac{1}{\sigma^2} Q(\alpha, \beta)$ має χ^2 -розподіл з n степенями свободи.

Неважко перевірити справедливість рівності

$$\begin{aligned} Q(\alpha, \beta) &= \sum_{k=1}^n [\xi_k - \check{\alpha} - \check{\beta}(x_k - \bar{x}) + (\check{\alpha} - \alpha) + (\check{\beta} - \beta)(x_k - \bar{x})]^2 = \\ &= \sum_{k=1}^n [\xi_k - \check{\alpha} - \check{\beta}(x_k - \bar{x})]^2 + n(\check{\alpha} - \alpha)^2 + nS_1^2(\check{\beta} - \beta)^2 . \end{aligned}$$

Тому величини

$$\frac{1}{\sigma^2} \sum_{k=1}^n [\xi_k - \check{\alpha} - \check{\beta}(x_k - \bar{x})]^2 = \frac{n\check{\sigma}^2}{\sigma^2} , \quad \frac{\sqrt{n}(\check{\alpha} - \alpha)}{\sigma} , \quad \frac{\sqrt{n}S_1(\check{\beta} - \beta)}{\sigma}$$

незалежні і перша з них має χ^2 -розподіл з n степенями свободи.

Далі, величини

$$\frac{\sqrt{n}(\check{\alpha} - \alpha)}{\sigma}, \quad \frac{\sqrt{n}S_1(\check{\beta} - \beta)}{\sigma}$$

є лінійними комбінаціями незалежних нормальних величин і, отже, також нормальні.

З наведених вище формул випливає, що їх середні дорівнюють 0, а дисперсії рівні 1.

Тому відношення величин

$$\frac{\sqrt{n}(\check{\alpha} - \alpha)}{\sqrt{\frac{n}{n-2}}\check{\sigma}} = \frac{\sqrt{n-2}(\check{\alpha} - \alpha)}{\check{\sigma}}, \quad \frac{\sqrt{n-2}S_1(\check{\beta} - \beta)}{\check{\sigma}}$$

мають розподіл Стюдента з $n - 2$ степенями свободи.

Отримані результати можна сформулювати наступним чином.

Теорема. Оцінки $\check{\alpha}$, $\check{\beta}$, $\check{\sigma}^2$, визначені формулами (3) – (4) параметрів α, β, σ^2 в задачі про просту нормальну регресію, є незміщеними, незалежними між собою і γ – довірчі інтервали для них мають такий вигляд:

$$\check{\alpha} - t_\gamma \frac{\check{\sigma}}{\sqrt{n-2}} < \alpha < \check{\alpha} + t_\gamma \frac{\check{\sigma}}{\sqrt{n-2}},$$

$$\check{\beta} - t_\gamma \frac{\check{\sigma}}{\sqrt{n-2} S_1} < \beta < \check{\beta} + t_\gamma \frac{\check{\sigma}}{\sqrt{n-2} S_1},$$

де t_γ - симетричний γ – довірчий рівень для розподілу Стюдента з $n - 2$ степенями свободи; $P\{|\tau| < t_\gamma\} = \gamma$; τ – випадкова величина, що має розподіл Стюдента з $n - 2$ степенями свободи.

Для дисперсії σ^2 маємо довірчий інтервал

$$\frac{n\check{\sigma}^2}{h'_\gamma} < \sigma^2 < \frac{n\check{\sigma}^2}{h''_\gamma}$$

де h'_γ і h''_γ - границі γ -довірчого інтервалу для χ^2 -розподілу з $n - 2$ степенями свободи:

$$P(\chi^2_{n-2} < h'_\gamma) = \frac{1-\gamma}{2}, \quad P(\chi^2_{n-2} > h''_\gamma) = \frac{1-\gamma}{2}.$$

У деяких випадках інтерес представляє оцінка ординати

$Y = \alpha + \beta(X - \bar{x})$ простої регресії при даному значенні X . Ця

оцінка задається виразом $\check{Y} = \check{\alpha} + \check{\beta}(X - \bar{x})$.

Щоб знайти довірчі інтервали для величини Y , зауважимо, що з незалежності величин $\check{\alpha}$, $\check{\beta}$, $\check{\sigma}^2$ впливає незалежність величин $\check{\alpha} - \alpha + (\check{\beta} - \beta)(X - \bar{x}) = Y - \check{Y}$ і $\check{\sigma}^2$.

Тому

$$E\check{Y} = Y, \quad D\check{Y} = D\check{\alpha} + (X - \bar{x})D\check{\beta} = \frac{\sigma^2}{n} \left(1 + \frac{(X - \bar{x})^2}{S_1^2} \right).$$

Отже, відношення

$$\tau = \frac{\sqrt{n-2} (\check{Y} - Y)}{\check{\sigma} \sqrt{1 + \frac{(X - \bar{x})^2}{S_1^2}}}$$

має розподіл Стюдента з $n - 2$ степенями свободи.

Довірчі границі для величини Y мають вигляд

$$\check{\alpha} + \check{\beta}(X - \bar{x}) \pm t_\gamma \frac{\check{\sigma}}{\sqrt{n-2}} \sqrt{1 + \frac{(X - \bar{x})^2}{S_1^2}},$$

де $\pm t_\gamma$ – симетричні – довірчі границі для розподілу Стюдента з $n - 2$ степенями свободи.

Характеристичні функції

Характеристична функція випадкової величини.

Нехай ξ — випадкова величина з функцією розподілу $F(x)$.

Характеристичною функцією випадкової величини ξ

називають функцію

$$\varphi(z) = E e^{iz\xi} = \int_{-\infty}^{\infty} e^{izx} dF(x),$$

яка означена для усіх дійсних z .

Якщо випадкова величина ξ має щільність розподілу $p(x)$, то

$$\varphi(z) = \int_{-\infty}^{\infty} e^{izx} p(x) dx.$$

Кожна характеристична функція $\varphi(z)$ має такі властивості:

а) $\varphi(0) = 1$;

б) $\varphi(z)$ рівномірно неперервна на $(-\infty, \infty)$;

в) $\varphi(z)$ додатно означена, тобто для будь-яких комплексних чисел $c_1, \dots, c_r$ і будь-яких дійсних чисел $z_1, \dots, z_r$, $r \geq 1$ справджується співвідношення

$$\sum_{k,j=1}^r c_k c_j \varphi(z_k - z_j) \geq 0.$$

Характеристична функція суми незалежних випадкових величин дорівнює добутку характеристичних функцій.

Теорема Бохнера-Хінчина.

Для того щоб функція $\varphi(z)$ була характеристичною функцією деякого розподілу, необхідно і достатньо, щоб вона була неперервна, додатно означена і $\varphi(0) = 1$.

Теореми неперервності для характеристичних функцій (теореми Леві).

Теорема 1. Якщо послідовність функцій розподілу

$$F_1(x), \dots, F_n(x), \dots$$

слабо збігається, при $n \rightarrow \infty$, до функції розподілу $F(x)$, то послідовність характеристичних функцій

$$\varphi_n(z) = \int_{-\infty}^{\infty} e^{izx} dF_n(x)$$

збігається, при $n \rightarrow \infty$, до характеристичної функції

$$\varphi(z) = \int_{-\infty}^{\infty} e^{izx} dF(x),$$

Ця збіжність рівномірна відносно z в будь-якому скінченному проміжку.

Теорема 2. Нехай послідовність характеристичних функцій

$$\varphi_1(x), \dots, \varphi_n(x), \dots$$

збігається до неперервної функції $\varphi(z)$.

Тоді функція $\varphi(z)$ є характеристичною, тобто $\varphi(z) = \int_{-\infty}^{\infty} e^{izx} dF(x)$, де $F(x)$ — деяка функція розподілу, і відповідна послідовність функцій розподілу $F_n(x)$ слабо збігається до функції $F(x)$.

Перетворення Лапласа випадкової величини.

Якщо величина ξ невід'ємна ($F(0) = 0$), то часто зручніше використати замість характеристичної функції випадкової величини її перетворення Лапласа

$$\psi(s) = E e^{-s\xi} = \int_0^{\infty} e^{-s\xi} dF(x) .$$

Характеристична функція випадкового вектора.

Характеристичною функцією випадкового вектора $\vec{\xi} = (\xi_1, \dots, \xi_m)$ називають функцію

$$\varphi(\vec{z}) = E e^{i(\vec{z}, \vec{\xi})} = \int_{-\infty}^{\infty} e^{i(\vec{z}, \vec{x})} dF(x_1, \dots, x_m) ,$$

де $(\vec{z}, \vec{\xi}) = \sum_{k=1}^m z_k \xi_k$, $F(x_1, \dots, x_m)$ — функція розподілу вектора $\vec{\xi}$.

Задача 13.1

Довести, що в тому випадку, коли характеристична функція $\varphi(z)$ випадкової величини ξ абсолютно інтегровна

$$\int_{-\infty}^{\infty} |\varphi(z)| dz < \infty$$

то випадкова величина ξ має щільність розподілу $p(x)$, причому

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-izx} \varphi(z) dz .$$

Множинна регресія.

Узагальнимо отримані результати на той випадок, коли середнє значення нормальної випадкової величини залежить від d аргументів. Отже, маємо n незалежних нормальних випадкових величин ξ_k , $k = 1, \dots, n$ з однаковою

дисперсією σ^2 і з середніми $a_k = E\xi_k$, які мають вигляд

$$E\xi_k = \alpha_0 + \beta_1(x_{1k} - \bar{x}_1) + \dots + \beta_d(x_{dk} - \bar{x}_d), \quad k = 1, \dots, n \quad (7)$$

$$\bar{x}_j = \frac{1}{n} \sum_{k=1}^n x_{jk}.$$

Для логарифма функції правдоподібності маємо такий вираз:

$$f(\zeta, \alpha, \beta, \sigma^2) = -\frac{n}{2} \ln 2\pi\sigma^2 - \frac{Q(\alpha, \beta)}{2\sigma^2} \quad (8)$$

$$\zeta = (\xi_1, \xi_2, \dots, \xi_n), \quad Q(\alpha, \beta) = \sum_{k=1}^n [\xi_k - \alpha - \sum_{j=1}^d \beta_j (x_{jk} - \bar{x}_j)]^2,$$

Введемо векторні позначення. Нехай

$$y_j = (x_{j1} - \bar{x}_j, x_{j2} - \bar{x}_j, \dots, x_{jn} - \bar{x}_j), \quad j = 1, \dots, d,$$

$$\eta = \zeta - \alpha \mathbb{I} - \sum_{j=1}^d \beta_j y_j$$

Тут $\mathbb{I}$ - вектор в R^n , компоненти якого рівні 1. Вектор

$$\eta = \eta_1, \eta_2, \dots, \eta_n, \quad \eta_k = \xi_k - \alpha - \sum_{j=1}^d \beta_j (x_{jk} - \bar{x}_j)$$

має незалежні нормально розподілені компоненти з параметрами $(0, \sigma^2)$.

Вираз для $Q(\alpha, \beta)$ набуває вигляду

$$Q(\alpha, \beta) = \left| \zeta - \alpha \mathbb{I} - \sum_{j=1}^d \beta_j y_j \right|^2, \quad (9)$$

де $|\cdot|$ - норма в просторі R^n .

Завдання максимізації функції правдоподібності зводиться до мінімізації

квадратичної форми $Q(\alpha, \beta)$ і подальшої

максимізації функції $\ln f(\zeta, \alpha, \beta, \sigma^2)$ по σ^2 . Позначимо

$$Q^* = \min_{\alpha, \beta} Q(\alpha, \beta),$$

причому мінімум $Q(\alpha, \beta)$ досягається при

$$\alpha = \check{\alpha}, \quad \beta = \check{\beta}, \quad Q^* = Q(\check{\alpha}, \check{\beta})$$

Тоді

$$f^* = \max_{\alpha, \beta, \sigma^2} f(\zeta, \alpha, \beta, \sigma^2) = \max_{\sigma^2} \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp \left\{ -\frac{Q^*}{2\sigma^2} \right\}.$$

Функція в правій частині рівності досягає свого максимуму при $\sigma^2 = \check{\sigma}^2$, де

$$\sigma^2 = \frac{1}{n} Q^* = \frac{1}{n} \left| \zeta - \check{\alpha} - \sum_{j=1}^d \check{\beta}_j y_j \right|^2. \quad (10)$$

Залишається знайти значення $\check{\alpha}$ і $\check{\beta}$.

Задача мінімізації квадратичної форми (9) має простий геометричний сенс.

Нехай H -лінійне простір в R^n , натягнутий на $d + 1$ вектор $1, y_1, \dots, y_d$. Мінімум Q^* форми Q дорівнює квадрату довжини перпендикуляра, опущеного з кінця вектора ζ на підпростір H . Якщо цей мінімум досягається при $\alpha = \check{\alpha}$, $\beta_j = \check{\beta}_j$, $j = 1, \dots, d$, то

$$\check{\zeta} = \check{\alpha} + \sum_{j=1}^d \check{\beta}_j y_j$$

є проекцією випадкового вектора ζ на H , а $\zeta - \check{\zeta}$ – вектором, ортогональним H і $|\zeta - \check{\zeta}|^2 = Q^*$.

Для визначення проекції $\check{\zeta}$ маємо $d + 1$ лінійне рівняння (так звані нормальні рівняння методу найменших квадратів)

$$(\zeta - \check{\zeta}, \mathbb{I}) = 0, \quad (\zeta - \check{\zeta}, y_j) = 0, \quad j = 1, \dots, d. \quad (11)$$

Тут (x, y) – скалярний добуток векторів $x = x_1, \dots, x_n$, $y = y_1, \dots, y_n$: $(x, y) = \sum_{k=1}^n x_k y_k$. Зазначимо, що вектор $\mathbb{I}$ ортогональний векторам y_j , $j = 1, \dots, d$, $(y_j, \mathbb{I}) = \sum_{k=1}^n (x_{jk} - \bar{x}) = 0$, причому $(\mathbb{I}, \mathbb{I}) = n$.

З рівняння $(\zeta - \check{\zeta}, \mathbb{I}) = 0$ випливає, що

$$\check{\alpha} = \frac{1}{n} (\zeta, \mathbb{I}) = \check{\zeta}.$$

Інші рівняння в (11) можна записати у вигляді

$$\sum_{i=1}^d (y_i, y_j) \beta_i = (\zeta, y_j), \quad j = 1, \dots, d. \quad (12)$$

Припустимо, що детермінант цієї системи

$$\Gamma = \begin{bmatrix} (y_1, y_1) & (y_1, y_2) & \dots & (y_1, y_d) \\ \vdots & \ddots & & \vdots \\ (y_d, y_1) & \dots & \dots & (y_d, y_d) \end{bmatrix}$$

відрізняється від 0, $\Gamma \neq 0$. Позначимо через γ_i детермінант із Γ , якщо в ньому замінити i -й стовпець правими частинами системи (12).

Тоді $\check{\beta}_j = \frac{\gamma_j}{\Gamma}$, $j = 1, \dots, d$. Зауважимо, що

$$(\zeta, y_j) = \sum_{i=1}^d (y_i, y_j) \beta_i + (\eta, y_j).$$

Відповідно до цих рівнянь, зробимо заміну величин (ζ, y_j) для γ_i . Тоді маємо $\gamma_i = \beta_i \Gamma + \bar{\gamma}_i$, де $\bar{\gamma}_i$ – детермінант із Γ , якщо в ньому замінити i -й стовпець Γ на стовпець із величин (η, y_j) , $j = 1, \dots, d$. Далі

$$\bar{\gamma}_i = \sum_{j=1}^d \Gamma_{ij} (\eta, y_j) = \sum_{k=1}^n \left(\sum_{j=1}^d \Gamma_{ij} y_{jk} \right) \eta_k,$$

де Γ_{ij} – алгебраїчні доповнення елементів детермінанту.

Таким чином, отримуємо для величин $\check{\beta}_i$ таке представлення:

$$\check{\beta}_i = \beta_i + \frac{1}{\Gamma} \sum_{k=1}^n \left(\sum_{j=1}^d \Gamma_{ij} y_{jk} \right) \eta_k, \quad i = 1, \dots, d. \quad (13)$$

Крім того,

$$\check{\alpha} = \frac{1}{n} \sum_{k=1}^n \xi_k = \alpha + \frac{1}{n} \sum_{j=1}^d \beta_j \sum_{k=1}^n y_{jk} + \frac{1}{n} \sum_{k=1}^n \eta_k,$$

або

$$\check{\alpha} = \alpha + \frac{1}{n} \sum_{k=1}^n \eta_k. \quad (14)$$

Тепер неважко вказати моменти першого і другого порядку величин $\check{\alpha}$, $\check{\beta}$.
Маємо

$$E\check{\alpha} = \alpha, \quad E\check{\beta}_j = \beta_j, \quad D\check{\alpha} = \frac{\sigma^2}{n},$$

$$\begin{aligned} E(\check{\alpha} - \alpha)(\check{\beta}_j - \beta_j) &= E \frac{1}{n} \sum_{k=1}^n \eta_k \frac{1}{\Gamma} \sum_{k=1}^n \left(\sum_{j=1}^d \Gamma_{ij} y_{jk} \right) \eta_k = \\ &= \frac{\sigma^2}{n\Gamma} \sum_{k=1}^n \sum_{j=1}^d \Gamma_{ij} y_{jk} = 0, \end{aligned}$$

$$E(\check{\beta}_i - \beta_i)(\check{\beta}_r - \beta_r) = \frac{\sigma^2}{\Gamma^2} \sum_{k=1}^n \sum_{j=1}^d \Gamma_{ij} y_{jk} \left(\sum_{j=1}^d \Gamma_{rj} y_{jk} \right) =$$

$$\begin{aligned}
&= \frac{\sigma^2}{\Gamma^2} \sum_{jj'} \Gamma_{ij} \Gamma_{rj'} \sum_{k=1} y_{jk} y_{j'k} = \frac{\sigma^2}{\Gamma^2} \sum_j \Gamma_{ij} \sum_{j'} \Gamma_{rj'} (y_j, y_{j'}) = \\
&= \frac{\sigma^2}{\Gamma^2} \sum_j \Gamma_{ij} \Gamma \delta_{rj} = \frac{\sigma^2 \Gamma_{ir}}{\Gamma}.
\end{aligned}$$

Звернемося до формули $|\check{\zeta}|^2 = |\zeta|^2 + |\zeta_N|^2$. Вона має місце для довільних значень вектора $\zeta = (z_1, z_2, \dots, z_n)$.

При цьому $|\check{\zeta}|^2$ є квадратичною формою від $z_1, z_2, \dots, z_n$ рангу $d + 1$ (внаслідок того, що $\Gamma \neq 0$ і вектор $\mathbb{I}$ ортогональний векторам $\vec{y}_1, \dots, \vec{y}_d$, а $|\zeta_N|^2$ є квадратичною формою рангу $n - d - 1$, оскільки ζ_N - довільний вектор з ортогонального доповнення до H).

Тому величина $\frac{1}{\sigma^2} |\zeta_N|^2$ має χ^2 -розподіл з $n - d - 1$ степенями свободи, а будь-яка лінійна комбінація вектора $\check{\zeta}$ не залежить від ζ_N .

Перевірка статистичних гіпотез.

Одним із завдань математичної статистики є встановлення узгодженості даної послідовності спостережень випадкових величин або подій з гіпотезами про розподіл випадкової величини (або ймовірності події). перевіряються гіпотези виникають на підставі теоретичних міркувань або в результаті статистичного дослідження інших спостережень.

Опишемо основний процес, який використовується при перевірці гіпотези. Вибирається деяка величина, що характеризує відхилення частот спостережуваних подій від гіпотетичних ймовірностей. Оскільки ця величина є функцією від спостережень, вона є статистикою. Статистики, обрані для перевірки гіпотез (зазвичай вони невід'ємні), називаються критеріями. Нехай обраний деякий критерій K . Знайдемо його розподіл в припущенні, що розподіл спостережень збігається з гіпотетичним. Фіксуємо заздалегідь певний «рівень значущості» p , вважаючи, що в одиничному експерименті події з імовірністю, меншою p , практично не відбуваються. За p знайдемо таке число x_p , щоб

$$P\{K > x_p\} = p$$

(Для простоти вважаємо, що розподіл K неперервний). Нехай $\hat{K}$ – обчислене за спостереженнями значення критерію. Якщо $\hat{K} > x_p$, то в припущенні справедливості гіпотези відбулося «практично неможлива» подія, тому гіпотеза відкидається.

В протилежному випадку вважається, що спостереження не суперечать гіпотезі.

Проілюструємо загальну ідею на прикладі критерію А. М. Колмогорова.

Нехай перевіряється гіпотеза про те, що дані спостереження $x_1, \dots, x_n$ мають неперервну функцію розподілу $F(x)$.

В якості критерію візьмемо величину

$$\mathcal{D}_n = \sqrt{n} \sup_x |F(x) - \check{F}_n(x)|,$$

де $\check{F}_n(x)$ – емпірична функція розподілу, побудована за величинами $x_1, \dots, x_n$.

Особливістю критерію $\mathcal{D}_n$ є те, що його розподіл не залежить від функції розподілу $F(x)$.

Дійсно, вважаючи в (1) $x = F^{-1}(t)$, де $t \in [0,1]$, $F^{-1}(t)$ – обернена до F функція (вона існує, оскільки F неперервна і монотонна), отримаємо

$$\mathcal{D}_n = \sqrt{n} \sup_{0 \leq t \leq 1} |t - \check{\Phi}_n(t)|,$$

де

$$\check{\Phi}_n(t) = \frac{1}{n} \sum_{i=1}^n \chi\{x_i < F^{-1}(t)\} = \frac{1}{n} \sum_{i=1}^n \chi\{F(x_i) < t\}$$

є емпірична функція розподілу, побудована за величинами $F(x_i)$, які (незалежні від виду F , якщо тільки F неперервна) мають рівномірний розподіл на $[0, 1]$. Незалежність розподілу $\mathcal{D}_n$ від F уможливорює використання критерію, оскільки якби в кожному випадку було свій розподіл, то відповідні статистичні таблиці були б неоглядні.

Крім того, є інша аналогічна складність: розподіл $\mathcal{D}_n$ залежить від n . Отже, якщо зміниться число спостережень, то потрібно заново обчислювати розподіл $\mathcal{D}_n$. Виявляється, що при досить великих n розподіл $\mathcal{D}_n$ вже практично не залежить від n , так як існує граничний розподіл $\mathcal{D}_n$ при $n \rightarrow +\infty$. Цей граничний розподіл саме використовується в якості розподілу $\mathcal{D}_n$ вже при $n \geq 20$. Вибравши, наприклад, число λ так, щоб

$$\lim_{n \rightarrow \infty} P\{\mathcal{D}_n > \lambda\} = \frac{5}{100},$$

отримаємо наступне правило перевірки гіпотези з 5% -м рівнем значущості: якщо $\mathcal{D}_n^*$ величина, обчислена за спостереженнями, то при $\mathcal{D}_n^* > \lambda$ гіпотеза відкидається, а при $\mathcal{D}_n^* < \lambda$ – приймається (до наступних спостережень).

Граничний розподіл величини $\mathcal{D}_n$ знаходиться з використанням деяких методів теорії випадкових процесів, які тут не викладалися, тому детально він не розглядається.

Найбільш широке застосування в статистиці отримав критерій χ^2 .

Цей критерій використовується при перевірці гіпотези про можливість декількох несумісних подій (або, що те ж саме, про розподіл випадкової величини, що приймає скінчене число значень). Вже відзначалися переваги критерію Колмогорова: це - незалежність розподілу від гіпотези, що перевіряється, а також незалежність від числа спостережень.

Нехай в експерименті спостерігається одне з r подій: $A_1, \dots, A_r$, $p_1, \dots, p_r$ – гіпотетичні ймовірності цих подій.

Проведено n спостережень, при цьому подія A_k спостерігалася v_k раз, $k = 1, \dots, r$. Складемо вираз

$$\chi_n^2 = \sum_{i=1}^r \frac{(v_i - np_i)^2}{np_i}$$

Його граничний розподіл χ_n^2 збігається з χ^2 – розподілом з $r - 1$ степенями свободи, тобто з розподілом $\xi_1^2 + \dots + \xi_{r-1}^2$, де $\xi_1, \dots, \xi_{r-1}$ – незалежні однаково розподілені величини, які мають гауссовське розподіл із середнім 0 і дисперсією 1. Тому при $x > 0$ маємо

$$\lim_{n \rightarrow \infty} P\{\chi_n^2 < x\} = \frac{1}{2^{\frac{r}{2}} \Gamma(\frac{r}{2})} \cdot \int_0^x t^{\frac{r}{2}-1} e^{-\frac{1}{2}t} dt.$$

Таким чином, залежність від гіпотетичного розподілу входить через число можливих подій r . При використанні критерію r для перевірки гіпотези про розподіл неперервної величини можливі значення величини розбивають на скінчене число неперетинаючих областей $\mathcal{J}_1, \dots, \mathcal{J}_r$ і в якості подій A_k розглядають події $\{\xi \in \mathcal{J}_k\}$, де ξ - спостережувана величина.

При великих значеннях r можна наближено замінити розподіл χ^2 нормальним із середнім $r - 1$ і дисперсією $2(r - 1)$.

Нехай λ_p – таке значення, що

$$\frac{1}{2^{\frac{r}{2}} \Gamma\left(\frac{r}{2}\right)} \cdot \int_0^x t^{\frac{r}{2}-1} e^{-\frac{t}{2}} dt = p.$$

Тоді при $\chi_n^2 \leq \lambda_p$ гіпотеза вважається неспростованою, а при $\chi_n^2 > \lambda_p$ вона відкидається. Подивимося, що відбувається в тому випадку, коли істинні значення ймовірностей подій A_n не є рівними p_n .

Припустимо, що $P(A_i) = \bar{p}_i$, причому $\sum (p_i - \bar{p}_i)^2 > 0$. Тоді

$$\chi_n^2 = \sum_{i=1}^r \frac{(v_i - n\bar{p}_i)^2}{np_i} + 2 \sum_{i=1}^r \frac{(v_i - n\bar{p}_i)}{p_i} \frac{(p_i - \bar{p}_i)}{p_i} + n \sum_{i=1}^r \frac{(p_i - \bar{p}_i)^2}{p_i}$$

Оскільки

$$E \sum_{i=1}^r \frac{(v_i - n\bar{p}_i)^2}{np_i} = r - 1$$

то

$$\frac{1}{n} \sum_{i=1}^r \frac{(v_i - n\bar{p}_i)^2}{np_i} \rightarrow 0 \quad \text{за ймовірністю}$$

Далі

$$\begin{aligned} & E \left(\frac{1}{n} \sum_{i=1}^r (v_i - n\bar{p}_i) \left(\frac{p_i - \bar{p}_i}{p_i} \right) \right)^2 \leq \\ & \leq r \frac{1}{n^2} \sum_{i=1}^r E (v_i - n\bar{p}_i)^2 \left(\frac{p_i - \bar{p}_i}{p_i} \right)^2 = \\ & = r \frac{1}{n^2} \sum_{i=1}^r n\bar{p}_i (1 - \bar{p}_i) (p_i - \bar{p}_i)^2 \rightarrow 0. \end{aligned}$$

Отже

$$\frac{1}{n} \sum_{i=1}^r \frac{(v_i - n\bar{p}_i)(p_i - \bar{p}_i)}{p_i} \rightarrow 0 \quad \text{за ймовірністю}$$

Тоді

$$\frac{1}{n} \chi_n^2 \rightarrow \sum_{i=1}^r \frac{(p_i - \bar{p}_i)^2}{p_i}$$

Тому для великих n

$$\chi_n^2 \sim n \sum_{i=1}^n \frac{(p_i - \bar{p}_i)^2}{p_i}$$

і з досить великою ймовірністю $\chi_n^2 > x_p$. Таким чином, якщо гіпотеза не вірна, то ми, використовуючи критерій χ^2 , відкидаємо її при досить великому n . Імовірність відкинути гіпотезу, якщо вона не є істинною, називається потужністю критерію. Ця ймовірність залежить від того розподілу ймовірностей $(\bar{p}_1, \dots, \bar{p}_r)$, яке є істинним. Будь-який розподіл ймовірностей, що може виявитися істинним, але відрізняється від гіпотетичного, називається альтернативним, або альтернативою. Отже, потужність критерію залежить від альтернативи. Ми показали, що для критерію χ^2 при будь-якому виборі рівня значущості p потужність прямує до 1 при $n \rightarrow \infty$.

Рекомендована література

1. Клесов О.І. Теорія ймовірностей та математична статистика. ТВІМС 2018, 427 с.
2. Булдигін В.В., Буценко Ю.П., Диховичний О.О. Теорія ймовірностей, конспект лекцій, ТВІМСб Київ, 1999.