

### 3.4 Характеристичні функції

#### 3.4.1 Комплексні випадкові величини

Перш ніж перейти до вивчення апарату характеристичних функцій, який далі буде застосовано для доведення центральної граничної теореми, ознайомимось з поняттям комплексної випадкової величини.

**Означення.** Комплексною випадковою величиною  $Z(w)$  будемо називати комплексно значну функцію елементарної події

$$Z(w) = X(w) + iY(w),$$

де  $X(w), Y(w)$  – дійсні випадкові величини, що задані на імовірнісному просторі  $(\Omega, F, P)$ , а  $i$  – уявна одиниця,  $i^2 = -1$ .

При фіксованому  $w \in \Omega$   $Z$  є звичайною комплексною величиною.

Операції над комплексними випадковими величинами такі самі, що і над звичайними комплексними числами, з тією різницею, що тепер дійсна частина та коефіцієнт при уявній частині результату є функція від випадкових аргументів, тобто самі є випадковими величинами.

Так сума двох комплексних випадкових величин  $Z_1(w) = X_1(w) + iY_1(w)$  та  $Z_2(w) = X_2(w) + iY_2(w)$  визначається як

$$Z_1(w) + Z_2(w) = X_1(w) + X_2(w) + i(Y_1(w) + Y_2(w)).$$

Добуток цих величин є

$$Z_1(w) \cdot Z_2(w) = X_1(w) \cdot X_2(w) - Y_1(w) \cdot Y_2(w) + i(X_1(w)Y_2(w) + X_2(w)Y_1(w)).$$

Модуль комплексної випадкової величини визначається як

$$|Z(w)| = \sqrt{X^2(w) + Y^2(w)}.$$

Надалі символ залежності випадкових величин від елементарної події будемо знову пропускати.

**Означення.** Дві комплексні випадкові величини  $Z_1$  та  $Z_2$  називаються незалежними, якщо незалежними є двовимірні випадкові величини  $(X_1, Y_1)$  та  $(X_2, Y_2)$ , тобто

$$P(X_1 < x_1, Y_1 < y_1, X_2 < x_2, Y_2 < y_2) = P(X_1 < x_1, Y_1 < y_1)P(X_2 < x_2, Y_2 < y_2).$$

Зауважимо, що звідси виходить незалежність однієї з компонент  $(X_1, Y_1)$  від довільної компоненти  $(X_2, Y_2)$ . Покладіть, наприклад, в останній рівності  $y_1 = \infty, y_2 = \infty$ , і ви одержите

$$P(X_1 < x_1, X_2 < x_2) = P(X_1 < x_1) \cdot P(X_2 < x_2),$$

що означає незалежність  $X_1$  та  $X_2$ .

В загальному випадку  $n$  комплексних випадкових величин  $Z_1, \dots, Z_n$  їх незалежність визначається як незалежність (у сукупності) векторів  $(X_1, Y_1), \dots, (X_n, Y_n)$ , тобто коли

$$P(X_1 < x_1, Y_1 < y_1, \dots, X_n < x_n, Y_n < y_n) = \prod_{i=1}^n P(X_i < x_i, Y_i < y_i).$$

Знов можна переконатися (зробіть це самостійно), що коли з кожного вектора  $(X_i, Y_i), i = \overline{1, n}$  взяти по одній (довільній) компоненті, то одержані  $n$  випадкових величин будуть незалежними (у сукупності).

**Означення. Математичним сподіванням** комплексної випадкової величини  $Z = X + iY$  називається число

$$MZ = MX + iMY \quad (1)$$

(вважаємо надалі, що  $MX$  та  $MY$  існують та скінченні).

Неважко переконатися, що властивості адитивності та мультиплікативності математичного сподівання мають місце і у комплексному випадку.

Дійсно, для суми комплексних випадкових величин  $Z = Z_1 + Z_2$  маємо

$$M(Z_1 + Z_2) = M[(X_1 + X_2) + i(Y_1 + Y_2)] = M(X_1 + X_2) + iM(Y_1 + Y_2) = \\ iMX_1 + iMY_1 + MX_2 + iMY_2 = MZ_1 + MZ_2,$$

тобто математичне сподівання суми комплексних випадкових величин дорівнює сумі їх математичних сподівань.

Нехай  $Z_1$  і  $Z_2$  – незалежні комплексні випадкові величини, тоді (використовуємо «перехресну» незалежність компонент векторів  $(X_1, Y_1)$ ,  $(X_2, Y_2)$  та властивості адитивності та мультиплікативності математичного сподівання для дійсних випадкових величин):

$$M[Z_1 \cdot Z_2] = M[(X_1 X_2 - Y_1 Y_2) + i(X_1 Y_2 + X_2 Y_1)] = M(X_1 X_2 - Y_1 Y_2) + iM(X_1 Y_2 + X_2 Y_1) = \\ iMX_1MX_2 - MY_1MY_2 + iMX_1MY_2 + iMX_2MY_1 = (MX_1 + iMY_1) \cdot (MX_2 + iMY_2) = MZ_1 \cdot MZ_2$$

і, таким чином, математичне сподівання добутку незалежних комплексних випадкових величин дорівнює добутку математичних сподівань співмножників.

Зрозуміло, що ці властивості за індукцією переносяться на довільну скінченну кількість доданків та незалежних співмножників.

Доведемо важливу нерівність для комплексної випадкової величини (для дійсної випадкової величини ми вже робили це раніше):

$$|MZ| \leq M|Z|, \quad (2)$$

Розглянемо лише два випадки:

1. Нехай  $(X, Y)$  – дискретна випадкова величина, компоненти якої приймають скінченну кількість значень:  $x_l, l = \overline{1, s}$  та  $y_j, j = \overline{1, m}$ , відповідно. Задано їх сумісний розподіл  $p(x_l, y_j) = P(X = x_l, Y = y_j)$ ,  $l = \overline{1, s}, j = \overline{1, m}$ .

Тоді

$$MZ = \sum_{l=1}^s \sum_{j=1}^m x_l p(x_l, y_j) + i \sum_{l=1}^s \sum_{j=1}^m y_j p(x_l, y_j) = \sum_{l=1}^s \sum_{j=1}^m (x_l + i y_j) p(x_l, y_j).$$

$$|MZ| = \left| \sum_{l=1}^s \sum_{j=1}^m (x_l + i y_j) p(x_l, y_j) \right| \leq \sum_{l=1}^s \sum_{j=1}^m |x_l + i y_j| p(x_l, y_j) = \sum_{l=1}^s \sum_{j=1}^m \sqrt{x_l^2 + y_j^2} p(x_l, y_j) = M |Z|$$

У дискретному випадку нерівність доведено.

2. Неперервний випадок. Нехай  $(X, Y)$  – неперервна випадкова величина зі щільністю розподілу  $f(x, y)$ . Тоді

$$MZ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy + i \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + i y) f(x, y) dx dy$$

Далі

$$|MZ| = \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + i y) f(x, y) dx dy \right| \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x + i y| f(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sqrt{x^2 + y^2} f(x, y) dx dy = M |Z|$$

### 3.4.2 Характеристичні функції. Властивості

**Означення.** Характеристичною функцією дійсної випадкової величини  $X$  називається комплексно-значна функція  $\phi_X(t)$  дійсного аргумента  $t$ , що дорівнює

$$\phi_X(t) = M e^{iXt} \quad (1)$$

Якщо розкрити в (1)  $e^{itX}$  за формулою Ейлера<sup>1</sup>  $e^{itX} = \cos tX + i \sin tX$ , то одержуємо

$$\phi_X(t) = M \cos tX + i M \sin tX \quad (2)$$

---

<sup>1</sup>Ейлер Леонард (1707-1783) – видатний математик, фізик, механік і астроном, член Петербурзької АН та Паризької АН.

Для дискретної випадкової величини  $X$ , що приймає значення  $x_j, j=\overline{1,s}$ , з імовірностями  $p(x_j), j=\overline{1,s}$ , характеристична функція має вигляд

$$\phi_X(t) = \sum_{j=1}^s e^{ix_j t} p(x_j). \quad (3)$$

У випадку неперервної випадкової величини  $X$  зі щільністю розподілу  $f_X(x)$

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{ixt} f_X(x) dx, \quad (4)$$

тобто  $\phi_X(t)$  є ніщо інше, як перетворення Фур'є<sup>2</sup> щільності  $f_X(t)$ .

Розглянемо спочатку дві найпростіші властивості характеристичних функцій (до інших перейдемо після прикладів обчислення цих функцій для деяких розподілів).

1. Характеристична функція обмежена.

$$|\phi_X(t)| \leq 1, \phi_X(0) = 1.$$

Дійсно, з нерівності  $|MZ| \leq M|Z|$  маємо

$$|Me^{itX}| \leq M|e^{itX}| = M|\cos tX + i \sin tX| = M(\cos^2 tX + \sin^2 tX) = 1.$$

Це, зокрема, означає, що  $\phi_X(t)$  існує для довільних  $t$ .

2. Нехай  $Y = aX + b$ , де  $a$  та  $b$  – сталі. Тоді

$$\phi_Y(t) = e^{itb} \phi_X(at). \quad (5)$$

З означення характеристичної функції маємо

$$\phi_Y(t) = Me^{itY} = Me^{it(aX+b)} = e^{itb} Me^{itaX} = e^{itb} \phi_X(at).$$

**Приклад 1.** Біномний розподіл. Як ми пам'ятаємо, його задано ймовірностями

---

<sup>2</sup>Фур'є Жан Батіст Жозеф (1768-1830) – видатний французький математик, член Паризької АН, почесний член Петербурзької АН.

$$p(m) = P(X=m) = C_n^m p^m q^{n-m}, \text{ де } q=1-p, m=\overline{0, n}.$$

Характеристична функція дорівнює (дивись (3)):

$$\phi_X(t) = M e^{itX} = \sum_{m=0}^n C_n^m e^{itm} p^m q^{n-m} = \sum_{m=0}^n C_n^m (e^{it} p)^m q^{n-m} = (e^{it} p + q)^n.$$

Тут ми використали біном Ньютона.

**Приклад 2.** Пуассонівський розподіл. У цьому випадку

$$p(m) = P(X=m) = \frac{a^m}{m!} e^{-a}, \quad a > 0, \quad m = \overline{0, \infty}.$$

Обчислюємо характеристичну функцію:

$$\phi_X(t) = \sum_{m=0}^{\infty} e^{itm} \frac{a^m}{m!} e^{-a} = e^{-a} \sum_{m=0}^{\infty} \frac{(ae^{it})^m}{m!} = e^{-a} \cdot e^{ae^{it}} = e^{a(e^{it}-1)}.$$

**Приклад 3.** Вироджений розподіл:

$$P(X=c) = 1;$$

$$\phi_X(t) = e^{itc}.$$

**Приклад 4.** Рівномірний розподіл на  $[a, b]$ . Щільність розподілу має вигляд:

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b]; \\ 0, & \text{інакше} \end{cases}$$

Характеристична функція дорівнює (дивись (4)):

$$\phi_X(t) = \frac{1}{b-a} \int_a^b e^{itx} dx = \frac{1}{it(b-a)} \int_a^b de^{itx} = \frac{e^{itb} - e^{ita}}{it(b-a)}.$$

**Приклад 5.** Нормальний розподіл. Спочатку розглянемо нормований нормальний розподіл, коли щільність розподілу  $X$  задано формулою:

$$f(x) = n(x, 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty.$$

Характеристична функція дорівнює (вилучаємо повний квадрат та використовуємо інтеграл Пуассона)

$$\phi_X(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} \cdot e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} e^{-\frac{(x-it)^2}{2}} dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}} d(x-it) = e^{-\frac{t^2}{2}}.$$

Тепер поставимо задачу відшукати характеристичну функцію випадкової величини  $Y$  з загальним нормальним розподілом  $n(y, \nu, \sigma)$ . Очевидно, величина  $Y = \sigma X + \nu$  має цей розподіл (переконайтеся в цьому самостійно). Застосовуємо доведену раніше властивість (дивись формулу (5)) та одержуємо

$$\phi_Y(t) = e^{it\nu} \phi_X(t\sigma) = e^{it\nu} \cdot e^{-\frac{t^2\sigma^2}{2}} = e^{it\nu - \frac{t^2\sigma^2}{2}}. \quad (6)$$

Запам'ятаємо цей простий але дуже важливий результат.

**Приклад 6.** Нехай випадкова величина  $X$  має експоненційний (показниковий) розподіл з параметром  $\lambda > 0$ , тобто щільність розподілу має вигляд:

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

Характеристична функція експоненційного розподілу дорівнює:

$$\phi_X(t) = \int_0^{\infty} e^{itx} \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} e^{-(\lambda - it)x} dx = \frac{\lambda}{\lambda - it}.$$

Повернемось до властивостей характеристичних функцій.

3. Коли  $X_1, X_2, \dots, X_n$  – незалежні, то

$$\phi_{X_1 + \dots + X_n}(t) = \prod_{k=1}^n \phi_{X_k}(t).$$

Дійсно, з незалежності  $X_1, \dots, X_n$  виходить незалежність  $e^{itX_1}, \dots, e^{itX_n}$ .

Тоді можна використати властивість мультиплікативності математичного сподівання та одержати

$$\phi_{X_1 + \dots + X_n}(t) = M e^{it \sum_{k=1}^n X_k} = M \prod_{k=1}^n e^{itX_k} = \prod_{k=1}^n M e^{itX_k} = \prod_{k=1}^n \phi_{X_k}(t).$$

1. Нехай  $\nu_k = MX^k$  - початкові моменти  $k$ -го порядку випадкової величини  $X$ , та  $\nu_n < \infty$ . Тоді для усіх  $0 \leq k \leq n$  існують похідні  $\phi_X^{(k)}(t)$  і

$$\frac{d^k \phi_X(t)}{dt^k} = i^k M(X^k e^{itX}) \quad (8)$$

$$\phi_X^{(k)}(0) = i^k \nu_k, k = \overline{0, n}. \quad (9)$$

У дискретному випадку (8),(9) доведіть самостійно.

Розглянемо випадок неперервної випадкової величини  $X$ . Якщо  $k$ -ту похідну можна поміняти місцями з інтегралом, то отримуємо необхідний нам результат:

$$\frac{d^k \phi_X(t)}{dt^k} = \frac{d^k}{dt^k} \int_{-\infty}^{\infty} e^{itx} f(x) dx = \int_{-\infty}^{\infty} \frac{d^k (e^{itx})}{dt^k} f(x) dx = i^k \int_{-\infty}^{\infty} x^k e^{itx} f(x) dx = i^k M(X^k e^{itX}).$$

За допомогою метода математичної індукції покажемо, що в цьому випадку похідну та інтеграл можна міняти місцями.

Вірність (8) при  $k = 0$  очевидна, тому що

$$\phi_X^{(0)}(t) = \phi_X(t) = M e^{itX}.$$

Припустимо, що (8) вірна для деякого  $k < n$ . Доведемо, що вона вірна і для випадку  $k+1$ . Маємо:

$$\frac{\phi_X^{(k)}(t+h) - \phi_X^{(k)}(t)}{h} = i^k M X^k \frac{e^{itX}(e^{ihX} - 1)}{h} = i^k \int_{-\infty}^{\infty} x^k \frac{e^{itx}(e^{ihx} - 1)}{h} f(x) dx \quad (10)$$

$$\left| x^k \frac{e^{itx}(e^{ihx} - 1)}{h} \right| = \frac{|x|^k}{|h|} |e^{itx}| |e^{ihx} - 1| = \frac{|x|^k}{|h|} \left| \int_0^{hx} e^{iu} du \right| \leq \frac{|x|^k}{|h|} \int_0^{|hx|} du = \frac{|x|^k |hx|}{|h|} = |x|^{k+1};$$

$$M |X|^{k+1} = \int_{-\infty}^{\infty} |x|^{k+1} f(x) dx < \infty$$

(Останнє вже було доведено раніше: якщо скінченим є початковий момент  $n$ -го порядку випадкової величини  $X$ , то скінченим є початковий



момент  $n$  – го порядку випадкової величини  $|X|$  та скінчені усі моменти більш низьких порядків).

Далі використовуємо варіант теореми Лебега про мажоровану збіжність: якщо є послідовність функцій  $\psi_n(x), n=\overline{1, \infty}$  таких, що для

кожного  $n$   $|\psi_n(x)| \leq g(x), \int_{-\infty}^{\infty} g(x) dx < \infty$  та  $\lim_{n \rightarrow \infty} \psi_n(x) = \psi(x)$  у кожній точці  $x \in (-\infty, \infty)$ , то

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \psi_n(x) dx = \int_{-\infty}^{\infty} \lim_{n \rightarrow \infty} \psi_n(x) dx = \int_{-\infty}^{\infty} \psi(x) dx.$$

Цей результат дає нам змогу у лівій та правій частині (10) перейти до границі  $h \rightarrow 0$  за правилом Лопіталя

$$\lim_{n \rightarrow \infty} \frac{(e^{ihn} - 1)}{h} = \lim_{n \rightarrow \infty} (e^{ihn} - 1)' = \lim_{n \rightarrow \infty} ixe^{ihn} = ix.$$

Одержуємо

$$\phi^{(k+1)}(t) = i^{k+1} M X^{k+1} e^{itX}.$$

І, таким чином, ми довели вірність (8) для  $k+1$ , а згідно з правилом математичної індукції вона вірна для усіх  $k \leq n$ . Якщо покласти в (8)  $t=0$ , одержуємо (9).

2. Якщо  $v_n = M X^n < \infty$ , то можна довести, що (аналог розкладу Маклорена):

$$\phi(t) = \sum_{k=0}^n \frac{(it)^k}{k!} v_k + R_n(t), \quad (11)$$

де  $R_n(t) = o(t^n)$ , тобто  $|R_n(t)|/t^n \rightarrow 0$ .

**Приклад.** Ми вже переконалися що характеристична функція нормального розподілу дорівнює

$$\phi(t) = e^{ivt - \frac{t^2 \sigma^2}{2}}.$$

За допомогою властивостей характеристичної функції ще раз з'ясуємо імовірнісний зміст параметрів  $\nu$  та  $\sigma^2$ . Диференціювання  $\phi(t)$  дає:

$$\begin{aligned}\phi'(t) &= e^{ivt - \frac{t^2 \sigma^2}{2}} (iv - t\sigma^2); \\ \phi''(t) &= e^{ivt - \frac{t^2 \sigma^2}{2}} (iv - t\sigma^2)^2 - \sigma^2 e^{ivt - \frac{t^2 \sigma^2}{2}}.\end{aligned}$$

Звідси

$$\begin{aligned}\phi'(0) &= iMX = iv, \\ \phi''(0) &= i^2 MX^2 = -(\nu^2 - \sigma^2),\end{aligned}$$

тобто  $\nu = MX$ ,  $\nu^2 + \sigma^2 = MX^2$ , а отже  $DX = MX^2 - (\nu X)^2 = \sigma^2$ .

Між функціями розподілу випадкових величин та характеристичними функціями існує взаємно однозначна відповідність. Має місце наступний результат.

**Теорема** (формула обернення). Якщо  $a < b$  - точки неперервності функції розподілу  $F(x)$ , то

$$F_X(b) - F_X(a) = \lim_{A \rightarrow \infty} \frac{1}{2\pi} \int_{-A}^A \frac{e^{-ita} - e^{-itb}}{it} \phi_X(t) dt.$$

Формула обернення вірна і для усіх  $a > b$  (тобто коли  $a$  та  $b$  можуть бути точками розриву  $F_X(x)$ ), якщо в лівій частині останньої рівності замінити  $F_X(x)$  на так звану нормалізовану функцію  $F_X^0(x)$ , яка

у точках розриву  $F_X(x)$  дорівнює  $F_X^0(x) = \frac{F_X(x-0) + F_X(x+0)}{2}$ , а у точках неперервності співпадає  $F_X(x)$  (без доведення).

Зауважимо, що по нормалізованій функції  $F_X^0(x)$  функція розподілу  $F_X(x)$  однозначно відновлюється, коли у точках неперервності покласти  $F_X(x) = F_X^0(x)$ , а у точках розриву  $F_X^0(x) = F_X(x)$ .

Розглянемо окремо випадок неперервної випадкової величини  $X$  за щільністю розподілу  $f_X(x)$ . Тоді (4) (як зазначалося раніше) є перетворення

Фур'є функції  $f_X(x)$ . З курсу аналізу відомо, що коли  $\int_{-\infty}^{\infty} |\phi_X(t)| dt < \infty$ , то має місце формула обернення для перетворення Фур'є:

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_X(t) dt,$$

що однозначно визначає функцію щільності на характеристичній функції.

Для дискретної випадкової величини задачу обернення розв'яжемо у випадку цілочислової випадкової величини, що приймає значення  $k = \overline{-\infty, \infty}$  з імовірностями  $p_k = P(X=k)$ . Характеристична функція такої величини має вигляд (дивись (3)):

$$\phi_X(t) = \sum_{k=-\infty}^{\infty} p_k e^{ikt}.$$

множимо ліву та праву частину цієї рівності на  $e^{-imt}$  та інтегруємо від  $-\pi$  до  $\pi$  (по членне інтегрування ряду в правій частині законно, тому

що він збігається рівномірно:  $|p_k e^{ikt}| = p_k, \sum_{k=-\infty}^{\infty} p_k = 1$ ). Інтеграли

$$\int_{-\pi}^{\pi} e^{i(k-m)t} dt = 0, \text{ для } k \neq m. \text{ Дійсно,}$$

$$\int_{-\pi}^{\pi} e^{i(k-m)t} dt = \frac{1}{i(k-m)} e^{i(k-m)t} \Big|_{-\pi}^{\pi} = \frac{1}{i(k-m)} [e^{i(k-m)\pi} - e^{-i(k-m)\pi}] =$$

$$i \frac{1}{i(k-m)} [\cos(k-m)\pi - \cos(m-k)\pi + i \sin(k-m)\pi - i \sin(m-k)\pi] = 0, k \neq 1$$

а при  $k = m$

$$\int_{-\pi}^{\pi} e^{i(k-m)t} dt = 2\pi.$$

Звідси

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-imt} \phi_X(t) dt = p_m, m = -\infty, \infty.$$

Бачимо, що розподіл цілочислової випадкової величини повністю визначається її характеристичною функцією на відрізку  $[-\pi; \pi]$ .

### 3.4.3 Теореми про характеристичні функції. Центральна гранична теорема

Дуже важливими для розуміння граничних теорем є наступні результати.

Раніше вже було встановлено взаємно однозначну відповідність між множиною функцій розподілу  $\{F_X(x)\}$  та множиною характеристичних функцій. Виявляється, що ця відповідність не тільки взаємно однозначна, але й, в певному розумінні, взаємно неперервна.

В 3.1 вже йшла мова про слабку збіжність  $F_n(x) \Rightarrow F(x)$ , тобто збіжність функцій розподілу  $F_n(x)$  до функції розподілу  $F(x)$  у точках неперервності останньої. Одна з найважливіших властивостей

характеристичних функцій міститься у наступних двох теоремах (теоремах неперервності), які ми наведемо без доведення.

Нехай  $F_n(x)$ ,  $i=\overline{1,\infty}$ ,  $F(x)$  — функції розподілу,  $\phi_n(t)$ ,  $i=\overline{1,\infty}$ ,  $\phi(t)$  — відповідні їм характеристичні функції.

**Теорема 1 (пряма гранична теорема).** Якщо  $F_n(x) \Rightarrow F(x)$ , то  $\phi_n(t) \Rightarrow \phi(t)$  у кожній точці  $t$ .

**Теорема 2 (обернена гранична теорема).** Якщо  $\phi_n(t)$ ,  $i=\overline{1,\infty}$ , збігається у кожній точці  $t$  до деякої функції  $\phi(t)$ , що неперервна у нулі, то  $F_n(x) \Rightarrow F(x)$  і  $\phi(t)$  є характеристичною функцією, що відповідає розподілу  $F(x)$ .

Далі ми у найпростішому варіанті розглянемо один з основних результатів теорії ймовірностей. Який показує, що за деяких умов сума великої кількості випадкових величин наближено розподілена нормально.

Теорема (центральна гранична).

Нехай  $X_1, X_2, \dots, X_n, \dots$  — послідовність незалежних однаково розподілених випадкових величин зі скінченною дисперсією  $0 < DX = \sigma^2 < \infty$ , математичне сподівання яких  $MX_i = \nu$ . Тоді

$$\lim_{n \rightarrow \infty} P\left(\frac{\sum_{k=1}^n X_k - n\nu}{\sigma\sqrt{n}} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du \quad (1)$$

Інакше кажучи, послідовність  $Z_n = \frac{\sum_{k=1}^n X_k - n\nu}{\sigma\sqrt{n}}$  (послідовність центрованих та нормованих сум випадкових величин  $X_1, X_2, \dots, X_n$ )

збігається за розподілом до випадкової величини  $Z$ , що має нормований нормальний розподіл..

Доведення. Використовуємо метод характеристичних функцій.

Позначимо  $\bar{X}_k = \frac{X_k - \nu}{\sigma}$ ,  $\bar{S}_n = \sum_{k=1}^n \bar{X}_k$ . Очевидно, що  $M \bar{X}_k = 0$ ,

$$D \bar{X}_k = 1, \quad Z_n = \frac{1}{\sqrt{n}} \bar{S}_n.$$

Розклад характеристичної функції для випадкової величини  $\bar{X}_k$  має вигляд (дивись (11) §3.4.2)

$$\phi_{\bar{X}_k}(t) = 1 + it M \bar{X}_k + \frac{(it)^2}{2} M (\bar{X}_k)^2 + o(t^2) = 1 - \frac{t^2}{2} + o(t^2).$$

За властивістю характеристичної функції для суми незалежних випадкових величин маємо:

$$\phi_{\bar{S}_n}(t) = \prod_{k=1}^n \phi_{\bar{X}_k}(t) = \left(1 - \frac{t^2}{2} + o(t^2)\right)^n$$

Далі застосовуємо властивість характеристичної функції для лінійного перетворення випадкової величини (дивись (5) §3.4.2) та одержуємо

$$\phi_{Z_n}(t) = \phi_{\bar{S}_n}\left(\frac{t}{\sqrt{n}}\right) = \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n.$$

Переходимо до границі при  $n \rightarrow \infty$ . Маємо:

$$\phi_{Z_n}(t) \xrightarrow{n \rightarrow \infty} e^{-\frac{t^2}{2}}.$$

Бачимо, що послідовність характеристичних функцій  $\phi_{Z_n}(t)$  збігається для кожного  $t$  до функції  $e^{-\frac{t^2}{2}}$ , яка неперервна у нулі та є характеристичною функцією деякої випадкової величини  $Z$  з нормованим нормальним розподілом. Застосовуємо теорему 2 та одержуємо твердження центральної граничної теореми (зауважимо, що

збіжність в (1) має місце для кожного  $x < \infty$ , тому права частина (1) неперервна в кожній точці  $x$ ).

Як і у випадку інших граничних теорем, центральну граничну теорему на практиці застосовують у дограничній ситуації. А саме, для великих  $n$  (практично  $n \geq 40$ ) замість (1) використовують наближену рівність

$$P\left(\frac{\sum_{k=1}^n X_k - nv}{\sigma\sqrt{n}} < x\right) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du = \frac{1}{2} + \Phi_0(x) \quad (2)$$

де  $\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}} du$  — вже відома нам функція Лапласа.

**Приклад.** Нехай деяка величина вимірюється  $n=3200$  разів. Похибка у кожному  $k$ -ому вимірюванні  $X_k$  є випадковою величиною з математичним сподіванням  $\nu=3$  (систематична похибка) та дисперсією  $\sigma^2=2$  (випадкова похибка). Результати вимірювань сумуються та береться їх середнє арифметичне. Треба обчислити ймовірність того, що похибка вимірювань буде знаходитися в інтервалі  $(a;b)$ , де  $a=2,925$ ;  $b=3,075$ .

Очевидно, мова йде про обчислення  $p = P\left(a \leq \frac{1}{n} \sum_{k=1}^n X_k < b\right)$ . Робимо еквівалентні перетворення та застосовуємо центральну граничну теорему:

$$p = P\left(na \leq \sum_{k=1}^n X_k < nb\right) = P\left(\frac{na - nv}{\sigma\sqrt{n}} \leq \frac{\sum_{k=1}^n X_k - nv}{\sigma\sqrt{n}} < \frac{nb - nv}{\sigma\sqrt{n}}\right) =$$

$$\begin{aligned}
&= P\left(\frac{\sum_{k=1}^n X_k - nv}{\sigma\sqrt{n}} < \frac{nb - nv}{\sigma\sqrt{n}}\right) - P\left(\frac{\sum_{k=1}^n X_k - nv}{\sigma\sqrt{n}} < \frac{na - nv}{\sigma\sqrt{n}}\right) \approx \Phi_0\left(\frac{nb - nv}{\sigma\sqrt{n}}\right) - \Phi_0\left(\frac{na - nv}{\sigma\sqrt{n}}\right) = \\
&= \Phi_0(3) - \Phi_0(-3) = 2\Phi_0(3) = 2 \cdot 0,4986 = 0,9972
\end{aligned}$$

У цьому прикладі водночас встановлено, що

$$P\left(a \leq \frac{1}{n} \sum_{k=1}^n X_k < b\right) \approx \Phi_0\left(\frac{b-v}{\sigma/\sqrt{n}}\right) - \Phi_0\left(\frac{a-v}{\sigma/\sqrt{n}}\right),$$

тобто при великих  $n$  середнє арифметичне  $\frac{1}{n} \sum_{k=1}^n X_k$  наближено має нормальний розподіл з математичним сподіванням  $v$  та середнім квадратичним відхиленням  $\frac{\sigma}{\sqrt{n}}$ .

Аналогічним прийомом можна показати, що

$$P\left(a \leq \sum_{k=1}^n X_k < b\right) \approx \Phi_0\left(\frac{b-nv}{\sigma\sqrt{n}}\right) - \Phi_0\left(\frac{a-nv}{\sigma\sqrt{n}}\right),$$

тобто при великих  $n$  сума  $\sum_{k=1}^n X_k$  наближено має нормальний розподіл з математичним сподіванням  $nv$  та середнім квадратичним відхиленням  $\sigma\sqrt{n}$ , що підтверджує висловлену на початку параграфа тезу.

Наслідком центральної граничної теореми є наступний результат.

Знову розглянемо схему  $n$  незалежних випробувань Бернуллі, у кожному з яких подія  $A$  відбувається з імовірністю  $0 < p < 1$  або не відбувається з імовірністю  $q = 1 - p$ . Введемо випадкові величини

$$X_k = \begin{cases} 1, & \text{якщо подія } A \text{ відбувається у } k\text{-ому випробуванні,} \\ 0, & \text{в протилежному випадку,} \end{cases}$$



тобто,  $P(X_k=1)=p$ ,  $P(X_k=0)=q$ . Очевидно,  $MX_k=p$ ,

$DX_k=p-p^2=pq$ . Позначимо  $\mu_n = \sum_{k=1}^n X_k$  — кількість випробувань, в яких відбулася подія  $A$ .

Бачимо, що випадкові величини  $X_k$  задовольняють усім умовам центральної граничної теореми. Тому має місце теорема.

**Теорема (інтегральна теорема Муавра<sup>3</sup>-Лапласа).**

$$\lim_{n \rightarrow \infty} P\left(\frac{\mu - np}{\sqrt{npq}} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du.$$

При практичному застосуванні цієї теореми при великих (але скінченних)  $n$  знов використовують наближену рівність

$$P\left(\frac{\mu - np}{\sqrt{npq}} < x\right) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du = \frac{1}{2} + \Phi_0(x).$$

**Приклад.** Здійснюються випробування  $n=100$  транзисторів на надійність. Імовірність відмови одного транзистора за час випробувань  $T$  дорівнює  $p=0,2$ . Треба обчислити ймовірність того, що за час випробувань  $T$  відмовлять менше ніж  $b=28$  транзисторів.

Питання зводиться до обчислення ймовірності  $p = P(a \leq \mu < b)$ , де  $a=0$ . Застосовуємо теорему Муавра-Лапласа. Маємо:

$$P(a \leq \mu < b) = P\left(\frac{a - np}{\sqrt{npq}} \leq \frac{\mu - np}{\sqrt{npq}} < \frac{b - np}{\sqrt{npq}}\right) = P\left(\frac{\mu - np}{\sqrt{npq}} < \frac{b - np}{\sqrt{npq}}\right) - P\left(\frac{\mu - np}{\sqrt{npq}} < \frac{a - np}{\sqrt{npq}}\right) \approx \Phi_0\left(\frac{b - np}{\sqrt{npq}}\right) - \Phi_0\left(\frac{a - np}{\sqrt{npq}}\right) = \Phi_0(2) - \Phi_0(-5) = \Phi_0(2) + \Phi_0(5) \approx 0,4773 + 0,5 = 0,9773$$

---

<sup>3</sup> Муавр де Абрахам (1667 – 1754) — англійський математик, член Лондонського королівського товариства, Паризької та Берлінської АН.

### 3.4.4 Характеристичні функції випадкових векторів та центральна гранична теорема

Попередні означення та результати, одержані для одновимірних випадкових величин, природно розповсюджуються на випадкові вектори.

Нехай  $X = \begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix}$  —  $k$ -вимірний випадковий вектор-стовпець, а

$x = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$  — відповідний вектор дійсних детермінованих змінних.

Нехай  $t = (t_1, \dots, t_k)$  — вектор-рядок дійсних змінних, кожна з яких одержує значення від  $-\infty$  до  $+\infty$ .

**Означення.** Характеристичною функцією випадкового вектора  $X$  називається функція

$$\phi_X(t) = \phi_{X_1, \dots, X_k}(t_1, \dots, t_k) = Me^{itX} = Me^{i \sum_{m=1}^k t_m X_m} \quad (1)$$

(тут ми згадуємо правила множення вектора-рядка  $t$  на вектор-стовпець  $X$ ).

З урахуванням попередніх означень в одновимірному випадку та формул Ейлера маємо:

$$\phi_X(t) = Me^{itX} = M \cos tX + iM \sin tX \quad (2)$$

У випадку неперервної випадкової величини  $X$  зі щільністю розподілу  $f(x_1, \dots, x_k)$  характеристична функція має вигляд

$$\phi_X(t) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{i \sum_{m=1}^k t_m x_m} f(x_1, \dots, x_k) dx_1 \dots dx_k \quad (3)$$

Якщо  $X$  є дискретною  $k$ -вимірною випадковою величиною, що приймає значення  $x^{(1)}, \dots, x^{(s)}$ , з імовірностями  $p_r = P(X = x^{(r)})$ , то характеристична функція  $\phi_X(t)$  дорівнює

$$\phi_X(t) = \sum e^{itx^{(r)}} p_r, \quad (4)$$

де  $x^{(r)} = \begin{pmatrix} x_{r1} \\ \vdots \\ x_{rk} \end{pmatrix}, \quad r = \overline{1, s}.$

Розглянемо деякі властивості характеристичної функції  $k$ -вимірної випадкової величини:

1) Для кожного  $t \in R^k$

$$|\varphi_X(t)| \leq 1, \quad |\varphi_X(\bar{0})| = 1, \quad \text{де } \bar{0} = (0, \dots, 0).$$

Доводиться зовсім аналогічно одновимірному випадку, якщо використати (2).

2) Нехай  $Y = cX$ , де  $c$  — деяка стала. Тоді

$$\varphi_Y(t) = \varphi_X(ct) \quad (5)$$

Дійсно,

$$\varphi_Y(t) = Me^{itY} = Me^{itcX} = Me^{ictX} = \varphi_X(ct).$$

3) Якщо компоненти вектора  $X$  — незалежні випадкові величини  $X_m, m = \overline{1, k}$ , то характеристична функція дорівнює

$$\varphi_X(t) = \prod_{m=1}^k \varphi_{X_m}(t_m), \quad (6)$$

де  $\varphi_{X_m}(t) = Me^{it_m X_m}$  — характеристичні функції компонент  $X_m, m = \overline{1, k}$ .

Щоб переконатися в цьому, треба використати доведену раніш мультиплікативну властивість математичного сподівання добутку незалежних комплексних випадкових величин.

Маємо:

$$\varphi_X(t) = Me^{itX} = Me^{i \sum_{m=1}^k t_m X_m} = M \prod_{m=1}^k e^{it_m X_m} = \prod_{m=1}^k M e^{it_m X_m} = \prod_{m=1}^k \varphi_{X_m}(t_m)$$

**Приклад.** Треба відшукати характеристичну функцію вектора  $Y$ , що

має нормальний розподіл зі щільністю  $n(y, \nu, B)$ , де  $\nu = \begin{pmatrix} \nu_1 \\ \vdots \\ \nu_k \end{pmatrix}$  — вектор математичних сподівань компонент, а  $B$  — додатно означена матриця коваріацій.

Спочатку обчислимо характеристичну функцію вектора  $X$ , компоненти якого незалежні та мають нормований нормальний розподіл  $n(x_m, 0, 1)$ ,  $m = \overline{1, k}$ .

Ми знаємо, що характеристична функція компоненти  $X_m$  з таким розподілом дорівнює

$$\varphi_{X_m}(u_m) = e^{-\frac{1}{2}u_m^2}.$$

Згідно з властивістю (3) маємо:

$$\phi_X(u) = \prod_{m=1}^k e^{-\frac{1}{2}u_m^2} = e^{-\frac{1}{2} \sum_{m=1}^k u_m^2} = e^{-\frac{1}{2}uu^T}.$$

Щільність розподілу вектора  $X \in n(x, \bar{0}^T, E)$ .

Далі, з курсу алгебри відомо, що для додатно означеної симетричної матриці  $B^{-1}$  існує не вироджена матриця  $A$ , така що

$$A^T B^{-1} A = E.$$

Звідси виходить ланцюжок рівностей

$$A^T B^{-1} A = E \Rightarrow A^T B^{-1} = A^{-1} \Rightarrow B^{-1} = (A^T)^{-1} A^{-1} = (AA^T)^{-1}$$

і, таким чином,

$$B = AEA^T .$$

Якщо згадати теореми про лінійні перетворення нормально розподілених векторних випадкових величин, то можна зробити висновок,

що нормальний розподіл  $n(y, \nu, B)$  буде мати вектор

$$Y = AX + \nu .$$

Обчислюємо

$$\phi_Y(t) = Me^{itY} = Me^{it(AX + \nu)} = e^{it\nu} Me^{itAX} = e^{it\nu} e^{-\frac{1}{2}(tA)(tA)^T} ,$$

тому що  $tA = u$  .

Звідси

$$\phi_Y(t) = e^{it\nu} e^{-\frac{1}{2}tA A^T} = e^{it\nu - \frac{1}{2}tB t^T} \quad (7)$$

4) Якщо  $X(1), X(2), \dots, X(n)$  — незалежні  $k$ -вимірні випадкові вектори і  $X = X(1) + X(2) + \dots + X(n)$  , то

$$\phi_X(t) = \prod_{j=1}^n \phi_{X(j)}(t) \quad (8)$$

де  $\phi_{X(j)}(t) = Me^{itX(j)}$  — характеристичні функції векторів  $X(j)$  ,  $j = \overline{1, n}$  .

Це виходить з того, що

$$\phi_X(t) = Me^{itX} = Me^{i \sum_{j=1}^n tX(j)} = M \prod_{j=1}^n e^{itX(j)} = \prod_{j=1}^n Me^{itX(j)} .$$

5) За допомогою характеристичної функції обчислюються моменти.

Так, наприклад, якщо математичні сподівання  $MX_m$  ,  $m = \overline{1, k}$  , компонент вектора  $X$  скінченні, то

$$MX_m = \frac{1}{i} \frac{\partial \phi_X(t)}{\partial t_m} \Big|_{t=\bar{0}}, \quad m = \overline{1, k} \quad (9)$$

Дійсно (розглянемо неперервний випадок):

$$\begin{aligned} \frac{\partial \varphi_X(t)}{\partial t_m} &= \frac{\partial}{\partial t_m} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{i \sum_{m=1}^k t_m x_m} f(x_1, \dots, x_k) dx_1 \dots dx_k = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \frac{\partial}{\partial t_m} e^{i \sum_{m=1}^k t_m x_m} f(x_1, \dots, x_k) dx_1 \dots dx_k = \\ &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} i x_m e^{i \sum_{m=1}^k t_m x_m} f(x_1, \dots, x_k) dx_1 \dots dx_k ; \\ \frac{\partial \varphi_X(t)}{\partial t_m} \Big|_{t=\bar{0}} &= i \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} x_m f(x_1, \dots, x_k) dx_1 \dots dx_k = i MX_m . \end{aligned}$$

Законність заміни місцями похідної та інтеграла доводиться аналогічно одновимірному випадку. Випадок дискретної випадкової величини розгляньте самостійно.

Обчислимо  $MX_m X_l$ . За припущенням скінченності всіх  $MX_m X_l$  (з цього, зокрема, випливає і скінченність всіх  $MX_m$ ) маємо в неперервному випадку

$$\begin{aligned} \frac{\partial^2 \varphi_X(t)}{\partial t_m \partial t_l} &= \frac{\partial^2}{\partial t_m \partial t_l} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{i \sum_{m=1}^k t_m x_m} f(x_1, \dots, x_k) dx_1 \dots dx_k = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \frac{\partial^2}{\partial t_m \partial t_l} e^{i \sum_{m=1}^k t_m x_m} f(x_1, \dots, x_k) dx_1 \dots dx_k = \\ &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} i^2 x_m x_l e^{i \sum_{m=1}^k t_m x_m} f(x_1, \dots, x_k) dx_1 \dots dx_k \end{aligned}$$

Звідси:

$$MX_m X_l = - \frac{\partial^2 \phi_X(t)}{\partial t_m \partial t_l} \Big|_{t=\bar{0}}, \quad m, l = \overline{1, k} \quad (10)$$

6. Припустимо, що всі  $MX_m X_l$  скінченні.

Тоді можна довести, що має місце розклад

$$\varphi_X(t) = 1 + i \sum_{m=1}^k t_m MX_m - \frac{1}{2} \sum_{m=1}^k \sum_{l=1}^k t_m t_l MX_m X_l + o(|t|^2) \quad (11)$$

де  $|t| = |t_1| + |t_2| + \dots + |t_k|$ .

Доведення базується на розкладі функції  $\varphi_X(t)$  в ряд в околі точки  $t = \bar{0}$ .

Аналогічно одновимірному випадку мають місце наступні теореми, які ми наведемо без доведення.

**Теорема 1.** Існує взаємно однозначна відповідність між характеристичною функцією  $k$ -вимірної випадкової величини  $\phi_X(t)$  та функцією розподілу цієї величини  $F(x_1, \dots, x_k)$ , тобто якщо відома одна з них, інша по ній відновлюється однозначно.

Нехай  $X(n) = \begin{pmatrix} X_{n_1} \\ \vdots \\ X_{n_k} \end{pmatrix}$ ,  $n = \overline{1, \infty}$ , — деяка послідовність  $k$ -вимірних

випадкових величин з функціями розподілу  $F_n(x_1, \dots, x_k)$ , а  $X = \begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix}$  —  $k$ -вимірна випадкова величина з функцією розподілу  $F(x_1, \dots, x_k)$ .  
Означення слабкої збіжності  $F_n(x) \Rightarrow F(x)$  аналогічне одновимірному випадку, якщо точки неперервності визначаються по всім аргументам.

**Теорема 2.** Якщо послідовність характеристичних функцій  $\phi_{X(n)}(t)$  збігається у кожній точці  $t \in R^k$  до деякої функції  $\phi(t)$ , що неперервна у точці  $t = \bar{0}$ , то  $F_n(x) \Rightarrow F(x)$  та  $\phi(t)$  є характеристичною функцією, що відповідає  $F(x)$ .

Методом характеристичних функцій доводиться багатовимірний варіант центральної граничної теореми.

**Теорема 3.** Нехай  $X(1), X(2), \dots, X(n), \dots$  — послідовність  $k$ -вимірних

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix}$$

випадкових величин з вектором математичних сподівань та матрицею коваріацій  $B$ , елементи якої скінченні. Тоді послідовність

векторів  $Z_n = \frac{\sum_{j=1}^k (X(j) - v)}{\sqrt{n}}$  слабко збігається (збігається за розподілом) до  $k$ -вимірного випадкового вектора  $Z$ , що має нормальний розподіл  $n(z, \bar{0}^T, B)$ , тобто нормальний розподіл з нульовим вектором математичних сподівань та матрицею коваріацій  $B$ .

**Доведення.** Позначимо  $\phi_n(t)$  характеристичну функцію вектора  $\bar{X}(n) = X(n) - v$ . Оскільки  $M \bar{X}(n) = \bar{0}^T$  та  $M \bar{X}_m(n) \bar{X}_l(n) = \text{cov}(X_m(n), X_n(n)) = b_{me}$ , де  $b_{me}$  — елемент матриці коваріацій  $B$ , то за властивістю (2) та властивістю (6) характеристична функція

вектора  $\frac{\bar{X}(n)}{\sqrt{n}}$  дорівнює

$$\phi_n\left(\frac{t}{\sqrt{n}}\right) = 1 - \frac{1}{2n} \sum_{m,l=1}^k b_{me} t_m t_l + o\left(\frac{1}{n}\right).$$

Далі з того, що  $Z_n = \frac{\sum_{j=1}^k \bar{X}(j)}{\sqrt{n}}$  та незалежності  $\frac{\bar{X}(j)}{\sqrt{n}}$ ,  $j = \overline{1, n}$ , виходить (дивись властивість (4))

$$\phi_{Z_n}(t) = \left( \phi_n\left(\frac{t}{\sqrt{n}}\right) \right)^n \xrightarrow{n \rightarrow \infty} e^{-\frac{1}{2} \sum_{m,l=1}^k b_{me} t_m t_l} = e^{-\frac{1}{2} t B t^T},$$



тобто (дивись (7))  $\phi_{Z_n}(t)$  збігається до характеристичної функції розподілу з нульовим вектором математичних сподівань та матрицею коваріацій  $B$ . Застосування теорем 1 та 2 завершує доведення теореми.

## 4. Математична статистика

### 4.4 Емпірична функція розподілу

Теоретична функція розподілу  $F(x)$  невідома. Над випадковою величиною  $X$  проведено  $n$  випробувань та фіксуються результати випробувань.

Примітка. в статистиці  $x_1 \dots x_n$  – результати випробувань, а  $n$  – кількість випробувань.

Ці  $n$  чисел зветься **вибіркою об'ємом  $n$** .

По вибірці довжиною  $n$  будемо варіаційний ряд (проведемо сортування членів виборці по збільшенню, де  $x_{(1)}$  – найменше та  $x_{(n)}$  – найбільше).

Означення. Емпіричною функцією розподілу  $W_n(x)$  зветься:

$$W_n(x) = \frac{n_x}{n},$$

де  $n_x$  – кількість членів варіаційного ряду, менших за  $x$ .

$W_n(X < x)$  – еквівалент емпіричної функції розподілу.

1)  $\forall x_{(i)} \leq x \leq x_{(i+1)}, \quad i = \overline{1, n-1}$  – емпірична функція розподілу  $W_n(x)$  – постійна.

В точках  $x_{(i)}, \quad i = \overline{1, n-1}$  емпірична функція розподілу має розрив першого роду на величину  $\frac{k}{n}$ , де  $k$  – кількість членів варіаційного ряду рівних  $x_i$ .

$$2) \forall x \text{ при } n \rightarrow \infty W_n(x) \rightarrow F(x)$$

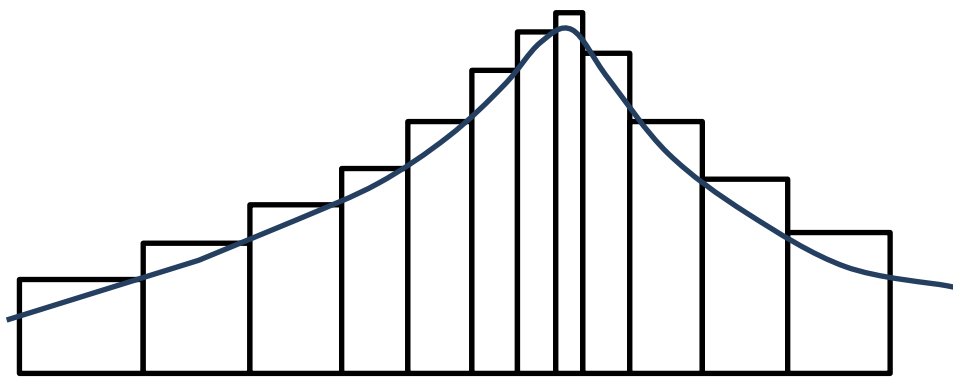
Це означає, що при достатньо великій кількості випробувань емпірична функція розподілу достатньо точно апроксимує  $F(x)$ .

Має місце **теорема Клименко**:

сходження емпіричної функції розподілу к теоретичної функції розподілу при  $n \rightarrow \infty$  рівномірно для усіх  $x$ .

#### 4.5 Гістограма (емпірична функція щільності)

Така ж сама задача: варіаційному ряду необхідно побудувати функцію щільності для його оцінки.



Прямокутна фігура апроксимує функцію тим краще, чим менша довжина відрізка. Проща кожного прямокутника рівна інтегралу від функції щільності по кінцям відрізка, що є основою прямокутника.

##### Алгоритм побудови гістограми:

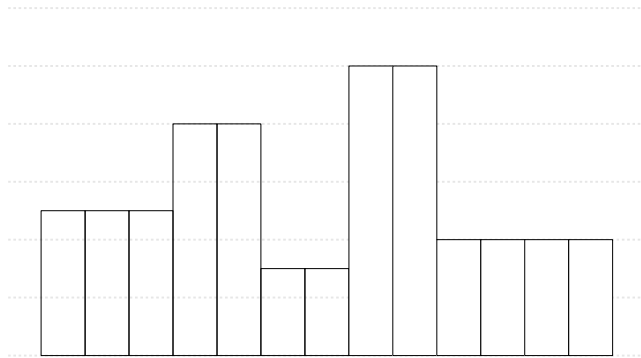
1. На числовій осі розглядають відрізок  $x(1) \dots x(n)$
2. Обираємо число  $a \in [8; 12]$

*Примітка.* Інтервал знайдено статистично

3. Відрізок  $x(1) \dots x(n)$  розіб'ємо на відрізки, кожен з котрих містить одне й теж саме число  $a$  членів варіаційного ряду та кінці відрізка співпадають з членами варіаційного ряду

4. На кожному відрізку, як на основі, будуємо прямокутники з однаковою площиною  $\frac{a}{n}$ . Побудована фігура і є гістограмою - емпіричною функцією розподілу невідомої функції щільності невідомої величини  $x$ .

5.



Неоднозначністю цього алгоритму є вибір кінців відрізків, на котрих будуються прямокутники. В деяких підручниках рекомендується відрізок  $(x(1), x(n))$  розбивати на відрізка рівної довжини (а не по кількості членів варіаційного ряду).

Насправді, гістограма сама по собі нікому не потрібна та використовується тільки для вирішення наступної задачі – I етап розділу мат. статистики «Перевірка складної гіпотези». Суть полягає в наступному: існують статистичні таблиці, де на графіку приведені усі найбільш відомі функції щільності випадкових величин і для кожної функції щільності наведені інженерні задачі, де найчастіше зустрічаються ці функції. Тоді на першому етапі дослідник будує гістограму, порівнює її зі усіма відомими графіками (усіх відомих йому теоретичних функцій щільності), та висуває гіпотезу, що досліджувана випадкова величина  $X$

має визначену функцію щільності. А далі необхідно перевірити цю гіпотезу.

#### 4.6 Теорія вибіркового методу.

Це перший розділ мат.статистики, де вирішується наступна задача: є неперервна випадкова величина  $X$  та відома її функція щільності з точністю до числових значень параметрів

Наприклад.  $\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\vartheta)^2}{2\sigma^2}}$  має два числових параметра:  $\vartheta$  та  $\sigma$ .

Використовуючи вибірку об'єму  $n$  (результати  $n$  випробувань), необхідно оцінити невідомі значення  $n$  параметрів (в нашому прикладі необхідно оцінити математичне сподівання та дисперсію).

Така постановка задачі позначає коли досліджувана випадкова величина задовольняє умовам граничної теореми (з інженерної точки зору)

Приклад. Виявилось, що інтегральна похибка вимірювань задовольняє загальній центральній граничній теоремі (є сумою 12 незалежних складових, кожна з яких не домінує над іншими). Тоді вважається, що інтегральна похибка вимірювань має нормальний розподіл та по результатам вимірювань необхідно оцінити математичне сподівання і дисперсію.

Існує два способи оцінки невідомих параметрів:

1) **точковий** – оцінкою параметра  $\theta_i, i=\overline{1,n}$  є числова скалярна функція, аргументи якої є членами вибірки  $\theta_i \approx \hat{\theta}_i(x_1, \dots, x_n)$  (функція  $\hat{\theta}_i(u_1, \dots, u_n)$ )

2) **інтервальна оцінка**. Вводяться дві функції  $n$  аргументів, в котрі підставляються члени вибірки (ряду):

$\underline{\theta}_i(x_1, \dots, x_n)$  та  $\overline{\theta}_i(x_1, \dots, x_n)$ .

Ці функції задовольняють умові:

з імовірністю  $1-\alpha$  ( $\alpha \leq 0.05$ ;  $\alpha \neq 0$ ) має виконуватися:

$$\underline{\theta}_i(x_1, \dots, x_n) \leq \theta \leq \bar{\theta}_i(x_1, \dots, x_n)$$

Математична статистика побудована на відомому прийомі:

маємо фізично існуючу випадкову величину  $X$  та  $n$  результатів випробування на нею ( $x_1 \dots x_n$  – вибірка об'єму  $n$ ). Можна вважати, що ці  $n$  результати –  $n$  повторів одного випробування над випадковою величиною  $X_1 X_2 \dots X_n$ , де  $X_i, i=\overline{1, n}$  – незалежні копії випадкової величини  $X$  (тобто число  $x_i, i=\overline{1, n}$  – результат випробування над  $X_i, i=\overline{1, n}$ )

#### 4.7 Побудова точкових оцінок

Означення. **Вибірковою характеристикою**, побудованою для оцінки невідомого параметра  $\theta_i, i=\overline{1, c}$ , зветься числова скалярна функція  $n$  випадкових аргументів, що є незалежними копіями фізично існуючої випадкової величини  $X$ :

$$\hat{\theta}_i(X_1, \dots, X_n)$$

Оцінка невідомого параметра береться як реалізація цієї випадкової величини:

$$\theta_i \approx \hat{\theta}_i(x_1, \dots, x_n),$$

де  $x_1$  – реалізація випадкової величини  $X_1, \dots, x_n$  – реалізація випадкової величини  $X_n$ ;

$x_1, \dots, x_n$  – вибірка, об'ємом  $n$ .

Властивості, яким повинна задовольняти вибіркова характеристика, щоб будь-яка її реалізація добре оцінювала параметр  $\theta_i, i=\overline{1, c}$ :

##### 1) Незміщеність

$$M \hat{\theta}_i(X_1, \dots, X_n) = \theta_i$$

В цьому випадку результати конкретних випробувань випадкової величини  $\hat{\theta}_i(X_1, \dots, X_n)$  будуть групуватись відносно свого математичного сподівання -  $\theta_i$ .

Обґрунтування:

$$M \hat{\theta}_i(X_1, \dots, X_n) = \theta_i$$

Якби  $M \hat{\theta}_i(X_1, \dots, X_n) \neq \theta_i$ , то результати конкретних випробувань  $\hat{\theta}_i$  групувалися б не навколо  $\theta_i$  були б поганою оцінкою  $\theta_i$ .

## 2) Ефективність

Вибіркова характеристика є **ефективною**, якщо вона є незміщеною і має найменшу дисперсію з усіх можливих незміщених вибірових характеристик, побудованих для оцінки  $\theta_i$ .

Обґрунтування: чим менша дисперсія, тим ближче групуються випробування  $\hat{\theta}_i$  відносно  $M \hat{\theta}_i$ .

## 3) Спроможність

Вибіркова характеристика  $\hat{\theta}_i(X_1, \dots, X_n)$  зветься **спроможною**, якщо при  $n \rightarrow \infty$  ряд вибірових характеристик збігається до  $\theta_i$ :

$$\lim_{n \rightarrow \infty} \hat{\theta}_i(X_1, \dots, X_n) = \theta_i$$

(з ймовірністю 1, по ймовірності, в середньостатичному)

Приклади.

1)  $v = MX$  – невідомий параметр.

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

$$M \bar{X}_n = \nu$$

2)  $\sigma^2$  – невідомий параметр.

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

$$MS_n^2 = M \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

$$= M \frac{1}{n} \sum_{i=1}^n (X_i^2 - 2X_i \bar{X}_n + \bar{X}_n^2)$$

$$=$$

$$\sigma^2 - 2D \bar{X}_n + D \bar{X}_n = \sigma^2 - D \bar{X}_n = \left[ D \frac{1}{n} \sum_{i=1}^n X_i^2 - \frac{\sigma^2}{n} \right] = \sigma^2 - \frac{\sigma^2}{n}$$

Вибіркова характеристика  $S_n^2$  є **зміщеною вибірковою характеристикою** відносно параметра  $\sigma^2$ . Але одночасно вона є асимптотично незміщеною на нескінченності:

$$MS_n^2 \rightarrow \sigma^2$$

$$n \rightarrow \infty$$

3) Вибіркова характеристика  $\bar{X}_n$  для оцінки вибіркового параметра  $\nu$  є спроможною. Відповідно до посиленого закону великих чисел:

$$\frac{1}{n} \sum_{i=1}^n X_i \rightarrow \nu \text{ (з ймовірністю 1)}$$

$$n \rightarrow \infty$$

4) Задача(приклад на ефективність).

Маємо  $X_1, \dots, X_n$  незалежних однаково розподілених випадкових величини з невідомим, але однаковим математичним сподіванням  $\nu$ , і різними, але відомими дисперсіями  $\sigma_i^2, i=1, \dots, n$ . Провели випробування, отримали результати випробувань  $x_1, \dots, x_n$ , де  $x_i$  – результат випробування над випадковою величиною  $X_i$ . За результатами випробувань треба оцінити  $\nu$ .



### Фізична інтерпретація постановки задачі

Є одна фізична величина, яка вимірюється  $n$  приладами, в кожного з яких нульова систематична помилка, а точність оцінюється дисперсіями. Кожним приладом виміряли цю фізичну величину. По результатам випробувань потрібно оцінити цю фізичну величину

### Розв'язок.

Будемо шукати розв'язок задачі в класі ефективних, лінійних, незміщених оцінок математичного сподівання.

Означення. Вибіркова характеристика зветься *лінійною*, якщо

$$v = M \sum_{i=1}^n g_i X_i$$

З умови незміщеності випливає, що  $\sum_{i=1}^n g_i = 1$ .

Тоді:

$$D \sum_{i=1}^n g_i X_i = g_1^2 \sigma_1^2 + g_2^2 \sigma_2^2 + \dots + g_{n-1}^2 \sigma_{n-1}^2 + (1 - g_1 - g_2 - \dots - g_{n-1})^2 \sigma_n^2$$

З умови ефективності відомо, що цей вираз має бути мінімальним. Так як цей вираз квадратичний, то він має один екстремум. Для його знаходження беремо часткові похідні по  $g_i$  і прирівнюємо їх до нуля:

$$2g_i \sigma_i^2 = 2g_n \sigma_n^2, \quad i=1, n-1$$

Тепер необхідно довести, що ця система має один розв'язок. Тоді будь-який розв'язок системи є єдиним. Для цього необхідно показати, що детермінант рівнянь лінійної системи відносно  $g_1, \dots, g_{n-1}$  дорівнює нулю. Показати самостійно.

Для знаходження коефіцієнтів  $g_1, \dots, g_n$  підставляємо будь-яке значення для  $g_n$  і з перших  $n-1$  рівнянь знаходимо  $g_1, \dots, g_{n-1}$ . Тоді:

$$\hat{g}_i = \frac{g_i}{\sum_{i=1}^n g_i}$$

Показати самим, що  $\hat{g}_i$  задовольняють усім рівнянням системи, тобто є розв'язком.

#### 4.8 Метод найбільшої правдоподібності

##### Постановка задачі.

##### а) Неперервний випадок:

$X$  – неперервна випадкова величина. Є функція щільності  $f(u, \theta_1, \dots, \theta_c)$  з невідомим параметром та вибірка  $x_1, \dots, x_n$ . Вважаємо  $x_1, \dots, x_n$  результатами випробувань над  $X_1 \dots X_n$ , де  $X_i$ ,  $i = \overline{1, n}$  – незалежні копії  $X$ . Знайдемо  $n$ -вимірну функцію щільності  $n$ -мірної випадкової величини  $X_1, \dots, X_n$ . Вона дорівнює:

$$f(u_1, \dots, u_n) = \prod_{i=1}^n f(u_i, \theta_1, \dots, \theta_c)$$

**Означення.** **Функція найбільшої правдоподібності** – числова скалярна функція  $c$ -аргументів  $\theta_1 \dots \theta_c$ , що дорівнює:

$$L(\theta_1, \dots, \theta_c) = \prod_{i=1}^n f(x_i, \theta_1, \dots, \theta_c)$$

**Примітка.** В  $n$ -вимірній функції щільності зафіксували аргументи – члени вибірки: замість  $u_i = x_i$ ,  $i = \overline{1, n}$ .

##### б) Дискретний випадок:

$$X \begin{pmatrix} y_1 & \dots & y_s \\ P_1(\theta_1, \dots, \theta_c) & \dots & P_s(\theta_1, \dots, \theta_c) \end{pmatrix}$$

##### Наприклад.

урна з шариками (білими / чорними)

$$P_1 = \frac{\theta}{N},$$

де  $\theta$  – кількість білих шариків, та

$$P_2 = 1 - \frac{\theta}{N}$$

Провели  $n$  випробувань, в яких  $Y_1$  настало  $m_1$  раз,  $Y_2 - m_2$  раз, ... ,  $Y_s - m_s$

раз ( $\sum_{i=1}^s m_i = n$ )

**Функцією найбільшої правдоподібності** називається вираз:

$$L(\theta_1, \dots, \theta_c) = \prod_{i=1}^n P_i^{m_i}(\theta_1, \dots, \theta_c)$$

В якості оцінки невідомих параметрів  $\theta_1 \dots \theta_c$  Роберт Фішер запропонував брати такі значення  $\theta_1 \dots \theta_c$ , на котрих досягається максимум функції найбільшої правдоподібності, тобто

$$\max_{\theta_1 \dots \theta_c} L(\theta_1, \dots, \theta_c)$$

Обґрунтування методу найбільшої правдоподібності.

В якості оцінки невідомих параметрів беруться ті, які максимізують ймовірність настання події, котре настало внаслідок випробування.

а) Дискретний випадок

$$P = \prod_{i=1}^s P_i^{m_i}(\theta_1, \dots, \theta_c) = L(\theta_1, \dots, \theta_c)$$

(тобто максимізуємо ймовірність того, що настали саме ті результати випробування, які маємо)

б) Неперервний випадок

Розглянемо ймовірність того, що внаслідок проведених випробувань настане подія:

$$P(x_i \leq X_i \leq x_i + \Delta x_i) = L(\theta_1, \dots, \theta_c) \prod_{i=1}^n \Delta x_i + o(\prod_{i=1}^n \Delta x_i), i = \overline{1, n}$$

( $X_i$  потрапило в відрізок  $x_i \leq X \leq x_i + \Delta x_i$ , де  $\Delta x_i \rightarrow 0$ )

Таким чином максимізація  $L$  по параметрам  $\theta_1 \dots \theta_c$  еквівалентна максимізації ймовірності потрапляння випадкової величини  $X_1 \dots X_n$  в прямокутник, в який вона попала внаслідок випробування. Отримали задачу на безумовний екстремум:

$$\frac{\partial L(\theta_1, \dots, \theta_c)}{\partial \theta_i} = 0, i = \overline{1, n}$$

Розв'язок. Необхідно знайти усі рішення цієї системи рівнянь та обрати те, котрому відповідає максимум функції найбільшої правдоподібності.

Задача спрощується, якщо розглядати

$$\frac{\partial \ln L(\theta_1, \dots, \theta_c)}{\partial \theta_i} = 0, i = \overline{1, n},$$

тобто обирають похідні не від  $L$ , а від  $\ln L$ , оскільки  $\ln \square$  — монотонно зростаюча функція, тому максимум  $\ln L$  не співпадає з максимумом  $L$ , але досягається на тих же самих значеннях аргументів. Оскільки  $L$  — добуток функцій, то  $\ln L$  — сума, тобто система рівнянь спрощується.

Теорема (без доведення).

Оцінки параметрів, отримані методом найбільшої правдоподібності є спроможними та асимптотично незміщеними, ефективними та розподіленими нормально (є протиріччя: оцінка розподілена нормально та є спроможною, але при  $n \rightarrow \infty$  нормальний розподіл вироджується в  $\delta$ -функцію).

Приклад.  $X$  має нормальний розподіл, тобто  $\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\nu)^2}{2\sigma^2}}$  та є вибірка  $x_1, \dots, x_n$ . Треба знайти параметри  $\nu$  та  $\sigma^2$ .

$$\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \prod_{i=1}^n e^{\frac{-(u_i-v)^2}{2\sigma^2}} = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} e^{\frac{-\sum_{i=1}^n (u_i-v)^2}{2\sigma^2}}$$

( $n$ -вимірна функція щільності).

Функція найбільшої правдоподібності має вигляд:

$$L(\theta_1, \dots, \theta_c) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} e^{\frac{-\sum_{i=1}^n (x_i-v)^2}{2\sigma^2}} = L(v, \sigma^2)$$

Якою б не була оцінка  $\sigma^2$ , найбільше значення  $L$  досягає при мінімуму

$$\sum_{i=1}^n (x_i - v)^2 \text{ по } v, \text{ тобто}$$

$$\min_v \sum_{i=1}^n (x_i - v)^2$$

Звідси випливає, що  $\hat{v} = \frac{1}{n} \sum_{i=1}^n x_i$  — оцінка  $v$ .

Примітка. Отримали реалізацію вибіркової характеристики оцінки математичного сподівання виду:

$$\acute{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

$$(\text{а її реалізація: } \frac{1}{n} \sum_{i=1}^n x_i)$$

Замість математичного сподівання  $v$  в  $L(v, \sigma)$  підставляємо  $\acute{X}_n$  та знаходимо  $\arg \max_{\sigma^2} \ln L(\acute{X}, \sigma^2)$ . Отримуємо:

$$\hat{\sigma}^2 = S_1^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \acute{X}_n)^2$$

(реалізація вибіркової характеристики, наведеної нижче)

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \acute{X}_n)^2$$

## 4.9 Інтервальні оцінки

Розглянемо наступні спеціальні розподіли:

### 4.9.1 Розподіл $\chi^2$

Означення. Випадкова величина має **розподіл  $\chi^2$  з  $n$ -степенями**

**свободи**, якщо її можна представити у вигляді суми  $\chi^2 = \sum_{i=1}^n X_i^2$ , де  $X_1 \dots X_n$  є

незалежними і розподіленими нормовано нормально  $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ .

Знайдемо функцію щільності розподілу  $\chi^2$ . Для цього побудуємо її функцію розподілу:

$$K_n(x) = P(\chi^2 \leq x) = \begin{cases} K_n(x), & P\left(\sum_{i=1}^n X_i^2 \leq x\right), \\ 0, & \end{cases}$$

(функція розподілу, а значить і функція щільності = 0 для від'ємних аргументів. Розглядаємо випадки, коли аргументи ф-ції розподілу є невід'ємними)

$$\left(\frac{1}{\sqrt{2\pi}}\right) \int \int_{\sum_{i=1}^n x_i^2 \leq x} e^{-\frac{\sum_{i=1}^n x_i^2}{2}} dx_1 \dots dx_n = P\left(\sum_{i=1}^n X_i^2 \leq x\right)$$

Примітка.  $x_i$  - аргументи  $n$ -вимірної ф-ції щільності  $n$ -вимірної випадкової величини  $X_1 \dots X_n$

$$0 < \theta < 1$$

$$I = \int \int_{\sum_{i=1}^n x_i^2 \leq x} dx_1 \dots dx_n$$

Розглянемо інтеграл

і зробимо заміну змінних

$$x_i = y_i \sqrt{x}, i = \overline{1, n}$$

$$dx_i = dy_i \sqrt{x}$$

$$I = x^{\frac{n}{2}} \iint_{\sum_{i=1}^n y_i^2 \leq 1} dy_1 \dots dy_n = C_n x^{\frac{n}{2}}$$

Тоді  $C_n$  - константа, від  $x$  не залежить) є

об'ємом  $n$ -вимірної сфери радіусом 1, з центром в початку координат

$$K_n(x+h) - K_n(x) = \left( \frac{1}{\sqrt{2\pi}} \right)^n C_n e^{\frac{-(x+\theta h)}{2}} ((x+h)^{\frac{n}{2}} - x^{\frac{n}{2}})$$

$$\lim_{\Delta x \rightarrow 0} \frac{K_n(x+h) - K_n(x)}{h} = \left( \frac{1}{\sqrt{2\pi}} \right)^n C_n e^{\frac{-(x+\theta \Delta x)}{2}} \frac{((x+h)^{\frac{n}{2}} - x^{\frac{n}{2}})}{h} = T_n x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$$

$$k_n(x) = T_n x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, x \geq 0, T_n = \frac{1}{\sqrt{2\pi}} C_n \frac{n}{2}$$

Знайдемо компактний вигляд для константи  $T$

$$k_n(x) = \begin{cases} T_n x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0 \\ 0, & x < 0 \end{cases}$$

$$\int_0^{\infty} T_n x^{\frac{n}{2}-1} e^{-\frac{x}{2}} dx = 1$$

,

$$T_n = \frac{1}{\int_0^{\infty} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} dx}$$

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

$$\Gamma(p) = (p-1)\Gamma(p-1)$$

$$\frac{x}{2} = y$$

$$x = 2y$$

$$dx = 2dy$$

$$T_n = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)}$$

Знаходимо мат. сподівання, дисперсію, характеристичну функцію розподілу  $\chi^2$

$$M(\chi^2) = \int_0^{\infty} x k_n(x) dx$$

$$M(\chi^2)^2 = \int_0^{\infty} x^2 k_n(x) dx$$

$$D(\chi^2) = M(\chi^2)^2 - (M\chi^2)^2$$

$$\varphi_{\chi^2}(t) = \int_0^{\infty} e^{itx} k_n(x) dx$$

Шукаємо перший початковий момент:

$$M(\chi^2) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} x^{\frac{n}{2}} e^{-\frac{x}{2}} dx = \frac{2^{\frac{n}{2}+1} \Gamma\left(\frac{n}{2}+1\right)}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} = n$$

Шукаємо другий початковий момент:

$$M(\chi^2)^2 = \frac{2^{\frac{n}{2}+2} \Gamma\left(\frac{n}{2}+2\right)}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} = n\left(\frac{n}{2}+1\right) = n^2 + 2n$$

$$DX = 2n$$

$$\varphi_x(t) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} x^{\frac{n}{2}-1} e^{-\frac{x}{2}(1-it)} dx = \frac{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right) (1-2it)^{\frac{n}{2}}} = (1-2it)^{-\frac{n}{2}}$$



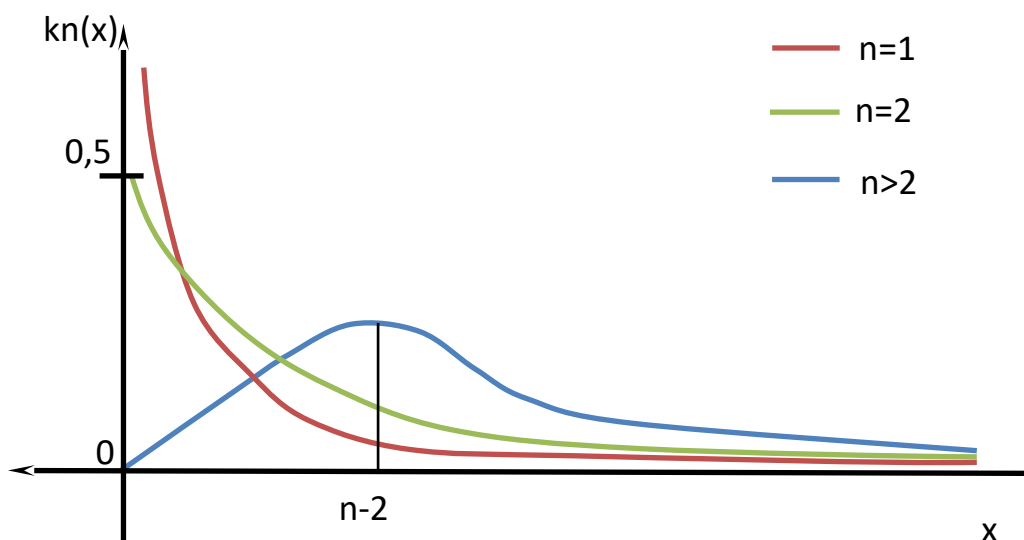
Наслідок. Сума двох незалежних випадкових величин, що має розподіл

$\chi^2$  з  $n_1$  та  $n_2$  ступенями свободи відповідно має розподіл  $\chi^2 + \chi^2$  сумарної

кількості ступенів свободи  

$$\varphi_{(\chi_1^2 + \chi_2^2)}(t) = (1 - 2it)^{-\frac{n_1}{2}} (1 - 2it)^{-\frac{n_2}{2}} = (1 - 2it)^{-\frac{n_1 + n_2}{2}}$$

Доведення завершено, тому що характеристична функція однозначно задає функцію щільності випадкової величини.



#### 4.9.2 Розподіл Ст'юдента

Означення. Випадкова величина з  $k$  ступенями свободи **має розподіл Ст'юдента**, якщо

$$T = \frac{Z\sqrt{k}}{\sqrt{V}}$$
, де  $Z$  розподілено нормовано нормально,  $V$  має розподіл  $\chi^2$  з  $k$  ступенями свободи.

Знайдемо ф-цію щільності випадкової величини Т.

Т приймає значення на всій числовій осі.

Як і в попередньому випадку

$$S_k(x) = P\left(\frac{Z\sqrt{k}}{\sqrt{V}} \leq x\right) = P\left(Z \leq x \frac{\sqrt{V}}{\sqrt{k}}\right) = P\left(ZV \in \left\{ \begin{array}{l} 0 \leq Z < \infty \\ -\infty < Z \leq \frac{\sqrt{V}}{\sqrt{k}} \end{array} \right\}\right)$$

Беремо інтеграл від двовимірної ф-ції щільності по цій області

$$\frac{1}{\sqrt{2\pi} 2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} \iint_{\substack{0 \leq v < \infty \\ -\infty < z \leq x \frac{\sqrt{v}}{\sqrt{k}}}} v^{\frac{k}{2}-1} e^{-\frac{v}{2}} e^{-\frac{z^2}{2}} dz dv = T_k \int_0^\infty v^{\frac{k}{2}-1} e^{-\frac{v}{2}} \left[ \int_{-\infty}^{x \frac{\sqrt{v}}{\sqrt{k}}} e^{-\frac{z^2}{2}} dz \right] dv$$

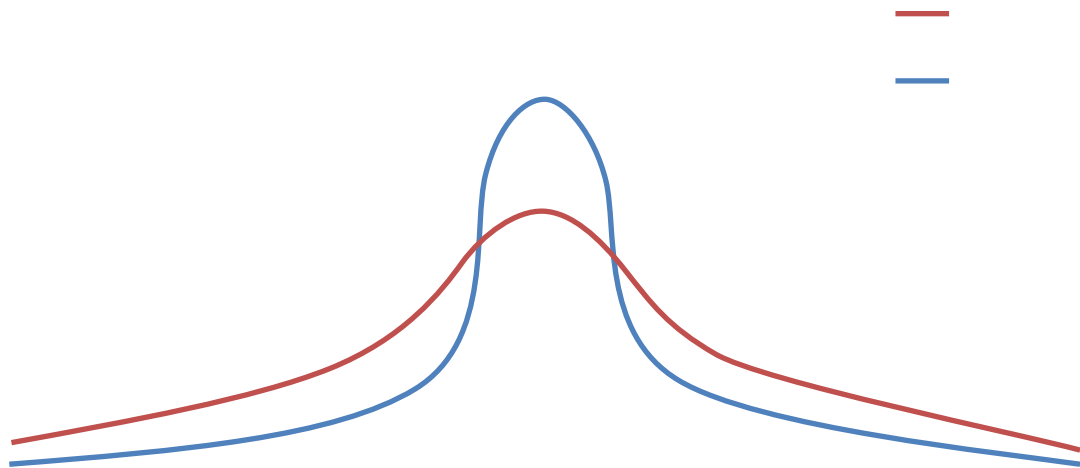
Знаходимо ф-цію щільності як похідну від ф-ції розподілу

$$S_k(x) = \frac{dS_k(x)}{dx} = T_k \int_0^\infty v^{\frac{k}{2}-1} e^{-\frac{v}{2}} \left[ \frac{d}{dx} \int_{-\infty}^{x \frac{\sqrt{v}}{\sqrt{k}}} e^{-\frac{z^2}{2}} dz \right] dv$$

Треба зробити таку заміну змінних, щоб у внутрішньому інтегралі був тільки х:

$$\left( \begin{array}{l} z = y \frac{\sqrt{v}}{\sqrt{k}} \\ dz = \frac{\sqrt{v}}{\sqrt{k}} dy \\ \frac{d}{dx} \int_{-\infty}^\infty f(u) du = f(x) \end{array} \right)$$

$$S_K(x) = \frac{T_k}{\sqrt{k}} \int_0^\infty v^{\frac{k-1}{2}} e^{-\frac{v}{2}} e^{-\frac{x^2 v}{2k}} dv = \frac{T_k}{\sqrt{k}} \int_0^\infty v^{\frac{k-1}{2}} e^{-\frac{v}{2} \left(1 + \frac{x^2}{k}\right)} dv = \frac{T_k 2^{\frac{k+1}{2}} \Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k} \left(1 + \frac{x^2}{k}\right)^{\frac{k+1}{2}}} = D_k \left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}}$$



При великих  $k$  розподіл Ст'юдента практично збігається з нормованим нормальним розподілом.

#### 4.9.3 Розподіл Фішера

Означення. Випадкова величина  $F$  має **розподіл Фішера** з  $n_1$  і  $n_2$  ступенями свободи якщо:

$$F = \frac{\frac{V_1}{n_1}}{\frac{V_2}{n_2}},$$

де  $V_1, V_2$  - незалежні випадкові величини, розподілені по закону  $\chi^2$  з  $n_1$  і  $n_2$  ступенями свободи відповідно.

Функція щільності випадкової величини  $F$ :

$$f_F(x) = \begin{cases} \frac{Cx^{\frac{n_1-n_2}{2}}}{(n_2+n_1x)^{\frac{n_1+n_2}{2}}}, & x \geq 0 \\ x < 0 \end{cases}$$

#### 4.9.4 Розподіл $\frac{nS^2}{\sigma^2}$

Випадкові величини  $X_1, \dots, X_n$  незалежні, розподілені нормально  $n(\mu, \sigma^2)$ .  
 . Випадкова величина

$$nS^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2,$$

де  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ .

Знайти випадкову величину  $\frac{nS^2}{\sigma^2}$ .

Розв'язання.

Виконаємо заміну:  $Z_i = X_i - \bar{X}_n = X_i - \frac{1}{n} \sum_{i=1}^n X_i$ ,  $i=1, \dots, n$ . Випадкові величини  $Z_i$  мають нормальний розподіл.

Випадкові величини  $Z_i$  залежні (зв'язані лінійно), бо  $\sum_{i=1}^n Z_i = 0$ . Тому знаходження розподілу  $\frac{nS^2}{\sigma^2}$  стає нетривіальним.

Покажемо, що випадкова величина  $nS^2$  не залежить від математичного сподівання  $X_i$ . Дійсно, у вираз  $nS^2$  підставимо  $X_i + M$ , де  $M$  – будь-яке дійсне число:

$$nS^2 = \sum_{i=1}^n \left( X_i + M - \frac{1}{n} \sum_{i=1}^n (X_i + M) \right)^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

Тому можна вважати, що  $M = 0$  (Насправді  $M = X_i$  може бути не рівним нулю, але це не впливає на  $nS^2$ ).

Розглянемо перетворення випадкового вектора  $X$

$$Y = AX,$$

$$\text{де } X = \begin{pmatrix} X_1 \\ \dots \\ X_n \end{pmatrix}, Y = \begin{pmatrix} Y_1 \\ \dots \\ Y_n \end{pmatrix}.$$

Розпишемо цю систему скалярно:

$$Y_1 = (X_1 - X_2) \frac{1}{\sqrt{2}}$$

$$Y_2 = (X_1 + X_2 - 2X_3) \frac{1}{\sqrt{6}}$$

...

$$Y_i = (X_1 + \dots + X_i - X_{i+1}) \frac{1}{\sqrt{i(i+1)}}$$

...

$$Y_{n-1} = (X_1 + \dots + X_{n-1} - X_n) \frac{1}{\sqrt{(n-1)n}}$$

$$Y_n = \frac{1}{\sqrt{n}} X_1 + \frac{1}{\sqrt{n}} X_2 + \dots + \frac{1}{\sqrt{n}} X_n = \sqrt{n} \dot{X}_n$$

Запишемо матрицю  $A$ :

$$\begin{pmatrix} \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} & 0 & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{\sqrt{i(i+1)}} & \dots & \left( \frac{1}{\sqrt{i(i+1)}} - \frac{i}{\sqrt{i(i+1)}} \right) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{\sqrt{(n-1)n}} & \dots & \dots & \dots & \dots & \left( \frac{1}{\sqrt{(n-1)n}} - \frac{n-1}{\sqrt{(n-1)n}} \right) \\ \frac{1}{\sqrt{n}} & \dots & \dots & \dots & \dots & \frac{1}{\sqrt{n}} \end{pmatrix}$$

Самим довести наступні **властивості матриці**  $A$ :

$$\sum_{j=1}^n a_{i_1 j} a_{i_2 j} = 0, \forall i_1, i_2, i_1 \neq i_2$$

$$\sum_{i=1}^n a_{i j_1} a_{i j_2} = 0, \forall j_1, j_2, j_1 \neq j_2$$

$$\sum_{j=1}^n a_{ij}^2 = 1, i=1, \dots, n$$

$$\sum_{i=1}^n a_{ij}^2 = 1, j=1, \dots, n$$

$A$  – ортогональна матриця (модуль детермінанта матриці рівний одиниці),  $Y = AX$ . З цього випливає, що

$$\sum_{i=1}^n X_i^2 = \sum_{i=1}^n Y_i^2$$

Тобто, лінійне перетворення матриці  $A$  є поворотом координат.

Властивості компонент випадкового вектора  $Y$ :

$$1) M Y_i = 0, D Y_i = \sigma^2, i=1, \dots, n.$$

Довести самим, використовуючи властивості матриці  $A$ .

2)  $Y_1 \dots Y_n$  –  $n$ -вимірний випадковий вектор, який має  $n$ -вимірний нормальний розподіл, тому що квадратна матриця  $A$  є ортогональною, а тому – невиродженою.

Наслідок. Випадкові величини  $Y_i$  розподілені нормально з

$$M Y_i = 0, D Y_i = \sigma^2, i=1, \dots, n.$$

$$3) \text{cov}(Y_i, Y_j) = 0, \forall i \neq j, \text{ бо:}$$

$$1. \text{cov}(Y_i, Y_j) = M(Y_i, Y_j) \text{ у нашому випадку.}$$

2.  $M(d_1 X_i d_2 X_j) = d_1 d_2 \text{cov}(X_i, X_j) = 0$ , так як випадкові величини  $X_i, X_j$  незалежні.

$$3. M(d_1 X_i d_2 X_i) = d_1 d_2 \sigma^2, \text{ бо } M X_i^2 = \sigma^2.$$

Наслідок. Випадкові величини  $Y_i$  розподілені нормально за законом  $n(x, 0, \sigma)$ ; є незалежними, бо  $\text{cov}(Y_i, Y_j) = 0$ .

Таким чином,

$$nS^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2 = \sum_{i=1}^n X_i^2 - 2n\bar{X}_n^2 + n\bar{X}_n^2 = \sum_{i=1}^n X_i^2 - n\bar{X}_n^2 = \sum_{i=1}^n Y_i^2 - \bar{Y}_n^2 = \sum_{i=1}^{n-1} Y_i^2$$

$$nS^2 = \sum_{i=1}^{n-1} Y_i^2$$

$$\frac{nS^2}{\sigma^2} = \sum_{i=1}^{n-1} \left( \frac{Y_i}{\sigma} \right)^2$$

Таким чином,  $\frac{nS^2}{\sigma^2}$  – розподіл  $\chi^2$  з  $n-1$  степенями свободи.

Примітка. у нас зменшилась на одиницю кількість доданків при

переході від  $X_i$  до  $Y_i$ , бо існує один лінійний зв'язок  $\sum_{i=1}^n (X_i - \bar{X}_n) = 0$

$$\bar{X}_n = \frac{\bar{Y}_n}{\sqrt{n}}$$

Випадкові величини  $\bar{X}_n, \frac{nS^2}{\sigma^2}$  – незалежні (показати самим).

#### 4.10 Інтервальні оцінки для параметрів нормального розподілу

Побудова довірчих інтервалів для параметрів нормального розподілу.

Є випадкова величина  $X, n(x, \nu, \sigma)$ . Провели  $n$  випробувань, маємо

вибірку  $x_1 \dots x_n$ , яку вважаємо результатом одного випробування над  $n$ -

вимірною випадковою величиною  $X_1 \dots X_n$

$$P(\theta_i(x_1 \dots x_n) \leq \theta_i \leq \bar{\theta}_i(x_1 \dots x_n)) = 1 - \alpha, \alpha \leq 0.05$$

Потрібно зробити 2 вибіркові характеристики. Розглянемо випадкову величину

$$\frac{\bar{X}_n - \nu}{\frac{\sigma}{\sqrt{n}}}, \text{ де } \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

Знайдемо розподіл цієї випадкової величини.  $\overline{X}_n$  має нормальний розподіл (дивись наслідок теореми про лінійне перетворення n-вімерного нормально розподіленого випадкового вектора. В нашому випадку rang A=1 )

$$M \dot{X}_n = v$$

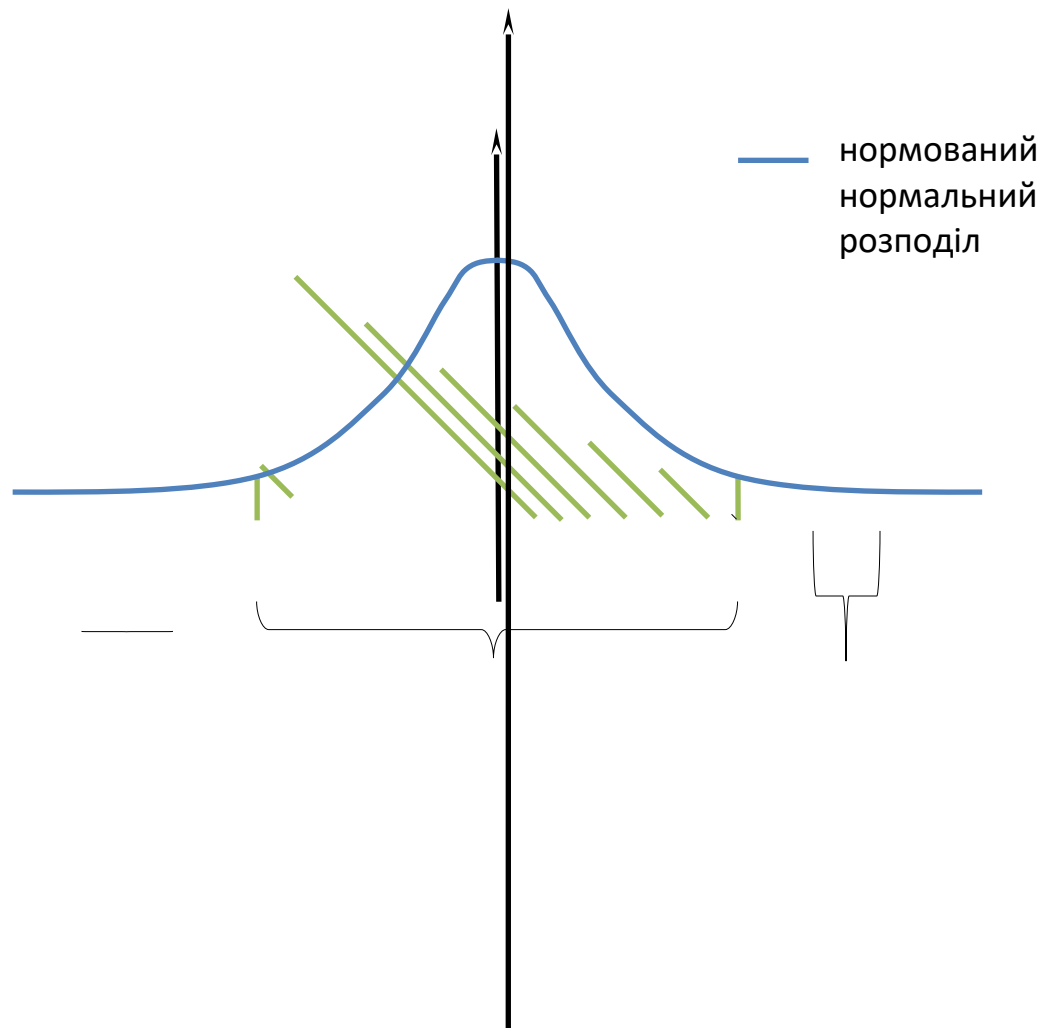
$$D X_n \frac{\sigma^2}{n} \text{Таким чином}$$

$\frac{\dot{X}_n - v}{\frac{\sigma}{\sqrt{n}}}$  має нормований нормальний розподіл.

Знайдемо таке  $t_\alpha$ , для якого виконується наступне рівняння:

$$P \left( \left| \frac{\overline{X}_n - v}{\frac{\sigma}{\sqrt{n}}} \right| \leq t_\alpha \right) = 1 - \alpha$$





$$\Phi_0(t_\alpha) = \frac{1-\alpha}{2}$$

$$t_\alpha = \Phi_0^{-1}\left(\frac{1-\alpha}{2}\right)$$

де  $\Phi_0^{-1}$  - обернена функція Лапласа від додатного значення аргумента.

Знаходимо  $t_\alpha$  по таблицям для оберненої функції Лапласа і будуємо інтервальну оцінку.

$$P\left(\bar{X}_n - t_\alpha \frac{\sigma}{\sqrt{n}} \leq \bar{X}_n \leq \bar{X}_n + t_\alpha \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

$$\bar{x}_n - t_\alpha \frac{\sigma}{\sqrt{n}} \leq v \leq \bar{x}_n + t_\alpha \frac{\sigma}{\sqrt{n}}$$

Будуємо інтервальну оцінку і з неї впливає невідомий параметр

$$\underline{\theta}_i(X_1 \dots X_n) = \bar{X}_n - t_\alpha \frac{\sigma}{\sqrt{n}}$$

$$\bar{\theta}_i(X_1 \dots X_n) = \bar{X}_n + t_\alpha \frac{\sigma}{\sqrt{n}}$$

Наслідок 1: обов'язково існує ненульова ймовірність помилитись, бо при обмеженому  $n$  і  $\alpha = 0, t_\alpha = +\infty$

Наслідок 2: це той випадок, коли до випробувань заздалегідь можна знайти кількість випробувань щоб оцінити невідоме мат сподівання із заданою точністю.

Інтервальна оцінка має еквівалентну назву «довірчий інтервал».

### *Інтервальна оцінка для математичного сподівання при невідомій дисперсії*

$$\frac{\frac{\bar{X}_n - \nu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{nS_n^2}{\sigma^2}}}, \quad \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i, \quad S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

має розподіл Ст'юдента з  $(n-1)$  ступенями свободи. Пояснення:  $\frac{\bar{X}_n - \nu}{\frac{\sigma}{\sqrt{n}}}$  -

нормований нормальний розподіл.  $\sqrt{\frac{nS_n^2}{\sigma^2}}$  - розподіл  $\chi^2$  з  $(n-1)$  ступенями свободи, вони незалежні.

По таблиці для розподілу Ст'юдента знаходимо таке число  $t_{\alpha, n-1}$ , для якого :

$$P\left(\left|\frac{\overline{X}_n - \nu}{\frac{\sigma}{\sqrt{n}}}\right| \leq t_{\alpha, n-1}\right) = 1 - \lambda$$

$$P\left(\overline{X}_n - t_{\alpha, n-1} \frac{S_n}{\sqrt{n-1}} \leq \nu \leq \overline{X}_n + t_{\alpha, n-1} \frac{S_n}{\sqrt{n-1}}\right) = 1 - \alpha$$

Наслідки:

- 1) Якщо  $n$  – невелике, то  $t_{\lambda, n-1} > t_{\lambda}$ , тобто довірчий інтервал більший
- 2) Заздалегідь нам невідома кількість випробувань, щоб отримати оцінку заданої точності

3 задача – інтервальна оцінка для дисперсії

Розглянемо випадковку величину  $\frac{nS_n^2}{\sigma^2}$  - має розподіл  $\chi^2$  з  $(n-1)$  ступенями свободи. За таблицями для розподілу  $\chi^2$  знаходимо числа  $\chi_1^2$  та  $\chi_2^2$ , такі що

$$P(\chi^2 \leq \chi_1^2) = \frac{\alpha}{2}$$

$$P(\chi^2 \geq \chi_2^2) = \frac{\alpha}{2}$$

Таким чином отримаємо:

$$\begin{array}{|c|c|c|} \hline + & + & + \\ \hline \end{array} \quad P(\chi_1^2 \leq \frac{nS_n^2}{\sigma^2} \leq \chi_2^2) = 1 - \alpha \quad \text{або}$$

$$P\left(\frac{nS_n^2}{\chi_2^2} \leq \sigma^2 \leq \frac{nS_n^2}{\chi_1^2}\right) = 1 - \alpha$$

Тоді з ймовірністю  $1 - \alpha$  виконується:

$$\frac{nS_n^2}{\chi_2^2} \leq \sigma^2 \leq \frac{nS_n^2}{\chi_1^2}$$

#### 4.11 Інтервальні оцінка для дискретної випадкової величини

Нехай дискретна випадкова величина задається таблицею:

$$X \begin{pmatrix} x_1 & \dots & x_s \\ p_1 & \dots & p_s \end{pmatrix}$$

Проведено  $n$  випробувань. Заздалегідь відомо, що  $P_i = \frac{n_i}{n}$ , де  $n_i$ —кількість випробувань, в кожному з яких настало  $x_i$ .  $n$  повторів одного випробування будемо вважати як  $n$  віртуальних незалежних копій цього одного випробування. З кожним  $j$ -им віртуальним випробуванням зв'яжемо

випадкову величину  $X_j^i = \begin{cases} 1, \text{ якщо настало } x_i \\ 0, \text{ якщо не настало } x_i \end{cases}$

$$\text{Тоді } X_j^i = \begin{pmatrix} 1 & 0 \\ p_i & 1-p_i \end{pmatrix}.$$

$$S_n^i = \sum_{j=1}^n X_j^i$$

- біноміальний розподіл. При цьому:

$$M S_n^i = n p_i$$

$$D S_n^i = n p_i (1 - p_i)$$

Розглянемо випадкову величину  $\frac{S_n^i}{n}$  – реалізація котрій є число  $\frac{n_i}{n}$ :

$$M \frac{S_n^i}{n} = p_i$$

$$D \frac{S_n^i}{n} = \frac{p_i(1-p_i)}{n}$$

Тоді можна вважати, що випадкова величина  $\frac{S_n^i}{n} - p_i$  має наближено нормальний розподіл, а величина:

$$\frac{\frac{S_n^i}{n} - p_i}{\sqrt{\frac{p_i(1-p_i)}{n}}}$$

має нормований нормальний розподіл.

Тоді по таблицям значень оберненої функції Лапласа знаходимо:

$$P\left(\left|\frac{\frac{S_n^i - P_i}{n}}{\sqrt{\frac{P_i(1-P_i)}{n}}}\right| < t_\alpha\right) = 1 - \alpha$$

$$\frac{S_n^i}{n} - t_\alpha \sqrt{\frac{P_i(1-P_i)}{n}} \leq P_i \leq \frac{S_n^i}{n} + t_\alpha \sqrt{\frac{P_i(1-P_i)}{n}}$$

$$\frac{n_i}{n} - t_\alpha \sqrt{\frac{P_i(1-P_i)}{n}} \leq P_i \leq \frac{n_i}{n} + t_\alpha \sqrt{\frac{P_i(1-P_i)}{n}}$$

Знайдемо верхн. оцінку  $P_i(1-P_i)$ :

$$\max_{P_i} P_i(1-P_i) = \frac{1}{4}$$

*Прим.: можемо заздалегідь знайти необхідну кількість випробувань для знаходження  $P_i$  з точністю.*

#### 4.12 Прості, нульові та складні гіпотези

Не маємо ніякої інформації про розподіл неперервної величини  $X$ .  
Необхідно знайти оцінку її функції щільності:

Методологія розв'язку:

- 1) Проводимо  $n$  випробувань та отримуємо вибірку об'єму  $n$
- 2) Будуємо гістограму
- 3) Порівнюючи гістограми зі всіма відомими графіками функції щільності висуваємо гіпотезу щодо невідомої функції щільності.

Гіпотези можуть бути простими та складними:

- Проста однозначно задає невідому функцію  $f(x)$

Зустрічаються випадки, коли  $n$  – достатньо велике число. В цьому випадку гіпотеза виявляє функцію щільності з точністю до числових значень параметрів. Вважаючи гіпотезу точною, використовуємо метод

найбільшої правдоподібності та оцінки числових параметрів (отримані методом найбільшої правдоподібності) вважаємо достатньо точними, оскільки кількість випробувань – велике число. Ці оцінки підставляємо в функцію щільності.

- Складна гіпотеза задає невідому функцію  $f(x)$  щільності з точністю до числових значень невідомих параметрів

Виникає тоді, коли кількість випробувань мале.

- Нульова гіпотеза  $H_0$

Виникає в наступному випадку: маємо  $n$  незалежних величин з невідомими параметрами (або параметром). Висувається гіпотеза, що один з параметрів цих випадкових величин має одне й те саме значення (наприклад значення математичні сподівання випадкових величин рівні).

#### 4.12.1 Методологія перевірки простої (нульової) гіпотези

Над випадковою величиною (випадковими величинами при  $H_0$ ) проводяться випробування та фіксуються результати, котрі вважаються результатом одного випробування над віртуальними копіями випробування (стандартний прийом мат. статистики). Кількість віртуальних випадкових величин дорівнює кількості проведених випробувань. Вводиться числова скалярна функція введених віртуальних випадкових величин, котра називається критерієм перевірки гіпотези.

Обов'язкова умова, що накладається на критерій перевірки гіпотези: проста гіпотеза або нульова гіпотеза однозначно задає функцію щільності критерію перевірки гіпотези (або функцію розподілу). Оскільки критерій перевірки – скалярна випадкова величина (числова скалярна функція випадкових аргументів), то її простір елементарних подій – числова вісь або відрізок числової вісі.

Простір елементарних подій критерію перевірки  $\Omega$  задається як:

$$D_{\text{доп}} \cup D_{\text{кр}}$$

де  $D_{\text{доп}}$  – область допустимих значень, а  $D_{\text{кр}}$  – область критичних значень. При цьому:

$$D_{\text{доп}} \cap D_{\text{кр}} = \{\emptyset\} = V$$

Критичну область беруть із умови:

$$P(K(\cdot) \in D_{\text{кр}}) = \alpha \leq 0.05$$

Де  $K(\cdot)$  - критерій перевірки гіпотези, та  $\cdot$  - випадковий аргумент.

Допустима область – область протилежна критичної області.

Підставив в критерій перевірки гіпотези замість значень віртуальних випадкових аргументів значення, які настали в ході випробування над віртуальними випадковими величинами, отримуємо реалізацію критерія перевірки гіпотези.

Якщо реалізація критерію належить критичної області, то вважається, що гіпотеза не є вірною (оскільки настала подія з малою ймовірністю наставання).

Нехай реалізація критерію належить допустимій області. Тоді продовжуємо дослідження (оскільки гіпотеза може бути невірною, але настали такі події, при яких реалізація критерію попала в допустиму область), тобто статистичні данні не суперечать гіпотезі. Дослідження продовжуються за допомогою інших критеріїв. Статистично доведено, що найбільш «жорстким» критерієм перевірки є критерій  $\chi^2$ . В процесі перевірки гіпотези можна зробити помилки першого и другого роду. Помилка першого роду – відкинути гіпотезу, що є вірною. Ймовірність помилки першого роду дорівнює  $\alpha$ . Помилка другого роду – прийняти гіпотезу, що є невірною. В загальному випадку знайти ймовірність помилки другого роду неможливо. Можна сформулювати методологічний прийом, що зменшує ймовірність помилки другого роду: таким чином обирати критичну область, щоб, чим більше була невірна гіпотеза, яка

перевіряється, тим більше була насправді ймовірність попадання критерію в критичну область.

#### 4.12.2 Класифікація критичних областей

1. Область великих значень критерію задається умовою

$$P(K(\cdot) \geq t_{\alpha}^1) = \alpha$$

використовується тоді, коли відомо, якщо гіпотеза хибна то у критерію є тенденція збільшення своїх значень

2. Область малих значень критерію

$$P(K(\cdot) \leq t_{\alpha}^2) = \alpha$$

3. Коли немає жодної інформації

А)  $P(|K(\cdot)| \geq t_{\alpha}^{31}) = \alpha$ , використовується тоді коли область симетрична відносно осі ординат

$$\text{Б) } P\left(\begin{matrix} K(\cdot) \leq t_{\alpha}^{32} \\ K(\cdot) \geq t_{\alpha}^{33} \end{matrix}\right) = \alpha$$

#### *Приклад*

Маємо дві незалежні випадкові величини  $X$  і  $Y$ , розподілені нормально:

$$X : n(x, \nu_x, \sigma_x)$$

$$Y : n(y, \nu_y, \sigma_y)$$

Дисперсії – відомі. Мат.сподівання – невідомі. Провели  $n_1, n_2$  випробувань:

$$X : x_1 \dots x_{n_1}$$

$$Y : y_1 \dots y_{n_2}$$

По результатам статистичних даних перевірити гіпотезу, що  $\nu_x = \nu_y$

Вважаємо, що  $n_1 + n_2$ -числа є результатом складного випробування над  $n_1 + n_2$ -вимірною випадковою величиною,  $Y_1 \dots Y_{n_1}, Y_2 \dots Y_{n_2}$

Будуємо критерій перевірки простої гіпотези



$$K(X_1 \dots X_{n_1}, Y_1 \dots Y_{n_2}) = \frac{\bar{X}_{n_1} - \bar{Y}_{n_2}}{\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}}$$

$$\bar{X}_{n_1} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i$$

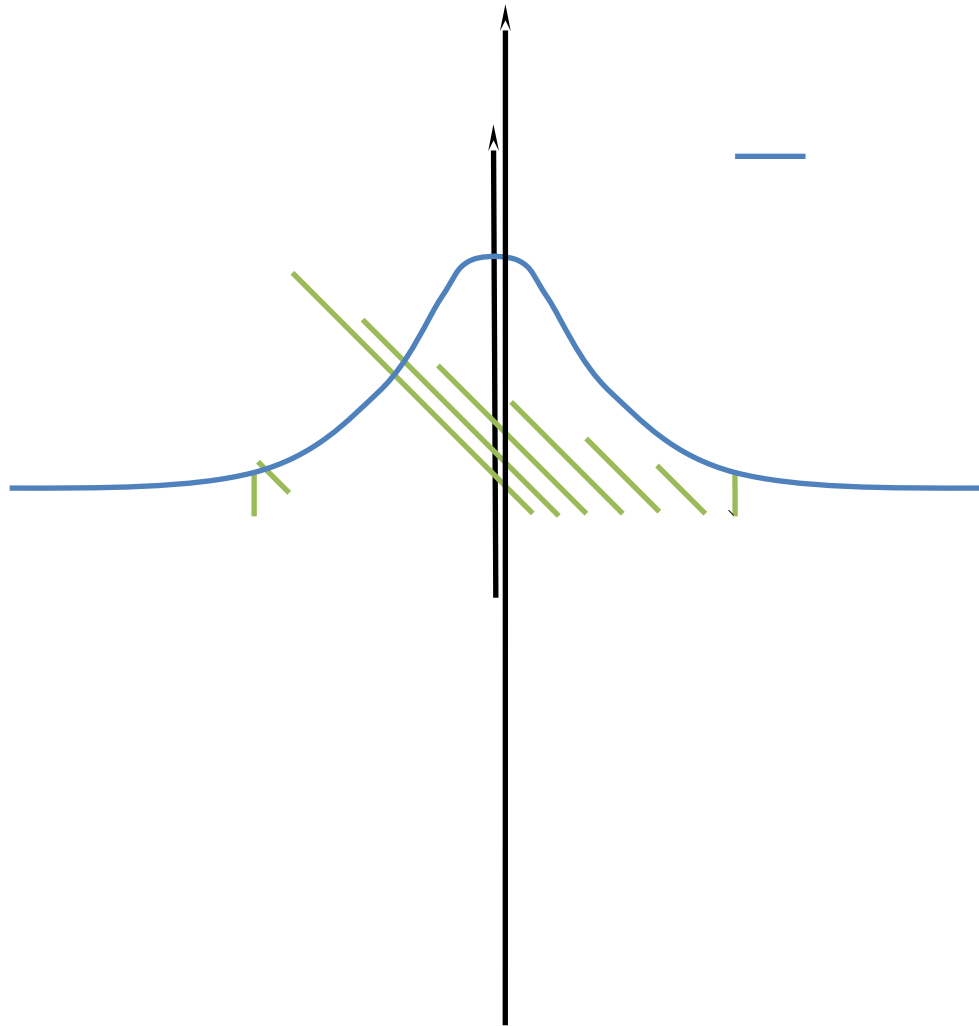
$$\bar{Y}_{n_2} = \frac{1}{n_2} \sum_{i=1}^{n_2} Y_i$$

В знаменнику корінь із дисперсії чисельника

В чисельнику випадкова величина, мат.сподівання якої дорівнює  $\nu_x - \nu_y$

Якщо  $H_0$  вірна, то  $\nu_x = \nu_y$  і критерій перевірки гіпотези має нормований нормальний розподіл.

Вибір критичної області для нормованого нормального розподілу:



$$v_x > v_y$$

$$\Phi_0(t_\alpha) = \frac{1}{2} - \alpha$$

$$t_\alpha^1 = \Phi_0^{-1}\left(\frac{1}{2} - \alpha\right)$$

Якщо  $v_y > v_x$ , то

$$t_\alpha^2 = -t_\alpha^1$$

Якщо немає жодної інформації, вибирається 3 критична область

$$t_\alpha = t_\alpha^1$$

$$\Phi_0(t_\alpha) = \frac{1 - \alpha}{2}$$

$$t_\alpha = \Phi_0^{-1}\left(\frac{1 - \alpha}{2}\right)$$

#### 4.12.3 Критерій знаків

Вирішуємо наступну задачу: є два незалежних випадкових об'єкти, що діють під впливом спільного фактору. Необхідно перевірити гіпотезу про те, що зміна значення фактора впливає на випадкові події однаково.

Статистична постановка задачі:

Маємо дві  $n$ -мірні випадкові величини:  $X_1 \dots X_N$  та  $X_1^1 \dots X_N^1$ , всі компоненти яких незалежні.

Перевіримо нульову гіпотезу про те, що розподіл  $X_i$  та  $X_i^1$  при довільному  $i$  однаковий.

Результати експериментів:

Для спрощення: випадкові величини  $X_i, X_i^1$  неперервні

Розглянемо випадкову величину  $K_N(+)$  і її реалізацію  $k_N(+)$

$k_n(+)$  - кількість додатніх різниць  $x_i - x_i^1, i = \overline{1, N}$ .

Якщо гіпотеза вірна, то випадкова величина  $K_n(+)$  має біноміальний розподіл. Це дозволяє сформулювати основну вадку критерію знаків: з того,

що додатня і від'ємна різниця має ймовірність настання  $\frac{1}{2}$  не впливає,

що випадкові величини  $X_i$  та  $X_i^1$  мають однаковий розподіл. Майже кожен критерій перевірки гіпотез перевіряє не саму гіпотезу а її наслідок. Критична область перевірки критерію гіпотези є:

$$P\left(\left|K_N(+)-\frac{N}{2}\right| \geq A\right) = \alpha \leq 0,05$$

точне значення числа  $A$  знаходиться в розв'язку алгебраїчного рівняння

$$\sum_{m=0}^{\frac{N}{2}-A} C_N^m \left(\frac{1}{2}\right)^N + \sum_{m=\frac{N}{2}+A}^N C_N^m \left(\frac{1}{2}\right)^N = \alpha$$

При достатньо великому N

$$\frac{K_N(+)-\frac{N}{2}}{\sqrt{\frac{N}{4}}} \text{ має нормований нормальний розподіл (т.Муавра - Лапласа)}$$

Тоді за наближенням для теореми Муавра-Лапласа:

$$P\left(\left|\frac{k_N(+)-\frac{N}{2}}{\sqrt{\frac{N}{4}}}\right| > t_\alpha\right) = \alpha$$

$$A = t_\alpha \sqrt{\frac{N}{4}}$$

$$\Phi_0(t_\alpha) = \frac{1-\alpha}{2}$$

$$t_\alpha = \Phi_0^{-1}\left(\frac{1-\alpha}{2}\right)$$

#### 4.12.4 Критерій $\chi^2$ для перевірки простої гіпотези

Є Найбільш «жорстким» і найбільш використовуваним критерієм для перевірки простої або складної гіпотези.

Постановка задачі: маємо випадкову величину  $X$ , вибірку  $x_1, \dots, x_n$  об'єму  $n$ , необхідно перевірити просту гіпотезу про функцію щільності  $f(x)$ .

Алгоритм перевірки гіпотези:

1) Оскільки при побудуванні гістограми розглядається відрізок  $(x_1; x_n)$ , то обирається число  $a \in [8; 12]$ , а цій відрізок б'ється на частини, в кожную з яких попало  $a$  членів варіаційного ряду (члені вибірки)

2) Знаходимо ймовірність того, що внаслідок випробування:

$$P(X \in \Delta_i) = P_i = \int_{y(i)}^{y(i+1)} f(x) dx$$

де  $y(i)$  – лівий кінець  $i$ -ого відрізка.

Оскільки  $X$  – неперервна, то усі члени ряду – різні, та  $x(1) = y(1)$ .

$$\int_{y(i)}^{y(i+1)} f(x) dx \approx f\left(y_i + \frac{1}{2}(y(i+1) - y(i))\right)(y(i+1) - y(i))$$

Теоретичний недолік критерію  $\chi^2$ : критерій перевіряє не вірна або невірна гіпотеза, а вірність числа  $P_i$  (тобто перевіряє не гіпотезу, а її наслідок).

С кожним  $\Delta_i$  відрізком зв'язуємо випадкову величину  $M_i$  – кількість випробувань з  $n$ , в кожному з яких випадкова величина прийняла значення в  $\Delta_i$  відрізка. Реалізація довільного  $M_i$  на кожному відрізку дорівнює  $a$ , оскільки так побудовані відрізки  $\Delta_i$ . Якщо  $n$  – кількість випробувань достатньо більше 20, то  $M_i$  мають нормальне розподілення (за теоремою Муавра – Лапласа).

Якщо гіпотеза є вірною (тобто числа  $P_i$  є вірними), то критерій  $\chi^2$  має  $r-1$  ступінь свободи.

$$\chi^2 = \sum_{i=1}^r \left( \frac{M_i - n p_i}{\sqrt{n p_i}} \right)^2$$

Реалізація цього критерію дорівнює:

$$\sum \left( \frac{a - n p_i}{\sqrt{n p_i}} \right)^2$$

тобто  $M_i$  для кожного  $i$  прийняла значення  $a$ .

В якості критичної області обирається область великих значень критерію, оскільки якщо  $P_i$  невірні, то оскільки у кожного доданка в сумі у чисельнику  $n^2$ , а у знаменнику –  $n$ , то чим більше невірні  $P_i$ , тим більше має тенденцію к збільшенню своїх значень критерій. Тобто в таблиці  $\chi^2$  з  $r-1$  ступеню свободи знайдеться таке  $\alpha$ , що

$$P(\chi^2 \geq \chi_\alpha^2) = \alpha$$

#### 4.12.5 Критерій $\chi^2$ для перевірки складної гіпотези

Постановка задачі: нехай гіпотеза задає функцію щільності  $f(x, \theta_1, \dots, \theta_c)$

Вважаючи складну гіпотезу вірною за методом найбільшої правдоподібності знаходимо оцінку невідомих параметрів  $\hat{\theta}_i$  і підставляємо знайдені оцінки в аналітичну функцію щільності, а далі повністю повторяємо для як для простої гіпотези. Остаточний вираз критерію має вигляд:

$$\chi^2 = \sum_{i=1}^r \left( \frac{M_i - n \hat{p}_i}{\sqrt{n \hat{p}_i}} \right)^2$$

Фішер строго довів, що якщо гіпотеза вірна, то критерій  $\chi^2$  має  $n - c - 1$  ступенів свободи.

#### 4.12.6 Критерій $n\omega^2$

$F(x)$  — функція неперервно розподіленої випадкової величини. Насправді гіпотезу можна висунути тільки по  $f(x)$ , просто по ній будується функція  $F(x)$  і  $F(x)$  перевіряється по  $n\omega^2$ .

$$W_n(x) = \begin{cases} 0, & x < x(1) \\ \frac{k}{n}, & x(k) < x \leq x(i+1) \\ 1, & x > x(n) \end{cases}$$

Примітка.  $X_i$  неперервна, тому числа  $x(i)$  різні.

Відомо, що при  $n \rightarrow \infty$ ,  $\omega_n(x) \rightarrow F(x)$  (нам відомо, що при  $\forall x W_n(x) \rightarrow F(x)$  з посиленого закону великих чисел).

$$\omega^2 = \int_{-\infty}^{\infty} (F(x) - \omega_n(x))^2 dF(x) = \int_{-\infty}^{x(i)} F(x)^2 dF(x) + \int_{x(i)}^{x(i+1)} \left[ \frac{k}{n} - F(x) \right]^2 dF(x) + \int_{x(n)}^{\infty} (1 - F(x))^2 dF(x)$$

Якщо гіпотеза вірна, то інтеграл прямує до 0.

Розглянемо випадкову величину  $n\omega^2$

Теоретично доведено, що при  $n \rightarrow \infty$  випадкові величини  $n\omega^2$  мають границю: випадкову величину з граничним розподілом, що є табульовним.

На практиці цей граничний розподіл використовують при  $n \geq 40$ .

Критична область – область великих значень критерія. Для заданого  $\alpha$  знаходиться за таблицями.

#### 4.12.7 Критерій грубих помилок вимірювання

Постановка задачі: є досліджувана випадкова величина. Відомо, що вона розподілена нормально. Провели  $n$  випробувань, отримали вибірку об'єму  $n$ .

Перевіряємо гіпотезу: найбільший член вибірки  $x_{(n)}$  є результатом випробування над  $X$ , а не результатом грубої помилки вимірювання.

$X, n(x, \nu, \sigma)$

$x_1 \dots x_n$

$x_{(1)} \dots x_{(n)}$

Задача: перевірити, що  $x_1, \dots, x_n$  дійсно є результатами випробувань над  $X$ .

Розв'язок: будуємо критичну область за допомогою числа  $x^*$ , тобто  $x_{(n)}$  вважаємо грубою помилкою вимірювання. Якщо максимальний член вибірки виявився більше  $A$ , то настала малоймовірна подія:

$$P(x_{(n)} \geq x) = \alpha \leq 0,05$$

$$P(x_{(n)} < x) = 1 - \alpha$$

Для того, щоб максимальний член вибірки був  $< A$ , необхідно і достатньо, щоб кожен член вибірки був  $< A$ , а результатом  $n$  повторів

випробувань вважаємо результат композиції  $n$  незалежних віртуальних випробувань, тоді

$$(N(x, \nu, \sigma))^n = 1 - \alpha$$

$$(N(\frac{x - \nu}{\sigma}, 0, 1))^n = 1 - \alpha$$

Примітка: записуємо функцію розподілу, як інтеграл від функції

щільності, в якому виконана заміна змінних  $\frac{x - \nu}{\sigma} = z$ .

$$\left( \frac{1}{2} + \Phi_0 \left( \frac{x - \nu}{\sigma} \right) \right)^n = 1 - \alpha$$

$$\frac{1}{2} + \Phi_0 \left( \frac{x - \nu}{\sigma} \right) = (1 - \alpha)^{\frac{1}{n}}$$

$$\frac{x - \nu}{\sigma} = \Phi_0^{-1} \left( (1 - \alpha)^{\frac{1}{n}} - \frac{1}{2} \right)$$

$$x = \Phi_0^{-1} \left( (1 - \alpha)^{\frac{1}{n}} - \frac{1}{2} \right) \sigma + \nu$$

#### 4.13 Багатовимірний дисперсійний аналіз

Багатовимірний дисперсійний аналіз досліджує вплив множини факторів на досліджуваний випадковий об'єкт.

*Приклади:*

1) *однофакторний*. Є різні фізичні величини, що вимірюються одним приладом, але різними операторами. Треба дослідити чи впливає 1 фактор – вимірювання різними операторами на систематичну похибку результату вимірювань.

2) *двофакторний*. Фізична величина вимірюється різними приладами і різними операторами. Дослідити чи впливають два фактори : вимірювання різними приладами та різними операторами на похибку вимірювань.



#### 4.14 Однофакторний дисперсійний аналіз

Постановка задачі.

Маємо  $m$  незалежних випадкових величин  $X_1 \dots X_n$  розподілених по нормальному закону з однаковою, але невідомою дисперсією і невідомими математичними сподіваннями.

Зафіксована вибірка

$$\begin{bmatrix} x_{11} & \dots & x_{n1} \\ \vdots & \ddots & \vdots \\ x_{1n} & \dots & x_{nn} \end{bmatrix}$$

Маємо фізичну величину. Кожен з  $m$  операторів робить по  $n$  вимірювань. Треба перевірити нульову гіпотезу.

$$\sigma_{x_1}^2 = \sigma_{x_2}^2 = \dots = \sigma_{x_n}^2 = \sigma^2$$

(в усіх випадкових величин однакові математичні сподівання). Фактор не впливає на математичне сподівання випадкової величини.

*Розв'язок:*

$$\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij} - \text{середнє арифметичне по } i\text{-ій виборці.}$$

$$\bar{x} = \frac{1}{n \cdot m} \sum_{i=1}^m \sum_{j=1}^n x_{ij} - \text{загальне середнє арифметичне}$$

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m \bar{x}_i - \text{загальне середнє арифметичне є середнім арифметичним}$$

середнього арифметичного по виборкам

Має місце наступна тотожність однофакторного дисперсійного аналізу

$$\sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x})^2 = \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_i + \bar{x}_i - \bar{x})^2 = \sum_{i=1}^m \sum_{j=1}^n (\bar{x}_i - \bar{x})^2 + \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2$$

$$Q_1 = \sum_{i=1}^m \sum_{j=1}^n (\bar{x}_i - \bar{x})^2$$

$$Q_2 = \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2$$

$$Q = Q_1 + Q_2$$

Самим довести рівність.

Якщо перевіряєма гіпотеза є вірною, то  $\bar{x}_i$  і  $\bar{x}$  оцінюють одне і те саме. Якщо перевіряєма гіпотеза є вірною, то  $Q_1$  є невеликою і тим меншою чим більше кількість випробувань.

Показати, що величина  $Q_2$  від перевіряємої гіпотези не залежить (підставити замість  $x_{ij}$  числа  $x_{ij} + c_i$  і показати, що  $Q_2$  не змінюється.

Методологія перевірки нульової гіпотези ґрунтується на порівнянні  $Q_1$  та  $Q_2$ .

Використовуємо стандартний прийом і вважаємо, що  $m \cdot n x_{ij}$  є результатом 1 віртуального складного випробування над  $m \cdot n$  вимірною випадковою величиною  $\{X_{ij}\}_{i=1, \dots, m; j=1, \dots, n}$ , де  $X_{ij}$  мають нормальний розподіл і незалежні.

Використовуючи тотожність

$$\sum_{i=1}^m \sum_{j=1}^n (X_{ij})^2 = \sum_{i=1}^m \sum_{j=1}^n (\bar{X}_i - \bar{X})^2 + \sum_{i=1}^m \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2$$

$$\bar{X}_i = \frac{1}{n} \sum_{j=1}^n X_{ij}$$

$$\bar{X} = \frac{1}{n \cdot m} \sum_{i=1}^m \sum_{j=1}^n X_{ij}$$

Якщо вірна гіпотеза, то випадкова величина  $\frac{Q}{\delta^2}$  має розподіл  $\chi^2$  з  $m \cdot n - 1$  ступенями свободи (повна аналогія з  $\frac{n \cdot S^2}{\sigma^2}$ )

$\frac{Q_1}{\delta^2}$  розподілена по закону  $\chi^2$  з  $n - 1$  ступенями свободи.

$\frac{Q_2}{\delta^2}$  розподілена по закону  $\chi^2$  з  $m \cdot (n - 1)$  ступенями свободи.

Для фіксованого  $\sum_{j=1}^m \frac{(X_{ij} - \bar{X}_i)^2}{\sigma^2}$  має розподіл  $\chi^2$  з  $n-1$  ступенями

свободи  $\Rightarrow \sum_{i=1}^m \left( \frac{\sum_{j=1}^n (X_{ij} - \bar{X}_i)^2}{\sigma^2} \right)$  має розподіл  $\chi^2$  з  $m \cdot (n-1)$  ступенями свободи, бо

для різних  $i$  випадкові величини незалежні.

Наслідок (властивість характеристичної функції розподілу  $\chi^2$ : сума незалежних випадкових величин, розподілених по розподілу  $\chi^2$ , має розподіл  $\chi^2$  з сумарною кількістю ступеней свободи).

Покажемо що випадкові величини  $Q_1$  та  $Q_2$  незалежні. Дійсно, оскільки  $X_{ij}$  розподілені однаково якщо вірна гіпотеза, що перевіряється, то  $\bar{X}_1 \dots \bar{X}_n$  незалежні між собою.

Аналогічно незалежні  $n \cdot S_1^2 \dots n \cdot S_n^2$ , а також  $\bar{X}_i$  та  $S_j^2$ ,  $\forall i \neq j$ .

$n \cdot S_n^2$  та  $\bar{X}_n$  також незалежні (дивись розподіл  $\frac{ns^2}{\sigma^2}$ ).

$Q_1$  та  $Q_2$  є числовими скалярними функціями незалежних аргументів, тому є незалежними (див. відповідну лекцію курсу).

Критерій перевірки нульової гіпотези є наступним

$$F = \frac{\frac{Q_1}{\sigma^2}}{\frac{m \cdot (n-1)}{\frac{Q_2}{\sigma^2}}}$$

Це розподіл Фішера з  $m-1, m \cdot (n-1)$  ступенями свободи.

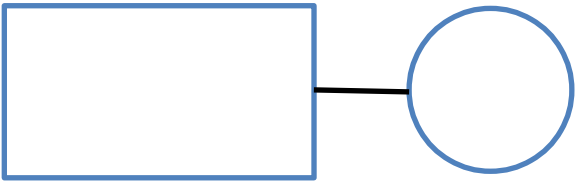
Критична область- область великих значень критерію.

#### 4.15 Регресійний аналіз

Задача регресії:

Постановка задачі:

Маємо фізично існуючий об'єкт, вхід котрого – детермінований та скалярний. Об'єкт лінійно перетворює вхід на вихід за законом:  $\alpha + \beta x$ . На виході якої



аддитивно додається випадкова величина  $\Delta$ , у якої  $M \Delta = 0$ ,  $\sigma^2 < \infty$  та розподілена нормально.

На виході вимірюється не  $\alpha + \beta x$ , а  $\alpha + \beta x + \Delta$ . На вхід подається  $n$

чисел і кожному числу відповідно міряється вихід. Маємо вибірку

$x_1$	$y_1$
$\vdots$	$\vdots$
$x_n$	$y_n$

По статистичним даним необхідно знайти оцінки  $\alpha$  та  $\beta$ .

В загальному вигляді:  $y = \alpha + \beta x + \Delta$ .

У результаті випробувань:  $y_i = \alpha + \beta x_i + \delta_i$ ,  $i = 1, \dots, n$ ,  $\delta$  – реалізація  $\Delta$ .

Використовуючи загальний прийом математичної статистики можна вважати, що числа  $y_i$  та  $\delta_i$  – реалізації случайних величин  $Y_i$  та  $\Delta_i$ , де  $Y_i = \alpha + \beta x_i + \Delta_i$  а  $\Delta_i$  – незалежні віртуальні копії  $\Delta$ .  $\Delta_i$  має нормальне розподілення з нульовим математичним сподіванням та дисперсією  $\sigma^2$ , невідомою нам. Звідси  $Y_i$  – незалежні та розподілені нормально за законом  $n(x, \alpha + \beta x_i, \sigma)$ .

Розглянемо  $n$ -мірну випадкову величину  $\Delta_1 \dots \Delta_n$  та знайдемо її функцію щільності.

$\Delta$  має нормальне розподілення  $\frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}}$ .

$n$ -мірна функція щільності рівняється добутку одновимірних, звідки

$$f(u_1, \dots, u_n) = \left( \frac{1}{2\pi\sigma} \right)^n e^{-\frac{\sum_{i=1}^n u_i^2}{2\sigma^2}}$$

Оцінки невідомих  $\alpha$ ,  $\beta$  та  $\sigma^2$  знаходимо за методом найбільшої правдоподібності. Для цього по  $n$ -мірної функції будуємо функцію найбільшої правдоподібності  $L(\alpha, \beta, \sigma^2)$ . Для цього в  $n$ -мірну функцію щільності необхідно підставити реалізації  $\Delta_i$ ,  $i=1, n$ . Числа  $\delta_i$  нам невідомі, але відомо, що  $\delta_i = y_i - \alpha - \beta x_i$ . Тоді:

$$L(\alpha, \beta, \sigma^2) = \left( \frac{1}{2\pi\sigma^2} \right)^{\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (y_i - \alpha - \beta x_i)^2}{2\sigma^2}}$$

В силу структури функції найбільшої правдоподібності  $\hat{\alpha}$  та  $\hat{\beta}$  знаходяться з умови:

$$\min_{\alpha, \beta} \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2$$

Візьмемо похідні по  $\alpha$  та  $\beta$ :

$$\alpha: \sum_{i=1}^n (y_i - \alpha - \beta x_i) = 0$$

$$\beta: \sum_{i=1}^n (y_i - \alpha - \beta x_i) x_i = 0$$

Оскільки вхід об'єкту детерміновано і на вхід можна подавати довільні значення, то для спрощення вважаємо, що:

$$\sum_{i=1}^n x_i = 0$$

(тобто маємо можливість реалізувати активний експеримент)

Звідси слідує, що:

$$\hat{\alpha} = \frac{\sum_{i=1}^n y_i}{n}$$

та

$$\hat{\beta} = \frac{\sum_{i=1}^n y_i x_i}{\sum_{i=1}^n x_i^2}$$

Якщо підставити в функцію правдоподібності значення  $\hat{\alpha}$  та  $\hat{\beta}$ , то отримуємо оцінку  $\sigma^2$ :

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{\alpha} - \hat{\beta} x_i)^2$$

Таким чином вибіркві характеристики для  $\hat{\alpha}$  та  $\hat{\beta}$  мають вигляд:

$$\hat{\alpha} = \frac{\sum_{i=1}^n Y_i}{n}$$

$$\hat{\beta} = \frac{\sum_{i=1}^n Y_i x_i}{\sum_{i=1}^n x_i^2}$$

Побудуємо інтервальні оцінки для коефіцієнтів  $\alpha$  та  $\beta$ :

$$Y_i = \alpha + \beta x_i + \Delta_i$$

$$Y_i x_i = (\alpha + \beta x_i + \Delta_i) x_i$$

Перше рівняння сумуємо по  $n$  та ділимо на  $n$ . Отримуємо:

$$\frac{\sum_{i=1}^n Y_i}{n} = \sum_{i=1}^n \frac{(\alpha + \beta x_i + \Delta_i)}{n}$$

$$\hat{\alpha} - \alpha = \frac{\sum_{i=1}^n \Delta_i}{n}$$

Аналогічно:

$$\frac{\sum_{i=1}^n Y_i x_i}{\sum_{i=1}^n x_i^2} = \frac{\sum_{i=1}^n (\alpha + \beta x_i + \Delta_i) x_i}{\sum_{i=1}^n x_i^2}$$

$$\hat{\beta} - \beta = \frac{\sum_{i=1}^n \Delta_i x_i}{\sum_{i=1}^n x_i^2}$$

Якщо дисперсія була би відома, то можна було побудувати довірчий інтервал для  $\alpha$  та  $\beta$  (зробить самостійно)

Але дисперсія невідома, тому продовжуємо дослідження.

$$\eta_1 = \frac{\sum_{i=1}^n \Delta_i}{\sqrt{n}}$$

$$\eta_2 = \frac{\sum_{j=1}^n \Delta_j x_j}{\sqrt{\sum_{i=1}^n x_i^2}}$$

Можна показати, що  $\eta_1$  та  $\eta_2$  мають від матриці:

$$\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} \cdots \\ \cdots \end{pmatrix} \begin{pmatrix} \Delta_i \\ \vdots \\ \Delta_n \end{pmatrix}, \text{rang}(\cdots) = 2,$$

$\Delta_1, \dots, \Delta_n$  мають  $n$ -мірне нормальний розподіл ( $\Delta_i$  – незалежні та нормовані). Звідси випадкові величини  $\eta_1$  та  $\eta_2$  мають одномірний

нормальний розподіл, а  $\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}$  – двомірний нормальний розподіл. Очевідно,

що:

$$M \eta_1 = M \eta_2 = 0$$

$$D \eta_1 = D \eta_2 = \sigma^2$$

Покажемо, що  $\text{cov}(\eta_1, \eta_2) = 0$ , а це означає, що  $\eta_1$  та  $\eta_2$  – незалежні, розподілені нормально з параметрами  $n(x, 0, \sigma^2)$ .

Дійсно  $\text{cov}(\eta_1, \eta_2) = M(\eta_1, \eta_2)$ , так як  $M \eta_1 \cdot M \eta_2 = 0$

$$M(\eta_1 \cdot \eta_2) = \frac{1}{\sqrt{n \sum_{i=1}^n x_i^2}} M \sum_{i=1}^n \sum_{j=1}^n (\Delta_i \Delta_j x_j) = 0$$

оскільки:

$$M \sum_{i=1}^n \sum_{j=1}^n (\Delta_i \Delta_j x_j) = \sigma^2 \sum_{i=1}^n x_i = 0$$

Прим.: математичне сподівання добутку двох незалежних величин рівняється добутку математичних сподівань, та  $M \Delta_i^2 = \sigma^2$  так як  $M \Delta_i = 0$ .

Розглянемо  $n\hat{\sigma}^2 = \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta}x_i)^2$ ,  $\hat{\alpha} = \frac{\sum_{i=1}^n Y_i}{n}$  та  $\hat{\beta} = \frac{\sum_{i=1}^n Y_i x_i}{\sum_{i=1}^n x_i^2}$ .

Можна показати, що  $\frac{n\hat{\sigma}^2}{\sigma^2}$  має розподіл  $\chi^2$  з  $(n-2)$  степенями свободи та випадкові величини  $\frac{n\hat{\sigma}^2}{\sigma^2}$ ,  $\eta_1$  та  $\eta_2$  — попарно незалежні.

Таким чином випадкові величини  $\frac{\sqrt{n}(\hat{\alpha} - \alpha)}{\sigma}$  та  $\frac{(\hat{\beta} - \beta)\sqrt{\sum_{i=1}^n x_i^2}}{\sigma}$  розподілені за нормованим нормальним законом та є незалежними, а  $\frac{n\hat{\sigma}^2}{\sigma^2}$  — за розподілом  $\chi^2$  з  $(n-2)$  степенями свободи та попарно ці величини також незалежні.

Довірчий інтервал для коефіцієнту  $\alpha$ .

Побудуємо випадкову величину:

$$\frac{\frac{\sqrt{n}(\hat{\alpha} - \alpha)\sqrt{n-2}}{\sigma}}{\sqrt{\frac{n\hat{\sigma}^2}{\sigma^2}}} = \frac{(\hat{\alpha} - \alpha)\sqrt{n-2}}{\hat{\sigma}}$$

Ця величина має розподіл Стюдента  $(n-2)$  степенями свободи. Невідома дисперсія скоротилася.

За таблицями розподілу Стюдента з  $(n-2)$  степенями свободи знаходимо таке  $t_{\alpha, n-2}$ , що:

$$P\left(\left|\frac{(\hat{\alpha} - \alpha)\sqrt{n-2}}{\hat{\sigma}}\right| \leq t_{\alpha, n-2}\right) = 1 - \alpha$$

$$\alpha \leq 0.05$$

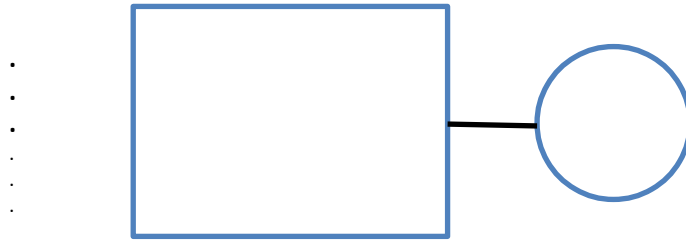
Можна переміщаючи знаменник вправо та розкривши модуль побудувати довірчий інтервал для  $\alpha$ . Побудуємо для  $\beta$ :



$$\frac{(\hat{\beta}-\beta)\sqrt{\sum_{i=1}^n x_i^2(n-2)}}{\frac{\sigma}{\sqrt{\frac{n\hat{\sigma}^2}{\sigma^2}}}} = \frac{(\hat{\beta}-\beta)\sqrt{\sum_{i=1}^n x_i^2(n-2)}}{\sqrt{n}\hat{\sigma}}$$

$$P\left(\left|\frac{(\hat{\beta}-\beta)\sqrt{\sum_{i=1}^n x_i^2(n-2)}}{\sqrt{n}\hat{\sigma}}\right| \leq t_{\alpha, n-2}\right) = 1-\alpha$$

#### 4.15.1 Узагальнена задача регресії



##### Постановка задачі

На «вхід» об'єкта подається детермінований вектор  $\dot{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ . Об'єкт

перетворює «вхід» на «вихід» по закону:

$$y = \sum_{j=0}^r \theta_j f_j(\dot{x}),$$

де  $f_j(\dot{x})$  – відомі функції.  $\theta_j$  невідомі коефіцієнти.

##### Наприклад

1) Двовимірний поліном другого ступеня:

$$\sum_{j=0}^r \theta_j f_j(\dot{x}) = \theta_0 + \theta_1 x_1 + \theta_2 x_1 x_2 + \theta_3 x_1^2$$

$$2) \sum_{j=0}^r \theta_j f_j(x) = \theta_0(x_1 + x_2) + \theta_1(x_1 + x_1^2) + \theta_2 x_1 x_2 + \theta_3 x_1^2$$

На виході аддитивно додається випадкова величина  $\Delta$ , така що  $M\Delta=0$ ,  $\sigma^2 < \infty$  та розподілено нормально за законом  $n(x, 0, \sigma)$ .

Таким чином, отримали випадкову величину

$$Y(x) = \sum_{j=0}^r \theta_j f_j(x) + \Delta$$

За результатами випробувань  $x_1 y_1 \dots x_n y_n$  необхідно знайти оцінку параметрів  $\theta_1 \dots \theta_r$ .

Випадкова величина  $\Delta$  має нормальний розподіл  $n(x, 0, \sigma)$ , а з цього випливає що оцінка параметрів зводиться до знаходження

$$\min_{\theta_1 \dots \theta_r} \sum_{i=0}^n \left( y_i - \sum_{j=0}^r \theta_j f_j(x_i) \right)^2,$$

якщо використовувати метод найбільшої правдоподібності.

*Примітка*

Це можна довести, повторюючи попередній результат, підставивши в  $n$ -вимірну функцію щільності  $\delta_i = y_i - \sum_{j=0}^r \theta_j f_j(x_i)$ .

Якщо  $\Delta$  розподілено довільно (не нормально), метод найбільшої правдоподібності приводить до зовсім іншої задачі максимізації.

#### 4.15.2 Метод найменших квадратів для скалярного детермінованого

##### ВХОДУ

*Постановка задачі*

Лінія регресії має вигляд:

$$Y(x) = \sum_{j=0}^r \theta_j x^j + \varepsilon,$$

де  $\varepsilon$  – випадкова величина, що має довільний розподіл і її математичне сподівання  $M\varepsilon=0$ , дисперсія  $\sigma^2 < \infty$  – невідома,  $\theta_j, j=0, \dots, r$  невідомі коефіцієнти.

Після проведення випробувань отримано  $x_i, y_i, i=1, \dots, n$ , де  $y_i = \sum_{j=0}^r \theta_j x_i^j + \delta_i$ ,  $\theta_j$  – реалізація випадкової величини  $\epsilon$ . Виходячи з методу найменших квадратів знаходження  $\theta_j, j=0, \dots, r$  зводиться до наступної задачі:

$$\min_{\theta_1, \dots, \theta_r} \sum_{i=0}^n \left( y_i - \sum_{j=0}^r \theta_j x_i^j \right)^2 \quad (1)$$

*Якісне пояснення суті методу найменших квадратів*

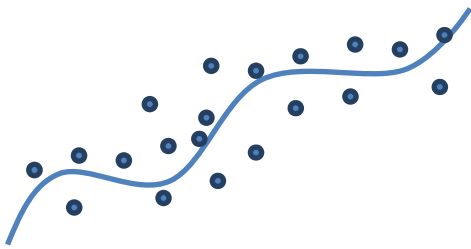


рис.1

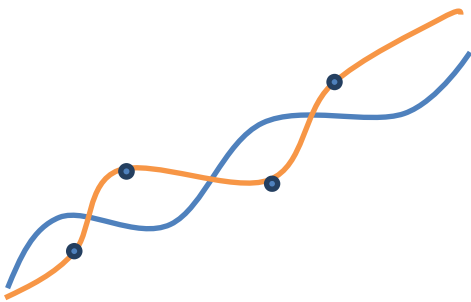


рис.2

Якщо кількість експериментів прямує до нескінченності (достатньо велика, рис.1), то формула (1) повинна дати точні значення коефіцієнтів лінії регресії. В цьому випадку

$$\min_{\theta_1, \dots, \theta_r} \sum_{i=0}^n \left( y_i - \sum_{j=0}^r \theta_j x_i^j \right)^2 = \sum_{i=1}^n \delta_i^2$$

Якби оцінки коефіцієнтів, отримані методом найменших квадратів, не давали точні значення, (1) була б гарантовано більшою, бо на неї накладалася б помилка від того, що лінія регресії знайдена неточно. А ми знаходимо ці оцінки мінімізуючи суму.

Якщо випробувань небагато, можна отримати погані чи навіть абсурдні результати (рис.2). Дійсно, якщо кількість випробувань менша, ніж кількість коефіцієнтів, то (1) дорівнює нулю і лінія регресії, побудована по знайденим коефіцієнтам, проходить через всі точки експерименту.

*Рішення*

$$y_i = \sum_{j=0}^r \theta_j x_i^j + \delta_i, i=1, \dots, n \quad (2)$$

$$Y_i = \sum_{j=0}^r \theta_j x_i^j + \Delta_i, i=1, \dots, n \quad (3)$$

Введемо позначення:

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \theta = \begin{pmatrix} \theta_0 \\ \vdots \\ \theta_r \end{pmatrix}, \delta = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_n \end{pmatrix}, \Delta = \begin{pmatrix} \Delta_1 \\ \vdots \\ \Delta_n \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^r \\ 1 & x_2 & x_2^2 & \dots & x_2^r \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^r \end{pmatrix}$$

Тоді можна переписати у векторній формі:

$$(2): y = A\theta + \delta$$

$$(3): Y = A\theta + \Delta$$

$$(1): \min_{\theta} (y - A\theta)^T (y - A\theta)$$

Використовуючи  $(AB)^T = B^T A^T$ :

$$(y - A\theta)^T (y - A\theta) = y^T y - y^T A\theta - \theta^T A^T y + \theta^T A^T A\theta = y^T y - 2\theta^T A^T y + \theta^T A^T A\theta$$

*Примітка:*  $y^T A \theta = \theta^T A^T y$  – бо це числа.

Тоді (1):  $\min_{\theta} (y^T y - 2 y^T A \theta + \theta^T A^T A \theta)$ .

Так як оптимальний розв'язок – це внутрішня точка, то необхідно вирішити систему:

$$\frac{\partial (y^T y - 2 y^T A \theta + \theta^T A^T A \theta)}{\partial \theta_j} = 0, j=0, \dots, r.$$

Запишемо цю систему рівнянь у векторній формі, використовуючи мнемонічне правило взяття похідної по вектору:

$$-2 A^T y + 2 A^T A \theta = 0.$$

Система рівнянь:

$$A^T y = 2 A^T A \theta.$$

Звідки:

$$\hat{\theta} = (A^T A)^{-1} A^T y$$

$$\hat{\theta} = (A^T A)^{-1} A^T Y$$

В другому рівнянні  $\hat{\theta}$  – випадковий вектор а в першому  $\hat{\theta}$  – його реалізація.

#### 4.15.3 Властивості оцінок, отриманих методом найменших квадратів

##### **1) Незміщеність**

Доведення

$$M \Delta = \begin{pmatrix} M \Delta_1 \\ \vdots \\ M \Delta_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Тоді з  $Y = A\theta + \Delta$ ;  $MY = A\theta$ . Знаходимо математичне сподівання оцінок:

$$M \hat{\theta} = M \left( (A^T A)^{-1} A^T Y \right) = (A^T A)^{-1} A^T M Y = (A^T A)^{-1} A^T A \theta = \theta$$

**2) Теорема Маркова.** Оцінки, отримані методом найменших квадратів, є ефективні в класі всіх лінійних незміщених оцінок.

Доведення

Означення. Оцінка  $\theta^{\hat{i}} = \begin{pmatrix} \theta_0^{\hat{i}} \\ \vdots \\ \theta_r^{\hat{i}} \end{pmatrix}$  є лінійною, якщо

$$\theta^{\hat{i}} = UY,$$

$U$  – довільня числова матриця, розмір  $U - (r+1) \times n$

З умови незміщеності:

$$M \theta^{\hat{i}} = \theta$$

чи

$$M(UY) = UMY = UA\theta = \theta$$

Тоді:

$$UA = E, (4)$$

де  $E$  – одинична матриця розміру  $(r+1) \times (r+1)$ .

Примітка.  $U$  і  $A$  – прямокутні матриці.

При умові виконання (4) теорема Маркова означає:

$$D \hat{\theta}_j \leq D \theta_j^{\hat{i}}, j=0, r.$$

Доведення

$Y$  випадкового вектора  $Y$  компоненти – незалежні випадкові величини  
 $Y$  з дисперсією  $DY_i, i=1, n$ . Тоді коваріаційна матриця вектора  $Y$ :

$$\text{cov} Y = (\text{cov}(Y_i Y_j))^n = \sigma^2 E$$

Знайдемо  $\text{cov} \theta^{\hat{i}}$  - довільня незміщена лінійна векторна оцінка.

$$\text{cov} \theta^{\hat{i}} = \text{cov} UY = U \sigma^2 E U^T = \sigma^2 U U^T$$

Примітка.  $\text{cov} AX = A \text{cov} X A^T$

Знайдемо:

$$\text{cov} \hat{\theta} = \text{cov} \left( (A^T A)^{-1} A^T Y \right) = (A^T A)^{-1} A^T \sigma^2 E A (A^T A)^{-1} = \sigma^2 (A^T A)^{-1} A^T A (A^T A)^{-1} = \sigma^2 (A^T A)^{-1}$$

Примітка. матриця  $(A^T A)^{-1}$  є симетричною, тому збігається з  $\left( (A^T A)^{-1} \right)^T$ .

Для спрощення, далі позначимо :  $(A^T A)^{-1} = K$ .

Має місце наступна тотожність:

$$U U^T = K + (U - K A^T)(U - K A^T)^T$$

### Доведення

$$K + (U - K A^T)(U - K A^T)^T = K + U U^T - U A K - K A^T U^T + K A^T A K = \left| \begin{array}{l} U A = E \\ A^T U^T = E \end{array} \right| = K + U U^T - K - K + K =$$

Що й необхідно було довести.

Помножимо вираз на  $\sigma^2$ :

$$\sigma^2 U U^T = \sigma^2 K + \sigma^2 (U - K A^T)(U - K A^T)^T$$

або

$$\text{cov } \theta^{\hat{}} = \text{cov } \hat{\theta} + \left| \begin{array}{l} \text{Матриця з додатними елементами} \\ \text{на головній діагоналі} \end{array} \right|$$

Дійсно, на головній діагоналі матриці  $\text{cov } \theta^{\hat{}}$  стоять  $D \theta^{\hat{}}_j$ , на головній діагоналі  $\text{cov } \hat{\theta}$  стоять  $D \hat{\theta}_j$ . Тобто:  $D \hat{\theta}_j \leq D \theta^{\hat{}}_j, j=0', r$ , що й необхідно було довести.

$$3) R^T R = Y^T Y - Y^T \hat{Y}, \text{ де } R = Y - A \hat{\theta}, \hat{Y} = A \hat{\theta}.$$

### Доведення

$$\hat{\theta} = (A^T A)^{-1} A^T Y$$

$\hat{Y}$  – випадковий вектор;  $\hat{\theta}$  – випадковий вектор.

$$\hat{Y} = A \hat{\theta}$$

$$\hat{y} = A \hat{\theta}$$

$$\hat{y}_i = \hat{\theta}_0 + \hat{\theta}_1 x_i + \dots + \hat{\theta}_r x_i^r, i=1', n$$

$$\hat{y}(x) = \hat{\theta}_0 + \hat{\theta}_1 x + \dots + \hat{\theta}_r x^r - \text{ця функція і є оцінкою регресії.}$$

$\hat{y}_i$  – значення оціночної регресії, коли вхід -  $x_i$ .

$$R = Y - \hat{Y}$$

$$r = y - \hat{y}$$

$$r^T r = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$R^T R$  - зветься **залишковою сумою квадратів**.

ІІІ реалізація  $r^T r$  має властивість: якщо це число велике, то ми отримали погані оцінки чи невірно вибрали ступінь полінома лінії регресії (менше, ніж вона є насправді).

Примітка. Вважається, що ступінь полінома  $r$  – відома. Насправді це не так. Але дослідник обирає з якісних міркувань  $r$ , дивлячись на графік.

І навпаки, якщо сума квадратів  $r^T r$  мала, значить взяли занадто високий ступінь.

$$R^T R = (Y - A\hat{\theta})^T (Y - A\hat{\theta}) = Y^T Y - Y^T A\hat{\theta} - \hat{\theta}^T A^T Y + \hat{\theta}^T A^T A\hat{\theta} = \underbrace{[Y^T A\hat{\theta} = \hat{\theta}^T A^T Y]}_{\text{скаляр}} = Y^T Y - 2Y^T A\hat{\theta} + \hat{\theta}^T A^T A\hat{\theta}$$

#### 4.15.4 Метод найменших квадратів на ортогональних поліномах

Насправді дослідник не знає ступені поліному  $r$ . Єдине рішення цього – брати  $r$  із запасом. Вважають, що метод найменших квадратів дає оцінку 0 при тех. членах поліному, котрих насправді нема. Але у цьому випадку матриця  $A^T A$  Астановиться практично виродженою та знаходження її оберненої фактично веде до помилок.

Ця проблема вирішується за допомогою використання ортогональних поліномів.

Приклад: нехай лінія регресії має вигляд (модель 1):

$$Y(x) = \theta_0 + \theta_1 x + \varepsilon$$

Будемо шукати лінію регресії в наступному вигляді (модель 2):

$$Y(x) = \phi_0 + \phi_1(x - \bar{x}) + \varepsilon$$

Де  $\phi_0$  та  $\phi_1$  – невідомі коефіцієнти.

Зв'язок між коефіцієнтами наступний:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$\theta_0 = \phi_0 - \phi_1 \bar{x}$$

$$\theta_1 = \phi_1$$

Для моделі (2) використовуємо метод найменших квадратів. В цьому випадку можна показати, що:



$$\begin{pmatrix} \phi_0 \\ \phi_1 \end{pmatrix} = (A^T A)^{-1} A^T Y$$

де

$$A = \begin{pmatrix} 1 & x_1 - \bar{x} \\ 1 & x_2 - \bar{x} \\ \dots & \dots \\ 1 & x_n - \bar{x} \end{pmatrix}$$

Дійсно, рівняння має вигляд:

$$y_i = \phi_0 + \phi_1(x_i - \bar{x}) + \sigma_i$$

Необхідно мінімізувати:

$$\min_{\phi_0, \phi_1} (y_i - (\theta_0 + \theta_1 x_i))^2$$

Побудуємо матрицю  $A^T$ :

$$A^T = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 - \bar{x} & x_2 - \bar{x} & \dots & x_n - \bar{x} \end{pmatrix}$$

Тоді:

$$A^T A = \begin{pmatrix} n & 0 \\ 0 & \sum_{i=1}^n (x_i - \bar{x})^2 \end{pmatrix}$$

Звідси:

$$(A^T A)^{-1} = \begin{pmatrix} \frac{1}{n} & 0 \\ 0 & \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2} \end{pmatrix}$$

Загальний випадок:

$$Y(x) = \theta_0 + \theta_1 x + \dots + \theta_r x^r + \varepsilon$$

Знайдемо модель у вигляді моделі 3:

$$Y(x) = \theta_0 G_0(x) + \theta_1 G_1(x) + \dots + \theta_r G_r(x) + \varepsilon$$

де

$$G_j(x) = g_{j,0} + g_{j,1}x + \dots + g_{j,j-1}x^{j-1} + x^j$$

Коефіцієнти  $g_{ij}$  повинні задовольняти наступній системі рівнянь:

$$\sum_{l=1}^n G_l(x_l) G_p(x_l) = 0, \forall l < p$$

яка зветься системою Грама-Шмідта.

Можна показати, що якщо використовувати метод найменших квадратів для моделі 3, то отримуємо формулу  $\hat{\theta} = (A^T A)^{-1} A^T Y$ , де матриця  $A$  має вигляд:

$$A = \begin{pmatrix} G_0(x_1) & \cdots & G_r(x_1) \\ \vdots & \ddots & \vdots \\ G_0(x_n) & \cdots & G_r(x_n) \end{pmatrix}$$

$$y_i = \theta_0 G_0(x_i) + \theta_1 G_1(x_i) + \dots + \theta_r G_r(x_i) + \delta_i$$

Тоді  $(A^T A)$  має вигляд:

$$A^T A = \begin{pmatrix} \sum_{i=1}^n G_0^2(x_i) & 0 & \dots & 0 \\ 0 & \sum_{i=1}^n G_1^2(x_i) & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & \sum_{i=1}^n G_r^2(x_i) \end{pmatrix}$$

$$(A^T A)^{-1} = \begin{pmatrix} \frac{1}{\sum_{i=1}^n G_0^2(x_i)} & 0 & \dots & 0 \\ 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{1}{\sum_{i=1}^n G_r^2(x_i)} \end{pmatrix}$$

Складність рішення залишилась: необхідно вміти вирішувати систему рівнянь Грама-Шмідта. Цю складність зняв американський вчений Форестер у 1951 році.

#### 4.15.5 Метод найменших квадратів на нормованих ортогональних поліномах

Модель задається у вигляду:

$$Y(x) = \omega_0 Q_0(x) + \omega_1 Q_1(x) + \dots + \omega_r Q_r(x) + \varepsilon$$

$$y_i = \omega_0 Q_0(x_i) + \dots + \omega_r Q_r(x_i) + \delta_i, \quad i=1, n$$

$$Q_j(x) = q_{j,0} + q_{j,1}x + q_{j,2}x^2 + \dots + q_{j,j}x^j, \quad j=0, r$$

*Прим.: кожен  $j$ -ий нормований поліном задається  $j+1$  коефіцієнтом.*

Визначення: Нормованим ортогональним поліномом зветься ортогональний поліном, що задовольняє:

$$\sum_{i=1}^n Q_j^2(x_i) = 1, \quad j=0, r$$

В цьому випадку матриця  $(A^T A)^{-1}$  – одинична, та  $\hat{\omega} = A^T Y$  – вектор оцінок, де

$$A^T = \begin{pmatrix} Q_0(x_1) & \dots & Q_0(x_n) \\ \vdots & \ddots & \vdots \\ Q_r(x_1) & \dots & Q_r(x_n) \end{pmatrix}$$

Форстер вивів рекурентну формулу для  $\alpha Q_j(x)$ :

$$\lambda Q_j(x) = x Q_{j-1}(x) - \alpha Q_{j-1}(x) - \beta Q_{j-2}(x)$$

де

$$\alpha = \sum_{i=1}^n x_i Q_{j-1}^2(x_i)$$

$$\beta = \sum_{i=1}^n x_i Q_{j-1}(x_i) Q_{j-2}(x_i)$$

$$\lambda = \sqrt{\sum_{i=1}^n (x_i Q_{j-1}(x_i) - \alpha Q_{j-1}(x_i) - \beta Q_{j-2}(x_i))^2}$$

Форестер довів, що такий вибір коефіцієнтів  $\alpha$  та  $\beta$  робить поліноми  $Q_j$  – ортогональними (тобто для них виконується система рівнянь Грама-Шмідта), а такий вибір  $\lambda$  – нормованими.

$$\text{Очевидно, що } Q_0(x) = \frac{1}{\sqrt{n}}$$

Розглянемо властивості коефіцієнтів, лінія регресії, що отримані методом найменших квадратів на нормованих ортогональних поліномах.

$$\hat{\omega} = A^T y$$

$$Y_i = \sum_{j=0}^r \omega_j Q_j(x_i) + \Delta_i, \Delta_i \text{ незалежні копії } \varepsilon$$

$$\widehat{W} = A^T Y$$

$$D Y_i = \sigma^2, Y_i - \text{незалежні випадкові величини}$$

Властивості:

$$1) W_j = \sum_{i=1}^n Q_j(x_i) Y_i, j=0, \dots, r$$

2)  $M \widehat{W}_j = w_j, j=0, \dots, r$  (повторюється виведення першої властивості загального методу найменших квадратів із заміною  $\theta$  на  $\omega$ )

$$3) D \widehat{W}_j = \sigma^2, \text{ (оскільки } D \left( \sum_{i=1}^n Q_j(x_i) Y_i \right) = \sigma^2 \sum_{i=1}^n Q_j^2(x_i) = \sigma^2)$$

Точність оцінки не залежить від кількості випробувань.

4) Побудова лінії регресії на нормованих ортогональних поліномах незручна, тому перейдемо до звичайної:  
 $\hat{\theta}_j = \widehat{W}_r q_{r,j} + \widehat{W}_{r-1} q_{r-1,j} + \dots + \widehat{W}_j q_{j,j}$  Прим.: аналогічно прикладу де  $\theta = \Phi_0 - \Phi_1 \dot{x}$ ;  
 $\Phi_1 = \theta_1$ .

$$5) \text{cov}(\widehat{W}_l, \widehat{W}_p) = 0, \forall l \neq p. \quad \text{Покажемо, що } M(\widehat{W}_l \widehat{W}_p) = M(\widehat{W}_l) M(\widehat{W}_p):$$

$$\widehat{W}_l = \sum_{i=1}^n Q_l(x_i) Y_i, \widehat{W}_p = \sum_{j=1}^n Q_p(x_j) Y_j, M \widehat{W}_l = \sum_{i=1}^n M Y_i Q_l(x_i), M \widehat{W}_p = \sum_{j=1}^n M Y_j Q_p(x_j)$$

$$M \widehat{W}_l M \widehat{W}_p = \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n M Y_i M Y_j Q_l(x_i) Q_p(x_j) + \sum_{i=1}^n (M Y_i)^2 Q_l(x_i) Q_p(x_i)$$

$$M(\widehat{W}_l \widehat{W}_p) = \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n M Y_i M Y_j Q_l(x_i) Q_p(x_j) + \sum_{i=1}^n M(Y_i)^2 Q_l(x_i) Q_p(x_i) \text{ Прим.:}$$

$$M(Y_i Y_j) = M(Y_i) M(Y_j), \quad \text{бо } Y_i \text{ та } Y_j - \text{незалежні.}$$

Тоді:

$$\sum_{i=1}^n Q_l(x_i) Q_p(x_i) = 0 \text{ з умови ортогональності.}$$

$$\widehat{\omega}_{r+1} = \sum_{i=1}^n Q_{r+1}(x_i) y_i \widehat{W}_{r+1} = \sum_{i=1}^n Q_{r+1}(x_i) Y_i$$
$$R^T R = Y^T Y - Y^T \hat{Y} = Y^T Y - Y^T A \hat{W} = Y^T Y - \hat{W}^T \hat{W} = \sum_{i=1}^n Y_i^2 - \sum_{j=0}^r \hat{W}_j^2 \text{Прим.:} \quad \hat{Y} = A \hat{W} \quad \text{та}$$

Спосіб доведення: повторити цей результат для загального метода найменших квадратів.

Якщо  $\varepsilon$  – нормальне, то звідси слідує наступний результат:

$$\widehat{W} = A^T Y$$

246

наслідком відповідної теореми компоненти вектору  $\widehat{W}$  мають  $r+1$ -мірний нормальний розподіл.

$\widehat{W}_j$  розподілено на законом  $n(x, \omega_j, \sigma)$ ,  $j=0, \dots, r$ .

Із того, що виконується властивість 5 слідує, що  $\widehat{W}_j$ —незалежні.

Нехай ступінь поліному  $r$  відома. Тоді модель запишемо у вигляді:

$$2) Y_i = \omega_0 Q_0(x_i) + \omega_1 Q_1(x_i) + \dots + \omega_{n-1} Q_{n-1}(x_i) + \Delta_i,$$

де  $\omega_{r+1} = \dots = \omega_{n-1} = 0$ .

Методом найменших квадратів на нормованих ортогональних поліномах знаходимо оцінки невідомим коефіцієнтам на моделі (1) та моделі (2). В моделі (1) оцінка  $r+1$  коефіцієнту  $\widehat{\omega}_j$ ,  $j=0, \dots, r$  дає випадкові величини, розподілені за нормальним законом  $n(x, \omega_j, \sigma)$ . У моделі (2) оцінки  $\widehat{W}_j$ ,  $j=0, \dots, r$ , незалежні, розподілені нормально, причому для перших  $r+1$  — по закону  $n(x, \omega_j, \sigma)$  та збігаються з оцінкою коефіцієнтів моделі (1). Тобто  $\widehat{\omega}_j$ ,  $j=0, \dots, r$  за першою та другою моделлю тотожні. А для  $j=r+1, \dots, n-1$  — розподіл  $n(x, 0, \sigma)$ . Для другої моделі розглянемо залишкову суму квадратів:

$$R_{n-1}^T R_{n-1} = \sum_{i=1}^n Y_i^2 - \sum_{j=0}^{n-1} \widehat{W}_j^2 = 0$$

оскільки  $n$  випробувань та  $n$  невідомих, то лінія регресії пройде по усім точками. Це означає, що

$$R_r^T R_r = \sum_{i=1}^n Y_i^2 - \sum_{j=1}^r \widehat{W}_j^2 = \sum_{j=r+1}^{n-1} \widehat{W}_j^2$$

Тобто залишкова сума квадратів для моделі (1) рівна:

$$\sum_{j=r+1}^{n-1} \widehat{W}_j^2$$

Тоді випадкова величина  $R^T R$  не залежить від перших  $r+1$   $\widehat{W}_j^2$ ,  $j=0, \dots, r$ , через які вона виражається.

Тоді випадкова величина  $\frac{R_r^T R_r}{\sigma^2}$  має розподіл  $\chi^2$  з  $(n-r-1)$  степенями свобод.

Алгоритм знаходження  $r$  – ступеню регресії:

(1) Візуально аналізувати експериментальні данні. Знаходимо таке  $r$  із запасом, що гарантовано  $\omega_{r+1}=\dots=\omega_{n-1}=0$  та знаходимо методом найменших квадратів оцінки коефіцієнтам  $\widehat{\omega}_j, j=0',r$

(2) Послідовно крок за кроком для  $\forall j=0',r$  перевіряємо гіпотезу, що  $M(\widehat{W}_j)=0$ , тобто коефіцієнт рівен нулю.

Для перевірки гіпотези будуємо критерій:

$$F_j = \frac{\frac{\widehat{W}_j^2}{\sigma^2} (n-r-1)}{\frac{R_r^T R_r}{\sigma^w}}, j=0',r$$

Якщо гіпотеза, що перевіряється, вірна, то критерій має розподіл Фішера з  $1, (n-r-1)$  ступенями свободи. В цьому випадку в чисельнику розподіл  $\chi^2$  з однією ступінню свободи. Критична область – область великих значень (знаменник не залежить, а чисельник тим більше, чим більше невірна гіпотеза)

Проведемо  $r+1$  перевірок гіпотези та відкидаємо усі  $\omega_i$ , для яких критерій попав у допустиму область.

## 5.

### 5.4 Марківські ланцюги

Проводиться система випробувань. В кожному випробуванні може настати одна з  $A_1 \dots A_k$  несумісних подій. Відомо, що у випробуванні з

номером  $S$  настала подія  $A_i$ . Тоді умовна ймовірність того, що в випробуванні з номером  $S+n$  настане подія  $A_j$  не залежить від того, які події настали у випробуваннях з номерами менше  $S$ .

Іншими словами, випробування називається ланцюгом Маркова, якщо умовна ймовірність настання будь-якої події залежить лише від останньої інформації про події, що настали в випробуваннях з меншими номерами.

Ланцюг Маркова називається однорідним, якщо ця умовна ймовірність залежить від параметрів  $n, A_i, A_j$ , і не залежить від  $S$ .

Для будь-якого  $S$  ця ймовірність одна і та сама.

Для однорідного ланцюга Маркова вводяться такі ймовірності:

$$\pi_1 = (P_{ij})_1^k$$

$P_{ij}$  – ймовірність того, що в наступному випробуванні настане подія  $A_j$ , якщо в даному випробуванні настала подія  $A_i$ .

Властивості елементів матриці  $\pi_1$ :

$$1) 0 \leq P_{ij} \leq 1$$

2) Сума елементів будь-якого рядка буде дорівнювати 1 (за 3 аксіомою Колмагорова)

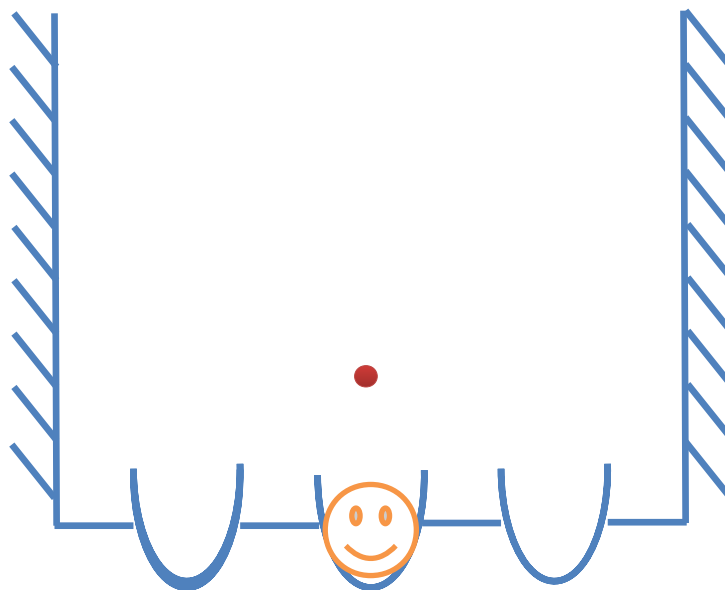
Примітка:  $\bigcup_i A_i = U, A_i \cap A_j = V, \forall i \neq j$

2 означення ланцюга Маркова

Маємо систему, яка в будь-який момент часу знаходиться в будь-якому з  $k$ -станів. В фіксований момент часу  $t_1, t_2 \dots t_s \dots t_{s+n}$  система миттєво, випадковим чином змінює свої стани. Між цими моментами часу система стани не змінює, при чому виконується умовна ймовірність того, що в момент часу  $t_{s+n}$  система перейде миттєво в стан  $A_j$ , якщо в момент часу  $t_s$  вона перейшла в  $A_i$ , не залежить від того, в які стани переходила система до моменту часу  $t_s$ .



Приклад:



В одній із лунок знаходиться кулька.

В моменти часу  $t_1, t_2, \dots$  діє сила, яка з ймовірністю  $p$  переправляє кульку вправо та з ймовірністю  $1-p$  вліво.

Якщо кулька знаходиться біля стінки, то з ймовірністю 1 повертається в сусідню лунку.

Запишемо матрицю переходів  $\pi_1$ :

$$\pi_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ P & 0 & 1-P & 0 & 0 \\ 0 & P & 0 & 1-P & 0 \\ 0 & 0 & P & 0 & 1-P \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Розглядаємо однорідний ланцюг Маркова.

Означення:

Якщо в випробуванні з номером  $S$  настала подія  $A_i$ , то з ймовірністю  $P_{ij}(n)$  настане подія  $A_j$  при будь-якому значенні  $S$ .

$\pi_k = (P_{ij}(n))_1^k$  - квадратна матриця  $k$ -го порядку.

Виразимо матрицю  $\pi_k$  через матрицю  $\pi_1$ , тобто  $P_{ij}(n)$  через  $P_{ij}$ . Будемо використовувати наступну формулу:  $P(AB / C) = P(A / C)P(B / AC)$ .

Формула не виводиться а лише обґрунтовується через частість.

Фіксуємо випробування з номером  $S+m$ , де  $m < n$ , і робимо композицію  $n+1$  випробування:

композиція випробувань з номерами  $S, S+1, \dots, S+n$

Елементарною подією цього композиційного випробування буде символ довжини  $n+1$ , а на місці кожного символу стоятиме подія  $A_1 \dots A_k$ .

Подія  $C$  - складна подія в просторі елементарних подій композиційного випробування, що складається з усіх елементарних подій композиційного випробування, у яких перший символ дорівнює  $A_i$ .

Подія  $A$  складається з усіх елементів, в яких на  $m+1$  місці стоїть подія  $A_r$  (тобто в випробуванні з номером  $S+m$  настане подія  $A_r$ , в інших – що завгодно).

Подія  $B$  – подія, яка складається з усіх елементарних подій, в кожній з яких на останній позиції стоїть  $A_j$

Тоді

$$P(AB / C) = P_{ir}(m)P_{rj}(n-m)$$

$$\left( \begin{array}{l} P(B / AC) = P(B / A) - \\ \text{маємо однорідний} \\ \text{Марківський ланцюг} \end{array} \right)$$

$$P_{ij}(n) = \sum_{r=1}^k P_{ir}(m) P_{rj}(n-m)$$

(За третьою аксіомою Колмогорова)

В матричній формі  $\pi_n = \pi_m \cdot \pi_{n-m}$ . Виразимо  $\pi_n$  через  $\pi_1$ :  $\pi_n = (\pi_1)^n$  (доводиться мат. індукцією).

#### 5.4.1 Гранична теорема для однорідних ланцюгів Маркова

Нехай існує таке натуральне число  $S$ , що всі елементи матриці переходу  $\pi_s$  більші 0

$$\forall_{i,j=1,n} P_{i,j}(S) > 0$$

Тоді існують числа  $P_1 \dots P_n$

$$\forall i \exists \lim_{n \rightarrow \infty} P_{ij}(n) = P_j, \quad \sum_{j=1}^k P_j = 1$$

Якісне трактування теореми:

Якщо всі елементи матриці  $\pi_s$  строго  $> 0$ , то в якому б стані не знаходилася система у теперішній час, при необмеженій кількості кроків переходу із стану в стан існує границя ймовірності перебування системи в стані  $A_j$  і ця ймовірність не залежить в якому стані знаходиться система в теперішній час.

Доведення теореми

Базується на наступних 3-х фактах.

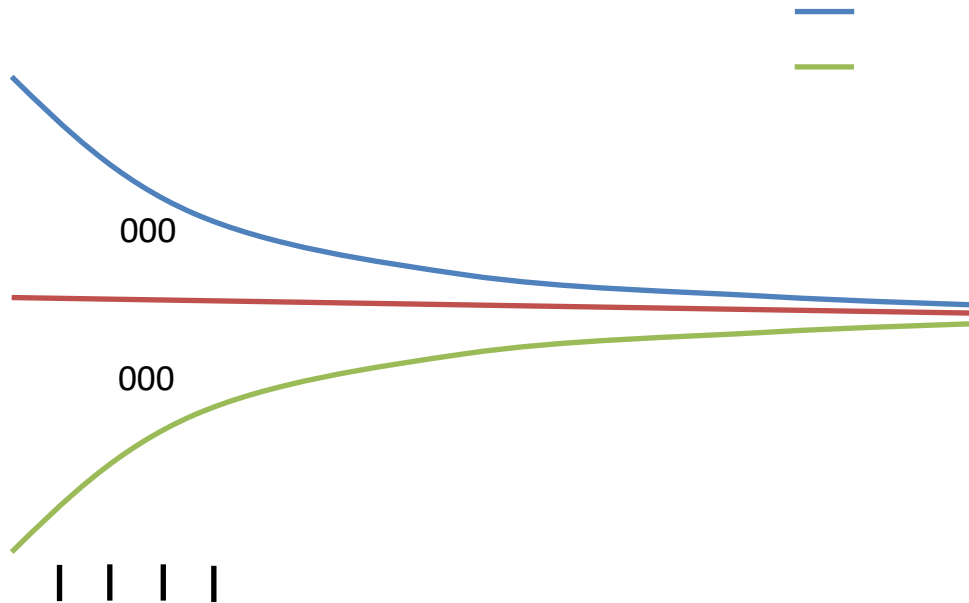
$$1) \min_i P_{ij}(n) \geq \min_i P_{ij}(n-1)$$

тобто зі збільшенням  $n$  мінімальна ймовірність не спадає.

$$\text{З цього випливає: } \exists \lim_{n \rightarrow \infty} \min_i P_{ij}(n) = \underline{P}_j$$

$$2) \max_i P_{ij}(n) \leq \max_i P_{ij}(n-1) \Rightarrow \exists \lim_{n \rightarrow \infty} \max_i P_{ij}(n) = \bar{P}_j$$

$$3) \exists \lim_{n \rightarrow \infty} \max_{i,l} |P_{ij}(n) - P_{lj}(n)| = 0 \Rightarrow \underline{P}_j = \bar{P}_j = P_j$$



Послідовне доведення цих фактів.

1) Використовуємо формулу  $\pi_n = \pi_m \cdot \pi_{n-m}$

Нехай  $m=1$ . Тоді 
$$P_{ij}(n) = \sum_{r=1}^k P_{ir} P_{rj}(n-1) \geq \min_r P_{rj}(n) \sum_{r=1}^k P_{ir} = \min_r P_{rj}(n-1)$$
 -

впливає нерівність першого факту.

2) Аналогічно доводиться 2 факт.

$$3) P_{ij}(n) - P_{lj}(n) = \sum_{r=1}^k P_{ir}(S) P_{rj}(n-S) - \sum_{r=1}^k P_{lr} P_{rj}(n-S) =$$

(Примітка: використана ф-ла 
$$P_{ij}(n) = \sum_{r=1}^k P_{ir}(m) P_{rj}(n-m)$$
 для  $m = S, \forall i, S,$

$P_{ij}(S) > 0$  (з умови теореми))

$$= \sum_{r=S}^k (P_{ir}(S) - P_{lr}(S)) P_{rj}(n-S) =$$

Далі отримуємо окремі результати:

$$\sum_{r=1}^k (P_{ir}(S) - P_{lr}(S)) = \sum_r^{(+)} (P_{ir}(S) - P_{lr}(S)) + \sum_r^{(-)} (P_{ir}(S) - P_{lr}(S)) = 0$$

$$h_{il} = \sum_r^{(+)} (P_{ir}(S) - P_{lr}(S)) = -\sum_r^{(-)} (P_{ir}(S) - P_{ls}(S))$$

$\sum^+$  - сума всіх невід'ємних різниць

$\sum^-$  - сума всіх від'ємних різниць

$$\forall i, l \quad 0 \leq h_{il} < 1, h = \max_{i,l} h_{il}, h < 1$$

В силу умови теореми (всі елементи матриці переходів  $\pi_s$  більше 0)

Повертаємось до виведення третього факту

$$\begin{aligned} &= \sum_r^{(+)} (P_{ir}(S) - P_{lr}(S)) P_{rj}(n-S) + \sum_r^{(-)} (P_{ir}(S) - P_{ls}(S)) P_{rj}(n-S) \leq \max_r P_{rj}(n-S) h_{il} - \\ &- \min_r P_{rj}(n-S) h_{il} \leq h \max_{i,l} |P_{ij}(n-2S)| \leq h^{\lfloor \frac{n}{S} \rfloor}, \lfloor \frac{n}{S} \rfloor - \text{ціле від ділення } n \text{ на } S \end{aligned}$$

Т.я.  $i, l$  – довільні отримані нерівність

$$\max_{i,l} |P_{ij}(n) - P_{lj}(n)| \leq h^{\lfloor \frac{n}{S} \rfloor}$$

Спрямуємо  $n \rightarrow \infty$  і враховуючи, що  $0 \leq h < 1$  отримали третій факт.

$$\text{Покажемо, що } \sum_{j=1}^k P_j = 1 :$$

$$\sum_{j=1}^k P_{ij}(n) = 1 \quad \text{як сума елементів будь-якого рядка матриці } \pi_n$$

Рівність вірна для будь-якого  $n$ .

$$\text{Так як } \exists \lim_{n \rightarrow \infty} \sum_{j=1}^k P_{ij}(n) = P_j \Rightarrow \sum_{j=1}^k P_j = 1$$

Знаходження чисел  $P_j$ :

Покажемо, що числа  $P_j$  задовольняють наступній системі рівнянь

$$\sum_{j=1}^k P_j = 1$$

$$P_j = \sum_{r=1}^k P_r P_{rj}, \quad j = \overline{i, k}$$

Використовується формула  $P_{ij}(n) = \sum_{r=1}^k P_{ir}(m) P_{rj}(n-m)$  де  $m=n-1$

$$P_{ij}(n) = \sum_{r=1}^k P_{ir}(n-1) P_{rj} \quad . \text{ Спрямуємо } n \rightarrow \infty$$

$$P_j = \sum_{r=1}^k P_r P_{rj}, \quad j = \overline{i, k}$$

Числа  $P_j$  існують, задовольняють  $k+1$  рівнянню і розв'язок єдиний.

Доводимо від супротивного: нехай є інший розв'язок  $x_1 \dots x_k$

$$\sum_{j=1}^k X_j = 1, \quad x_j = \sum_{r=1}^k x_r P_{rj}, \quad j = \overline{i, k}$$

Покажемо, що має місце сума:

$$x_j = \sum_{r=1}^k x_r P_{rj}(n), \quad j = \overline{i, k}$$

Доводимо, використовуючи ММІ: для  $n=1$  виконується, нехай виконується для  $n-1$ , покажемо що виконується для  $n$

$$x_j = \sum_{r=1}^k x_r P_{rj}(n-1) = \sum_{r=1}^k x_r \sum_{i=1}^k P_{rj}(n-1) P_{ir} = \sum_{i=1}^k P_{ij} \sum_{r=1}^k x_r P_{ri}(n-1) = x_j$$

Спрямуємо  $n \rightarrow \infty$ .

$$\text{Тоді } x_j = \sum_{r=1}^k x_r P_{rj}(n) \text{ для } n \rightarrow \infty \text{ отримуємо } x_j = \sum_{r=1}^k x_r P_j = P_j \sum_{r=1}^k x_r = P_j$$

Наша система має розв'язок, бо ми довели, що числа  $P_j$  існують і задовольняють цій системі рівнянь.

#### 5.4.2 Однорідні регулярні марківські процеси

Марківський процес — узагальнення Марківського ланцюга( використовується 2 означення Марківського ланцюга), а саме: система міняє свої стани не в конкретні моменти часу, а в будь-який час  $t$ .

$P_{ij}(t)$  - ймовірність того, що в момент часу  $t$  система перейде із стану  $A_i$  в стан  $A_j$ .  $P_{ij}(t) \equiv 0$  (оскільки  $P_{ij}(t)$  - ймовірність неперервної випадкової величини).

$P_{ij}(t, t + \Delta t)$  - ймовірність того, що система знаходиться в момент часу  $t$  в стані  $A_i$ , в момент часу  $t + \Delta t$  буде знаходитись в стані  $A_j$ .

$P_i(t)$  - ймовірність того, що в момент часу  $t$  система знаходиться в стані  $A_i$ .  $P_{ij}(t, t + \Delta t)$  не залежить від того, в яких станах знаходилась система до моменту часу  $t$

$$\sum_{i=1}^k P_i(t) = 1.$$

Марківський процес називається регулярним, якщо ймовірність того, що на відрізку довжини  $\Delta t$  буде 1 перехід системи із стану в стан є  $O(\Delta t) = a(t)\Delta t + o(\Delta t)$ ,  $0 < |a(t)| < \infty$  нескінченна мала порядку  $\Delta t$

а ймовірність того, що у відрізку часу  $\Delta t$  система перейде з стану в стан більше 1 раз рівна  $o(\Delta t)$  (нескінченно-мала більш високого порядку ніж  $\Delta t$ ).

Можна сказати, що якщо умова регулярності не виконується, то можливі два випадки:

- 4 На довільному інтервалі з ймовірністю 1 не буде жодного переходу
- 5 На довільному інтервалі з ймовірністю 1 буде необмежена кількість переходів.

З умови регулярності існує  $\lim_{\Delta t \rightarrow 0} \frac{P_{ij}(t, t + \Delta t) - P_{ij}(t)}{\Delta t} = \lambda_{ij}(t)$ , де  $\lambda_{ij}(t)$  - умовна інтенсивність переходу із стану  $A_i$  в стан  $A_j$  в момент часу  $t$ ,  $i \neq j$ ,  $i, j = \overline{1, k}$ .

Ця границя для регулярних марківських процесів завжди існує і записується як

$$\lambda_{ij}(t) = \lim_{\Delta t \rightarrow 0} \frac{P_{ij}(t, t + \Delta t) - P_{ij}(t)}{\Delta t} = \lambda_{ij}(t) \Delta t + o(\Delta t)$$

Регулярний Марківський процес зветься однорідним, якщо

$P_{ij}(t, t + \Delta t)$ , тоді

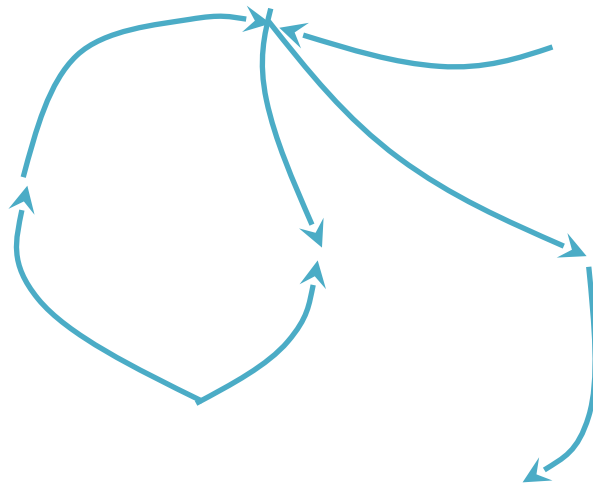
$$\lim_{\Delta t \rightarrow 0} \frac{P_{ij}(\Delta t) - P_{ij}(0)}{\Delta t} = \lambda_{ij}$$

### 5.4.3 Диференціальне рівняння Колмагорова

Зв'язує між собою  $P_i(t)$  та  $\lambda_{ij}$ .

Визначення: графік переходу, що задає регулярний однорідний марківський процес, називається орієнтований граф з  $k$ -вершинами(станами).

Вершина  $A_i$  зв'язана орієнтовним ребром з вершиною  $A_j$ , якщо система може перейти з  $A_i$  в  $A_j$  за один крок і під нею підписують  $\lambda_{ij} (\neq 0)$ .



Розглянемо приклад (див. малюнок).



Виведемо диф.рівняння Колмагорова для стану  $A_1$ : знайдемо ймовірність того, що  $P_1(t + \Delta t)$  - ймовірність того, що в момент  $t + \Delta t$  система знаход. в стані  $A_1$ .

Подія, ймовірність якої ми шукаємо, об'єднує наступні несумісні події

1)  $t$   $t + \Delta t$

$A_1$   $A_1$

2)  $t$   $t + \Delta t$

$A_2 \vee A_4$   $A_1$

3)  $t$   $t + \Delta t$

$A_6$   $A_1$

Використовуємо формулу  $P(AB) = P(A) \cdot P(B / A)$ .

1) Першу ймовірність знайдемо, як протилежну: в момент часу  $t$  знаходимось в стані  $A_1$ , а в  $t + \Delta t$  - не в  $A_1$ .

$$P_1(t)(1 - (\lambda_{13}\Delta t + \lambda_{15}\Delta t + o(\Delta t))), \text{ де } o(\Delta t) = o_1(\Delta t) + o_2(\Delta t) + o_3(\Delta t)$$

Використовуємо результат, що якщо  $A \subset B$ , то  $P(A) \leq P(B)$ , тобто  $o_3(\Delta t)$  - верхня оцінка ймовірності перебування в момент  $t + \Delta t$  в стані  $A_7$ , якщо вона в момент часу  $t$  була в стані  $A_1$ , де  $A$  - подія на інтервалі  $\Delta t$  перейти в стані  $A_7$ , а  $B$  - на інтервалі  $\Delta t$  у системі не менше 2-х переходів - ймовірність, яку позначили  $o_3(\Delta t)$ . (умова ординарності регулярного харківського процесу)

2) Знаходимо ймовірність другої події:

$$P_2(t)\lambda_{21}\Delta t + P_4(t)\lambda_{41}\Delta t + o(\Delta t)$$

3) Ймовірність третьої події

Верхня границя ймовірності третьої події  $o(\Delta t)$ , бо ця подія належить події: у марківського процесу має бути більше ніж 1 перехід із стану в стан на відрізок  $\Delta t$ .

В силу умови регулярності верхня оцінка ймовірності  $o(\Delta t)$

$$P(t+\Delta t) = P_1(t)(1 - \lambda_{13}\Delta t - \lambda_{15}\Delta t) + P_6(t)\lambda_{61}\Delta t + P_7(t)\lambda_{71}\Delta t + o(\Delta t)$$

Всі  $o(\Delta t)$  зібрані в 1, знак – неважливо.

$$\frac{P_1(t+\Delta t) - P_1(t)}{\Delta t} = \frac{-P_1(t)\lambda_{12}\Delta t - P_1(t)\lambda_{13}\Delta t + P_2(t)\lambda_{21}\Delta t + P_4(t)\lambda_{41}\Delta t}{\Delta t} +$$

$$\frac{o(\Delta t)}{\Delta t}, t \rightarrow \infty$$

$$\frac{P_1(t+\Delta t) - P_1(t)}{\Delta t} = \frac{-P_1(t)\lambda_{12}\Delta t - P_1(t)\lambda_{13}\Delta t + P_2(t)\lambda_{21}\Delta t + P_4(t)\lambda_{41}\Delta t}{\Delta t} + \frac{o(\Delta t)}{\Delta t}, t \rightarrow \infty.$$

Отримаємо:

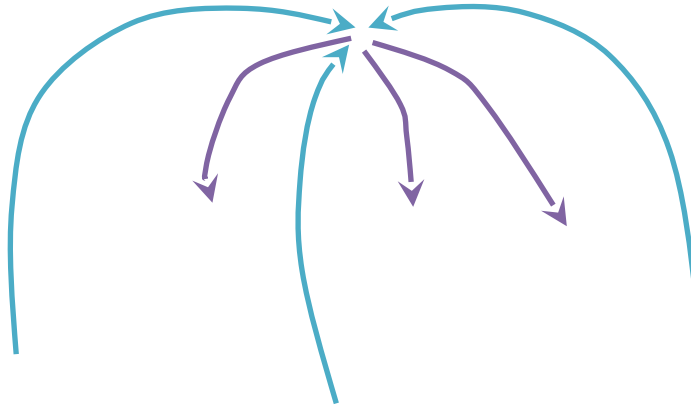
$$\frac{dP_1(t)}{dt} = -P_1(t)\lambda_{12} - P_1(t)\lambda_{13} + P_2(t)\lambda_{21} + P_4(t)\lambda_{41}$$

Границю справа існує, звідки існує границя зліва і вона є похідною

$$\frac{dP_1(t)}{dt}$$

Загальний випадок

Розглянемо фрагмент загального графу переходів, що задає загальний фрагмент марківського процесу:



обмежена кількість станів, з яких можна перейти в  $A_i$  за один крок і в які можна перейти з  $A_i$  за один крок.

Виводимо диференціальне рівняння Колмагорова:

Подія в момент часу  $t + \Delta t$  бути в стані  $A_i$  розкладатися на систему попарно – несумісних подій

$$1) \begin{matrix} t & t + \Delta t \\ A_i & A_i \end{matrix}$$

$$A_i$$

$$2) \begin{matrix} t & t + \Delta t \\ A_{m_1} \cup \dots \cup A_{m_{k_2}} & A_i \end{matrix}$$

$$A_{m_1} \cup \dots \cup A_{m_{k_2}}$$

$$3) \begin{matrix} t & t + \Delta t \\ \forall l, A_l & A_i \end{matrix}$$

$$\forall l, A_l$$

З стану  $A_l, \forall l$ , в стан  $A_i$  можна перейти не менш ніж за 2 кроки.

Використовуємо формули:  $P(AB) = P(A) \cdot P(B / A)$ , якщо  $A \subset B$  то  $P(A) \leq P(B)$ .

$$\text{Перша подія має ймовірність настання } P_i(t) \left( 1 - \sum_{l=1}^{k_1} \lambda_{ij_l} \Delta t + o(\Delta t) \right).$$

Примітка: в нашому випадку  $P(A) = P_i(t)$ ,  $P(B/A)$  - бути в момент часу  $t + \Delta t$  в стані  $A_i$  при умові, що в момент часу  $t$  були в стані  $A_i$  знаходиться через обернену подію бути в момент часу в будь-якому стані крім  $A_i$ .

Ймовірність 2-ї події:

$$\sum_{h=1}^{k_2} P_{m_h}(t) \lambda_{m_h l} \Delta t + o(\Delta t) \text{ (за 3 аксіом. теорії ймовірності)}$$

Ймовірність 3-ї події:

$$O(\Delta t)$$

Використовуємо: якщо  $A \subset B$  то  $P(A) \leq P(B)$

де подія  $A$  - знаходитись в момент часу  $t$  в будь-якому стані, з якого можна перейти не менше ніж за два кроки в момент часу  $t + \Delta t$  - бути в  $A_i$ , а  $B$  – на інтервалі  $\Delta t$  було більше ніж 1 перехід із стану в стан.

Таким чином, отримали наступне рівняння

$$P_i(t + \Delta t) = P_i(t) \left( 1 - \sum_{l=1}^{k_1} \lambda_{i j_l} \Delta t \right) + \sum_{h=1}^{k_2} P_{m_h}(t) \lambda_{m_h i} \Delta t + o(\Delta t)$$

Аналогічно прикладу:

$$\frac{dP_i(t)}{dt} = - \sum_{l=1}^{k_1} P_i(t) \lambda_{i j_l} + \sum_{h=1}^{k_2} P_{m_h}(t) \lambda_{m_h i}$$

*Мнемонічне правило побудови диф.рівняння Колмагорова*

Зліва – похідна від безумовної ймовірності. Справа – стільки членів, скільки дужок зв'язано зі станом переходів з вершиною  $A_i$  в графі переходів, що задає марківські процеси. Кожен член дорівнює добутку безумовної ймовірності перебування в момент часу  $t$  в стані, з якого виходить дужка, помножена на умовну інтенсивність переходу, приписана цій дужці.

Члени, що відповідають дужкам, що виходять із стану  $A_i$  і мають знак «-», інші мають знак «+».

Примітка: інколи відкидають одне з рівнянь і його заміняють алгебраїчним:

$$\sum_{i=1}^k P_i(t) = 1 \quad . \text{ Початкові умови: } P_i(0), \quad i = \overline{1, k} .$$

*Узагальнення*

диференціальне рівняння Колмагорова залишається таким самим і для випадку, коли станів марківського процесу злічена нескінченна кількість, при умові що обмежена кількість ребер входить і обмежена кількість ребер виходить. Дійсно доведення для цього випадку повністю збігається з приведеним вище.

### 5.5 Пуасонівський (найпростіший) потік подій

Потоком подій зветься послідовна поява в часі подій. Потоки подій бувають детермінованими та випадковими.

Детермінований потік подій – це такий потік подій у якого інтервали часу між сусідніми подіями є детермінованими в часі.

Випадковий потік подій – це такий потік подій у якого інтервали часу між сусідніми подіями є випадкові величини.

Пуасонівський (найпростіший) потік подій – це такий потік подій для якого виконуються наступні три умови:

1) *Стаціонарність*: ймовірність того, що на довільному часовому відрізку попаде певна кількість подій залежить тільки від довжини цього відрізку і не залежить від того, де він розташований на часовій вісі.

2) *Ординарність*: ймовірність того, що на часовий відрізок довжини  $\Delta t$  з'явиться одна подія є нескінченно мала порядку  $\Delta t$ , а дві чи більше подій – нескінченно мала більш високого порядку ніж  $\Delta t$ .

3) *Безпослідовність*: ймовірність того, що на довільному часовому відрізку з'явиться певна кількість подій не залежить від того скільки подій з'явилося на часових відрізках, що не перетинаються з даним.

Доводиться (перша тема другої частини спец. курсу) що в цьому випадку ймовірність того, що на часовому відрізку довжини  $t$  з'явиться  $m$  подій дорівнює:

$$\pi_m = \frac{(\lambda t)^m}{m!} e^{-\lambda t}$$

$\lambda$  – математичне сподівання кількості подій, що з'явиться на часовому відрізку довжини 1.

Зрозуміло, що з умови стаціонарності випливає, що інтервали часу між сусідніми подіями потоку Пуасона є реалізаціями тієї ж самої випадкової величини (це стане ясно з процедури знаходження функції розподілу цієї випадкової величини).

Примітка: в цьому означенні розглядається не пуасонівський потік, а будь-яка його реалізація.

Дійсно, нехай  $T \geq 0$  – випадкова величина, інтервал часу між сусідніми подіями потоку Пуасона.

Знайдемо її функцію розподілу:

$$F(t) = 0, t < 0$$

$$F(t) = P(0 \leq T < t) = 1 - P(T \geq t) = 1 - \pi_0 = 1 - e^{-\lambda t}$$

Отримали функцію експоненціального розподілу:

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

Знайдемо ймовірність того, що на інтервалі довжини  $\Delta t$  з'явиться одна подія. Для цього в формулі Пуасона підставимо  $m=1, t=\Delta t$ :

$$\lambda \Delta t e^{-\lambda \Delta t}$$

Виконавши розвинення  $e^{-\lambda \Delta t}$  в ряд Маклорена, отримаємо:

$$\lambda \Delta t (1 - \lambda \Delta t + o(\Delta t)) = \lambda \Delta t + o(\Delta t)$$

Таким чином отримали, що  $\lambda$  одночасно є і математичним сподіванням кількості подій, що з'явиться на відрізку довжини 1, і параметром умови ординарності.

### 5.6 Потоки Пальма

Випадковий потік подій зветься потоком Пальма, якщо випадкові величини (інтервал часу між сусідніми подіями) є незалежними.

Пуасонівський потік – це потік Пальма, оскільки в цьому випадку інтервали часу між сусідніми подіями є незалежні випадкові величини, що мають один і той самий розподіл (експоненціальний з параметром  $\lambda$ ). Факт незалежності цих випадкових величин буде доведений в спец. курсі.

### 5.7 Потоки Ерланга $k$ -го порядку

Потік Ерланга  $k$ -го порядку - це стаціонарний потік Пуассона, в якому залишається кожна  $k$ -та подія.

*Наприклад*

Потік Ерланга 3-го порядку виглядає наступним чином( червоні хрестики – події потоку Ерланга, всі хрестики – події потоку Пуассона):



*Ймовірнісні характеристики потоку Ерланга  $k$ -го порядку*

$T$  – випадкова величина – інтервал часу між настанням сусідніх подій потоку Ерланга  $k$ -го порядку.

$T_i$  – інтервал часу між сусідніми подіями потоку Пуассона, який породив потік Ерланга  $k$ -го порядку.

$$T = \sum_{i=1}^k T_i$$

$$MT_i = \frac{1}{\lambda}$$

$$DT_i = \frac{1}{\lambda^2}$$

$$MT = \frac{k}{\lambda}$$

$$DT = \frac{k}{\lambda^2}$$

Знайдемо функцію щільності для  $T$ :

$$P(t \leq T \leq t + \Delta t) = f(t) \Delta t + o(\Delta t)$$

Знайдемо імовірність цієї події  $P(t \leq T \leq t + \Delta t)$  – на інтервалі  $\Delta t$  настала  $k$ -та подія потоку Пуассона.

Подія складається з двох несумісних подій:

1) На  $t$  настала  $k-1$  подія потоку Пуассона, на  $\Delta t$  – єдина  $k$ -та подія.

2) На  $\Delta t$  настало більше однієї події  $(2, \dots, k)$ , а на  $t$  те, що залишилось з  $k$  подій – подія  $A$ . Ця подія належить події  $B$  – на  $\Delta t$  настало більше однієї події –  $A \subseteq B$ . З умови ординарності  $P(B) = o(\Delta t)$ .

Знайдемо ймовірність першої події з умови безпослідності Пуассонівського потоку:

$$\frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t} (\lambda \Delta t + o(t))$$

Тоді:

$$P(t \leq T \leq t + \Delta t) = \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t} (\lambda \Delta t + o(t) + o(t))$$

$$f(t) = \frac{\lambda t^{k-1}}{(k-1)!} e^{-\lambda t}$$

Інколи ймовірнісні характеристики потоку Ерланга  $k$ -ого порядку виражаються через його параметр. По аналогії з потоком Пуассона він



задається через  $\Lambda$  – математичне сподівання кількості подій потоку Ерланга  $k$ -го порядку, настає за  $t=1$ . Тоді:

$$\Lambda = \frac{\lambda}{k}; DT = \frac{1}{k \Lambda^2}$$

*Наслідки:*

1) Нехай  $k \rightarrow \infty$ ,  $\Lambda$  – не змінюється.

*Примітка:* необхідна умова існування:  $\lambda \rightarrow \infty$  при  $k \rightarrow \infty$ . Тоді:

$$MT = \frac{1}{\Lambda} DT = 0$$

Тобто в залежності від параметрів потоку Ерланга, його властивості змінюються від повної безпослідності ( $k=1$  - потік Ерланга перетворюється в потік Пуассона) до детермінованого потоку з однаковими фіксованими інтервалами часу між сусідніми подіями.

2) В багатьох реальних задачах реальний стаціонарний потік подій апроксимується (замінюється) відповідним потоком Ерланга.

*Наприклад*

Нехай знайдено оцінки  $MT = \hat{\nu}, DT = \hat{\sigma}^2$  для реального стаціонарного потоку. Тоді:

$$\Lambda = \frac{1}{\hat{\nu}}; \hat{\sigma}^2 = \frac{1}{k \Lambda^2}; k = \left\lfloor \frac{\hat{\nu}}{\hat{\sigma}^2} \right\rfloor, \lfloor \cdot \rfloor - \text{ціле відделення}$$

### 5.8 Зв'язок між регулярними однорідними марківськими процесами та пуасонівськими потоками

Зв'язок між регулярними однорідними марківськими процесами та потоками Пуассона

Маємо систему  $A_1 \dots A_k$  (система в 1 з  $k$  станів).

$$k_i < k, i \notin \{j_1, \dots, j_{k_i}\}$$

Нехай система в момент часу  $t$  миттєво перейшла в стан  $A_i$ . В цьому стані

На систему діє  $k_i$  незалежних пуасонівських потоків.

Якщо першим в систему, що знаходиться в стані  $A_i$ , приходить подія з потоку з параметром  $\lambda_{ij_i}$ , то система миттєво переходить в  $A_{ij_i}$ , далі ситуація повторюється.

Примітка. Пуасонівські потоки починають працювати не в момент часу  $t$ , а в  $\forall$  моменти часу, що передують  $t$ .

Покажемо, що переходи системи зі стану в стан задають регулярний однорідний марківський процес і параметри  $\lambda_{ij_1} \dots \lambda_{ij_k}$  є умовними інтенсивностями переходів марківського процесу.

$$P_{ij}(\Delta t) = \lambda_{ij} \Delta t + o(\Delta t)$$

Для доведення цього факту доведемо 2 теореми.

1. Умовний експонційний розподіл дорівнює безумовному.

*Доведення.*

$T$ -випадкова величина з експонційним розподілом з параметром  $\lambda$ .

$$\text{Тоді } P(T < t_2 / T \geq t_1) = P(T < t_2 - t_1)$$

Функція щільності експонційного розподілу :

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

$$t_2 > t_1$$

Ця умовна ймовірність не залежить від  $t_1$ .

$$A = \{T < t_2\}$$

$$B = \{T \geq t_1\}$$

$$t_1 < t_2$$

$$A \cdot B = \{t_1 \leq T < t_2\}$$

$$P(A \cdot B) = F(t_2) - F(t_1) = 1 - e^{-\lambda t_2} - 1 + e^{-\lambda t_1} = e^{-\lambda t_1} - e^{-\lambda t_2}$$

$$P(B) = 1 - P(\dot{B}) = 1 - 1 + e^{-\lambda t_1} = e^{-\lambda t_1}$$

Тоді :

$$P(A/B) = \frac{P(A \cdot B)}{P(B)} = \frac{e^{-\lambda t_1} - e^{-\lambda t_2}}{e^{-\lambda t_1}} = 1 - e^{-\lambda(t_2 - t_1)} = F(t_2 - t_1) = P(T < t_2 - t_1)$$

*Теорему доведено*

Теорема 2. Сума обмеженої кількості незалежних пуасонівських потоків є пуасонівським потоком.

*Доведення*

Доведемо, що сумарний потік має властивості :

1. *Стаціонарності*

2. *Ординарності*

3. *Безпіслядії*

Тоді за означенням він буде пуасонівським.

Розглянемо часовий відрізок довжини  $t$  і задамо подію  $A_k$  - у сумарного потоку на відрізку довжини  $t$  настало  $k$  подій.  $A_k$ - об'єднання попарно несумісних подій  $A_{ik}$ , де подія  $A_{ik}$ —  $i$  — та ситуація при якій у сумарного потоку на відрізку довжини  $t$  з'являється  $k$  подій.

В силу незалежності пуасонівських потоків

$$P(A_{ik}) = \prod_{j=1}^m P(A_{ik}^j)$$

$A_{ik}^j$ - кількість подій які мають настати у  $j$ -ому потоці на відрізку довжиною  $t$  у ситуації  $A_{ik}$ .

Тоді:

$$P(A_k) = \sum_i P(A_{ik}) = \sum_i \prod_{j=1}^m P(A_{ik}^j)$$

В силу незалежності пуасонівських потоків.

Ми довели стаціонарність та безпослідию.

*Примітки*

Так як цим властивостям відповідає  $P(A_{ik})$ , тім задовольняє і  $P(A_k)$ ,

Ординарність – ймовірність того, що на відрізку з’явиться 1 подія – нескінченно мала порядку  $\Delta t$ , а ймовірність того, що з’явиться більше 1 події нескінченно мала більш високого порядку ніж  $\Delta t$ .

Доведемо ординарність з початку. Доводимо другу частину умови ординарності:

на інтервалі  $\Delta t$  поява більш ніж однієї події є нескінченно мала більш високого порядку ніж  $\Delta t$ .

Розіб’ємо усю подію на дві несуміні події :

1. У кожного пуассонівського потоку на інтервалі  $\Delta t$  з’являється або одна подія, або нічого

2. У одного пуассонівського потоку з’являються на відрізку довжини  $\Delta t$  дві чи більше події, а у всіх інших що завгодно

Верхня оцінка другої події :  $m \cdot o(\Delta t) = o(\Delta t)$ .

*Примітка*

$$P(i=1 \dots m A_i) \leq \sum_{i=1}^m P(A_i)$$

Знаходимо верхню границю першої події.

$$\lambda = \max_i \lambda_i, i=1..m$$

Тоді верхня границя першої події дорівнює (використовується біноміальний розподіл):

$$\sum_{i=2}^m C_m^i \cdot (\lambda \cdot \Delta t + o(\Delta t))^i \cdot (e^{-\lambda \Delta t})^{m-i}$$

$$\text{де } (e^{-\lambda \Delta t})^{m-i} = (1 - \lambda \cdot \Delta t + o(\Delta t))^{m-i}$$

У цього степеневому поліному найменша ступінь  $(\Delta t)^2$ . Сума двох нескінченно малих більш високого порядку ніж  $\Delta t$  є нескінченно малою більш високого порядку ніж  $\Delta t$ .

Знаходимо ймовірність появи однієї події.

$$\sum_{i=1}^m (\lambda_i \cdot \Delta t + o(\Delta t)) \prod_{\substack{j=1, \\ j \neq i}}^m e^{-\lambda_j \Delta t} = \left( \sum_{i=1}^m \lambda_i \right) \cdot \Delta t + o(\Delta t)$$

$$\text{Де } e^{-\lambda_j \Delta t} = (1 - \lambda_j \cdot \Delta t + o(\Delta t))$$

*Теорему доведено*

Повертаємось до дослідження системи, що змінює свої стани під дією незалежних пуасонівських потоків.



Використовуючи першу теорему можна вважати, що всі пуасонівські потоки запустилися в момент часу  $t$  в якій система перейшла в стан  $A_i$ .

Дійсно за першою теоремою ймовірність настання подій пуасонівського потоку не залежить від того коли настала попередня, отже можна вважати, що всі пуасонівські потоки запустились.

Таким чином умова марковості виконана.

В силу того, що система може змінити стан в  $\forall$  момент часу, в ній функціонує марківський процес і залишилось довести, що він регулярний, ординарний і параметри  $\lambda_{ij}$  - це умовні інтенсивності переходу марківського процесу.

*Примітка.*

Показати, що виконується умова марковості – умовна ймовірність переходів із стану в стан не залежить від того, в якому стані перебувала система до моменту.

Розглянемо ймовірність того, що система перейде з  $A_i$  в  $A_j$  за  $\Delta t$ , тобто  $P_{ij}(\Delta t)$

Ця подія розкладається на дві несумісні :

1) На інтервалі  $\Delta t$  приходить подія з параметром  $\lambda_{ij}$  і на цьому інтервалі у всіх інших незалежних потоків події не з'являються.

$$\lambda_{ij} \cdot \Delta t + o(\Delta t) \prod_{\substack{l=1 \\ j_l \neq j}}^{k_i} e^{-\lambda_{il} \Delta t} = \lambda_{ij} \cdot \Delta t + o(\Delta t)$$

2) На інтервалі  $\Delta t$  з'явилась не одна, а декілька подій з різних пуасонівських потоків, але подія з потоку з параметром  $\lambda_{ij}$  прийшла першою. Якщо просумувати усі потоки що діють в системі, то сумарний потік за другою теоремою є пуасонівським і в нього повинно з'явитися більше однієї події, за умовою ординарності ймовірність цієї події  $o(\Delta t)$

$$P_{ij}(\Delta t) = \lambda_{ij} \cdot \Delta t + o(\Delta t)$$

Умовна інтенсивність марківського процесу – параметр відповідного пуасонівського потоку.

Покажемо, що ймовірність зміни станів більше одного разу  $o(\Delta t)$ .

Це так, тому щоб система на інтервалі  $\Delta t$  змінила більше одного стану необхідно щоб у сумарного пуасонівського потоку на інтервалі  $\Delta t$  з'явилося не менше двох подій. Тому верхня оцінка ймовірності цієї події  $o(\Delta t)$ .

### *Наслідок 1*

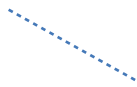
Нехай кількість станів системи злічена нескінченна, але у відповідному графі переходів, що задає марківський процес з кожною вершиною

пов'язана обмежена кількість ребер. Покажемо, що результат залишається незмінним.

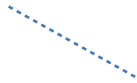
$$P_{ij}(\Delta t) = \lambda_{ij} \Delta t + o(\Delta t)$$

Те що процес є регулярним однорідним марківським доводиться практично так само. Зміниться лише наступне доведення : на інтервалі  $\Delta t$  ймовірність зміни більше одного стану – нескінченно мала  $o(\Delta t)$ .

Доведемо цей факт: розглянемо наступні фрагменти графу переходів



Відповідає ребрам, що виходять з вершини  $A_i$ .



Відповідає ребрам, що входять в вершину  $A_j$

Дамо еквівалентне означення події : на інтервалі довжиною  $\Delta t$  система змінює більш ніж один стан. Для  $\forall$  станів  $i, j$  перехід системи на інтервалі  $\Delta t$  з  $A_i$  в  $A_j$  більш ніж за один крок є нескінченно малою більшого порядку ніж  $\Delta t$ , тобто  $o(\Delta t)$ .

Подія з  $A_i$  перейти в  $A_j$  за два чи більше кроків належить наступній події : у сумарного пуасонівського потоку, що складається з усіх незалежних пуасонівських потоків, що діють на систему в стані  $A_i$  і всіх незалежних пуасонівських потоків, що приводять систему в  $A_j$  за один крок



(за теоремою 2 – це пуассонівський потік) на  $\Delta t$  повинно з'явитись дві чи більше події. За умовою ординарності ця ймовірність є  $o(\Delta t)$ .

### *Наслідок 2*

Той самий результат отримується якщо систему із стану в стан переводять події одиничні, що мають експоненційний розподіл.

$$\begin{cases} \lambda_{ij_i} \cdot e^{-\lambda_{ij_i} t}, t \geq 0 \\ 0, t < 0 \end{cases}$$

На систему діють не незалежні пуассонівські потоки, а незалежні випадкові величини, що мають експоненційний розподіл.

Дійсно, враховуючи те, що систему із стану в стан переводять перші події пуассонівських потоків, що зустрічаються в системі, за першою теоремою усі потоки запустились у момент часу  $t$ , та те, що інтервали часу появи першої події має експоненційний розподіл, то можемо вважати, що наші події є першими подіями, віртуально існуючих незалежних пуассонівських потоків.

### *Наслідок 3*

Той самий результат отримаємо, якщо систему із стану в стан переводять як перші події незалежних пуассонівських потоків так і незалежні випадкові величини, що мають експоненційний розподіл. Довести самим.

## 5.9 Гранична теорема для однорідних регулярних марківських процесів

Надається без доведення.

Нагадаємо граничну теорему для однорідних марківських ланцюгів (регулярні однорідні марківські процеси – узагальнення однорідних марківських ланцюгів).

Існує таке натуральне  $S$ , щовсі елементи  $P_{ij}(S) > 0$

$$\forall i \exists \lim_{n \rightarrow \infty} P_{ij}(n) = P_j$$

умова існування граничних ймовірностей наступна – в орієнтованому графі, що задає однорідний регулярний марківський процес, є орієнтований шлях з будь-якого стану в будь-який інший, включаючи сам цей стан.

Знаходження граничних ймовірностей

Так як  $P_i(t)$  задовольняє диф. рівнянням Колмагорова і існує  $\lim_{t \rightarrow \infty} P_i(t) = P_i$ ,

то на нескінченності  $\frac{dP_i(t)}{dt}$  вироджується і має границю 0. Таким чином, алгоритм знаходження  $P_i$  наступний:

Ліву частину заміняємо на 0, а праву частину – підставляємо замість  $P_i(t)$  граничні  $P_i$ . Отримаємо систему лінійних рівнянь. Для її розв'язку

треба викинути на будь-яке рівняння і замінити його  $\sum_{i=1}^n P_i = 1$ . Отримаємо систему k- рівнянь з k невідомими, у якій вектор – стовпець правої частини не є 0, бо одна з компонент дорівнює 1.

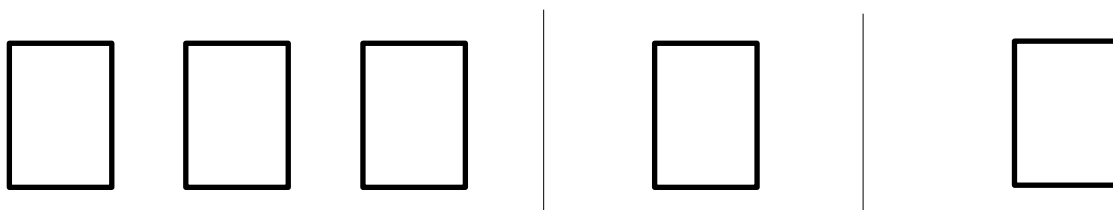
На практиці гранична ймовірність використовується наступним чином: знаходимо таке  $t'$ , для якого  $\forall t \geq t'$  виконується:  $P_i(t) \approx P_i$  (настання стаціонарного режиму).

Тоді, взявши інтервал часу довжини T, можна довести, що числа  $P_1T, P_2T, \dots, P_kT$  є мат.сподіваннями кількості часу, в яких система перебуває в станах  $A_1, A_2, \dots, A_k$  на інтервалі довжини T.

$P_i T, i = \overline{1, k}$  - мат.сподівання кількості часу, в якому система перебуває в стані  $A_i$  на інтервалі  $T$ .

### 5.10 Процес розмноження і загибелі

Регулярний марківський процес зветься процесом розмноження і загибелі(цим процесом моделюється життєвий цикл відповідних біологічних популяцій), який має наступний граф переходу:



Цей граф переходів задовольняє граничній теоремі. Знайдемо граничні ймовірності станів  $P_1, P_2 \dots P_n$ .

Для цього в рівняння Колмагорова в лівій частині кладемо  $P_i = 0$ , і з правої частини переносимо частину в ліву члени, що мають знак «-».

$$\begin{aligned} P_1 \lambda_{12} &= P_2 \lambda_{21} \\ P_2 \lambda_{23} &= P_3 \lambda_{32} \\ &\dots \\ P_i \lambda_{i,i+1} &= P_{i+1} \lambda_{i+1,i} \\ &\dots \\ P_{n-1} \lambda_{n-1,n} &= P_n \lambda_{n,n-1} \end{aligned}$$

Примітка: останнє рівняння для  $A_n$  збігається з передостаннім. Тому останнє рівняння замінюємо на  $\sum_{i=1}^n P_i = 1$

$$P_2 = \frac{\lambda_{12}}{\lambda_{21}} P_1$$

$$P_3 = \frac{\lambda_{23}}{\lambda_{32}} P_2 = \frac{\lambda_{12}\lambda_{23}}{\lambda_{21}\lambda_{32}} P_1$$

$$\dots$$

$$P_i = \frac{\lambda_{12} \cdot \dots \cdot \lambda_{i-1,i}}{\lambda_{21}\lambda_{32} \cdot \dots \cdot \lambda_{i,i-1}} P_1$$

Мнемонічне правило:

$$P_{i=2,n} = \frac{\lambda_{12} \cdot \dots \cdot \lambda_{i-1,i}}{\lambda_{21}\lambda_{32} \cdot \dots \cdot \lambda_{i,i-1}} P_1$$

$$P_1 \text{ знаходимо з рівняння } \sum_{i=1}^n P_i = 1 :$$

$$P_1 = \frac{1}{1 + \frac{\lambda_{12}}{\lambda_{21}} + \dots + \frac{\lambda_{12} \cdot \dots \cdot \lambda_{n-1,n}}{\lambda_{21} \cdot \dots \cdot \lambda_{n,n-1}}}$$

### 5.11 Системи масового обслуговування марківського типу

Системою масового обслуговування зветься система каналів обслуговування призначених для обслуговування заявок. СМО задає правила, по яким заявки попадають на обслуговування чи залишають систему не обслугованими.

На вхід поступає пуассонівський стаціонарний потік заявок на обслуговування. Задачею теорії систем масового обслуговування є знаходження аналітичних виразів ймовірнісних характеристик ефективності роботи СМО. Всі ймовірнісні характеристики знаходяться для стаціонарного режиму роботи СМО (ймовірності всіх станів є граничними – від часу не залежать).

*Класифікація СМО, які розглядаються в курсі:*

СМО з відмовами без черги – якщо заявка, що прийшла в систему, застає хоча б 1 канал вільний – стає на обслуговування. Якщо всі канали зайняті – не обслуговується.

Ймовірнісні характеристики її роботи

1) Абсолютна пропускна спроможність  $A$  – мат.сподівання кількості заявок, що обслуговують систему за одиницю часу.

2) Відносна пропускна спроможність  $q$  – відношення абсолютної пропускної спроможності до мат.сподівання кількості заявок, що поступають в систему за одиницю часу.

$$q = \frac{A}{\lambda}$$

3)  $P_{об}$  - ймовірність того, що заявка буде обслугована

4)  $P_{від}$  - ймовірність того, що заявка залишить СМО не обслугованою.

5) Мат.сподівання кількості зайнятих каналів.

СМО з відмовами і обмеженою чергою - якщо заявка застала всі канали зайнятими – встає в чергу останньою; якщо довжина черги дорівнює граничній – залишає систему не обслугованою.

Крім вищеназваних характеристик додаються додаткові:

6) Мат.сподівання довжини черги

7) Мат.сподівання часу перебування в черзі.

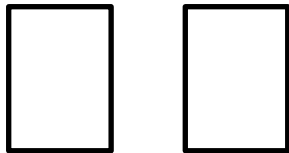
СМО без відмов – в цьому випадку черга є нескінченною. Далі буде показано, що якщо вхідний потік заяв пуассонівський, а канали обслуговують заявки незалежно один від одного закону і час обслуговування має експоненційний розподіл, то в СМО функціонує однорідний регулярний марківський процес.

### 5.11.1 Одноканальна СМО з відмовами

Є один канал обслуговування заявки, час обслуговування випадковий, що має параметр  $\mu$ .

На вхід системи поступає пуассонівський потік заяв з параметром  $\lambda$ . Заявка, попавши в систему і заставши канал вільним, миттєво поступає на обслуговування. Якщо канал зайнятий – залишає систему не обслуговуваною.

В системі функціонує однорідний регулярний марківський процес з наступним графом переходів



де стани системи:  $A_0$  - канал вільний,  $A_1$  - канал зайнятий.

*Примітка.* Показати, що маємо частковий випадок задачі (розділ 4.17, наслідок 3)

Запишемо диф. рівняння Колмагорова для стану  $A_0$

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t) + \mu P_1(t)$$

$$P_0(t) + P_1(t) = 1$$

$$P_0(0) = 1$$

$$\frac{dP_0(t)}{dt} = -P_0(t)(\lambda + \mu) + \mu$$

$$P_0(t) = \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t} + \frac{\mu}{\lambda + \mu}$$

$$P_1(t) = \frac{-\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t} + \frac{\lambda}{\lambda + \mu}$$

Можна строго довести, що відносна пропускна спроможність  $\eta$  – ймовірність обслуговування заявки в стаціонарному режимі.

Дійсно, це впливає з аналізу виразу для відносної пропускної спроможності

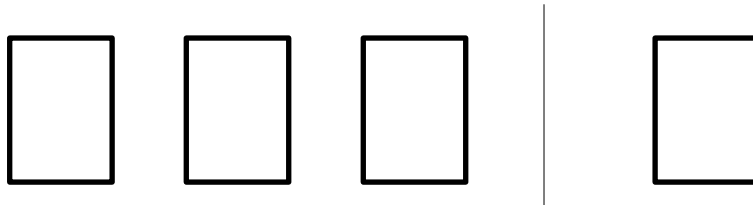
$q = P_0 = \frac{\mu}{\lambda + \mu}$ ,  $A = \lambda q$ . Ймовірність того, що заявка залишить систему не обслугованою -  $P_1 = \frac{\lambda}{\lambda + \mu}$

#### 5.11.2 Багатоканальна СМО з відмовами

Маємо  $n$ – незалежних каналів обслуговування, кожен з яких обслуговує заявку по експоненційному закону з параметром  $\mu$ . Якщо заявка застала хоча б 1 канал вільний – миттєво стає на обслуговування. Якщо всі канали зайняті – залишає систему не обслугованою.

$A_0$  - всі канали вільні,  $A_1$  - 1 канал зайнятий...  $A_n$  -  $n$  каналів зайняті.

Покажемо, що в системі функціонує однорідний регулярний марківський процес з наступним графом:



Так як з ймовірністю 1 (показується на спецкурсі) у пуасонівського потоку подій з'являються по одній, то в графі переходів система переводиться в сусідній стан справа.

Для системи в стані  $A_0$  і  $A_1$  повністю повторюється розмірковування попередньої задачі.

Нехай в момент часу  $t$  система перейшла в стан  $A_2$ .

Використовуємо першу теорему:

Розглядаємо систему в стані  $A_2$ . В ньому діє фізично існуючий пуассонівський потік заяв з параметром  $\lambda$ , тому верхні стрілки стану  $A_2$  (і всіх інших станів) мають параметр  $\lambda$ . На систему в стані  $A_2$  діють 2 незалежні випадкові величини – заявки, що обслуговуються. За I теоремою можна вважати, що обидва канали стали обслуговувати заявки в момент часу переходу системи в стан  $A_2$ . Якщо першими обслужать заявку один з 2 каналів, а не нова заявка прийде в систему, то система перейде в стан  $A_1$ .

Для того, щоб знайти умовну інтенсивність переходів цього ребра використовуємо прийом, який реалізовано в попередній лекції: будемо вважати заявку, що обслуговується першим каналом, є першою заявкою віртуального пуассонівського потоку обслуговуваних заявок з параметром  $\mu$ . Другу заявку, що обслуговується другим каналом також вважаємо першою заявкою віртуального пуассонівського потоку обслуговуваних заявок з параметром  $\mu$ .

Це дозволяє вважати, що на систему в стані  $A_2$  діє лише 2 незалежних пуассонівських потоки заяв:

Перший – фізично існуючий пуассонівський потік заявок на обслуговування з параметром  $\lambda$ .

Другий – сумарний віртуальний пуассонівський потік обслужених заявок з параметром  $2\mu$ .



$2\mu$  і є умовною інтенсивністю переходу з стану  $A_2$  в стан  $A_1$ .

Привести самим аналогічні розміркування для станів  $A_3, \dots, A_n$  і таким чином показати, що отримали частковий випадок задачі (розділ 4.17, наслідок 3)

$$\rho = \frac{\lambda}{\mu}$$

Цей граф є частковим випадком регулярного марківського процесу розмноження і загибелі.

Використовуємо вирази його граничних ймовірностей

$$P_1 = \rho \cdot P_0$$

$$P_2 = \frac{\rho^2}{2!}$$

...

$$P_n = \frac{\rho^n}{n!} P_0$$

$$P_0 = \frac{1}{1 + \rho + \frac{\rho^2}{2!} + \dots + \frac{\rho^n}{n!}}$$

$$q = 1 - P_n$$

$$A = \lambda(1 - P_n)$$

Знайдемо мат.сподівання кількості зайнятих каналів

$$MZ = 1 \cdot P_1 + 2 \cdot P_2 + \dots + n \cdot P_n$$

$$MZ = \frac{A}{\mu}$$

$A$ -мат.сподівання кількості заявок, що обслуговуються системою за 1 часу.  $A$  параметр  $\mu$  - можна вважати мат.сподіванням кількості заявок, який обслуговує канал в одиницю часу, якщо він працює без перерви.

Тоді дріб видає мат.сподівання кількості зайнятих каналів.

### 5.11.3 Одноканальна СМО з обмеженою чергою

Маємо канал, вхідний потік, довжина черги  $m$  – залишити систему необслугованою. Стани системи:

$A_0$  - канал вільний

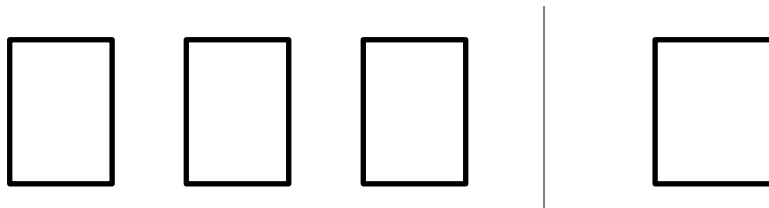
$A_1$  - канал зайнятий, довжина черги 0

$A_2$  - канал зайнятий, довжина черги 1

...

$A_m$  - канал зайнятий, довжина черги  $m$

Показати самим, що в системі ф-нує однорідний регулярний харківський процес з графом переходу



$$\rho = \frac{\lambda}{\mu}$$

$$P_1 = \rho P_0, P_2 = \rho^2 P_0, \dots, P_{m+1} = \rho^{m+1} P_0$$

$$P_0 = \frac{1}{1 + \rho + \rho^2 + \dots + \rho^{m+1}} = \frac{1 - \rho}{1 - \rho^{m+2}} \quad (\text{використовуємо формулу суми членів}$$

геометричної прогресії)

$$q = 1 - P_{m+1}$$

$$A = \lambda q$$

Знайдемо мат. сподівання довжини черги в стаціонарному режимі

$$MR = 1 \cdot P_2 + \dots + m P_{m+1} = \underbrace{\rho^2 P_0}_{P_2} (1 + 2\rho + 3\rho^2 + \dots + m\rho^{m-1}) = \rho^2 P_0 \frac{d(\rho + \rho^2 + \dots + \rho^m)}{\frac{d}{d\rho}(\rho + \rho^2 + \dots + \rho^m)}$$

$$\left( \frac{d}{d\rho} \left( \frac{\rho - \rho^{m+1}}{1 - \rho} \right) = \frac{(1 - (m+1)\rho^m)(1 - \rho) + \rho - \rho^{m+1}}{(1 - \rho)^2} = \frac{1 - \rho^m(m+1 - m\rho)}{(1 - \rho)^2} \right)$$

$$\rho^2 \frac{1 - \rho^m(m+1 - m\rho)}{(1 - \rho)(1 - \rho^{m+2})}$$

Знайдемо мат.сподівання часу перебування в черзі:

$$MT = \frac{1}{\mu} \cdot P_1 + \frac{2}{\mu} \cdot P_2 + \dots + \frac{m}{\mu} \cdot P_m$$

Пояснення: з ймовірністю  $P_1$  в стаціонарному режимі заявка застає канал вільним. Тоді за І теоремою можна вважати, що в цей момент часу

канал почав обслуговувати заявку і мат.сподівання  $\frac{1}{\mu}$  - мат.сподівання випадкової величини, що має експоненційний розподіл з параметром  $\mu$ .

З ймовірністю  $P_2$  заявка застає 1 канал зайнятий і одну заявку в черзі. За І теоремою можна вважати, що канал почав тільки обслуговувати заявку. Тоді мат.сподівання перебування в черзі – мат.сподівання суми двох випадкових величин:

1. Час обслуговування заявки каналом

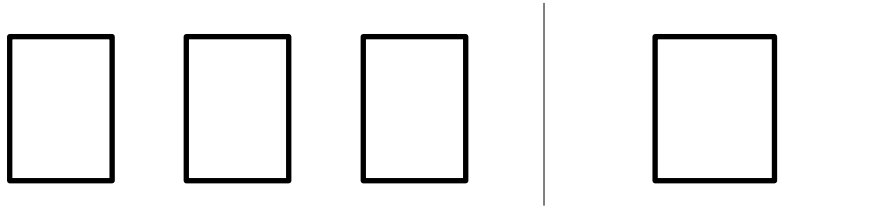
2. Час обслуговування заявки каналом, що перебуває першою в черзі.

Аналогічно пояснюються інші члени виразу для  $MT$ .

Остаточного отримуємо:

$$MT = \frac{\rho(1 - \rho^m(m+1 - m\rho))}{\mu(1 - \rho)(1 - \rho^{m+1})}.$$

#### 5.11.4 Одноканальна СМО з необмеженою чергою



$$P_i = \rho^i P_0, i=1, \dots, \infty$$

$$P_0 = \frac{1}{1 + \rho + \rho^2 + \rho^3 + \dots}$$

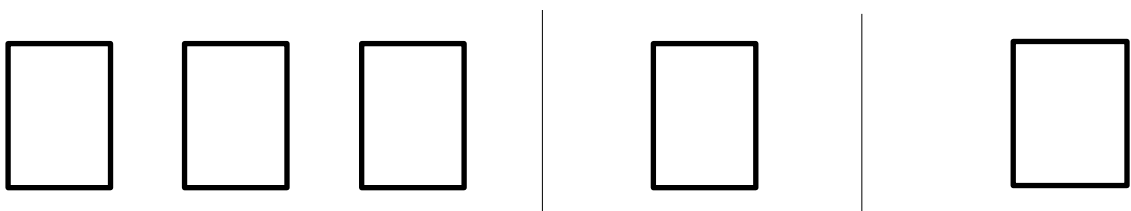
Будемо розглядати випадок, коли мат. сподівання черги обмежено (не є нескінченним). Умовою цього є  $\rho < 1$ .

Якщо  $\rho \geq 1$ , то  $P_0 = 0$ , а це означає, що в стаціонарному режимі з ймовірністю 1 довжина черги – нескінченність, а тому і мат. сподівання довжини черги дорівнює  $\infty$ .

Граничні характеристики для випадку  $\rho < 1$  отримати самостійно.

#### 5.11.5 Багатоканальна СМО з обмеженою чергою

В системі функціонує однорідний регулярний марківський процес з наступним графом переходів:



$A_n$  каналів, довжина черги  $m$

$\mu$  - параметр експоненційного розподілу часу обслуговування заявки,

$\alpha$  – параметр пуасонівського потоку заявок на обслуговування.

$$\rho = \frac{\lambda}{\mu}, \quad P_1 = \frac{\rho}{1!} P_0, \quad P_n = \frac{\rho^n}{n!} P_0, \dots, P_{n+r} = \frac{\rho^{n+r}}{n! n^r} P_0, \dots, P_{n+m} = \frac{\rho^{n+m}}{n! n^m} P_0$$

$$P_0 = \frac{1}{1 + \frac{\rho}{1!} + \dots + \frac{\rho^n}{n!} \left( \frac{\rho}{n} + \frac{\rho^2}{n^2} + \dots + \frac{\rho^m}{n^m} \right)}$$

$$q = 1 - P_{m+n}$$

$$A = \lambda q = \lambda (1 - P_{n+m})$$

$$MZ = \frac{A}{\mu} - \text{мат. сподівання числа зайнятих каналів.}$$

Знайдемо мат. сподівання числа довжини черги

$$MR = 1 \cdot P_{n+1} + 2 \cdot P_{n+2} + \dots + m \cdot P_{n+m} = P_{n+1} \left( 1 + 2 \left( \frac{\rho}{n} \right) + 3 \left( \frac{\rho}{n} \right)^2 + \dots + m \left( \frac{\rho}{n} \right)^{m-1} \right)$$

Позначимо  $\frac{\rho}{n} = \chi$  :

$$MR = P_{n+1} \frac{1 - \chi^m (m+1 - m\chi)}{(1 - \chi)^2}$$

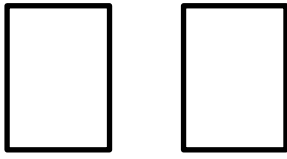
Примітка: якщо в даний момент часу всі канали стали обслуговувати заявки, то мат.сподівання часу до появи першої обслугов.заявки дорівнює

$$\frac{1}{n\mu}.$$

Пояснення: тому що можна вважати, що n-каналів генерують сумарний віртуальний потік обслуговуваних заявок з параметром  $n\mu$ , де за час, що має експансійний розподіл з параметром  $n\mu$  реалізується перша заявка.

$$\begin{aligned} MT &= \frac{1}{n\mu} P_n + \frac{2}{n\mu} P_{n+1} + \dots + \frac{m}{n\mu} P_{n+m-1} = \\ &= \frac{1}{n\mu} P_n \left( 1 + 2 \left( \frac{\rho}{n} \right) + 3 \left( \frac{\rho}{n} \right)^2 + \dots + m \left( \frac{\rho}{n} \right)^{m-1} \right) = \\ &= \frac{1}{n\mu} P_n \frac{1 - \chi^m (m+1 - m\chi)}{(1 - \chi)^2} \end{aligned}$$

#### 5.11.6 Багатоканальна СМО з необмеженою чергою



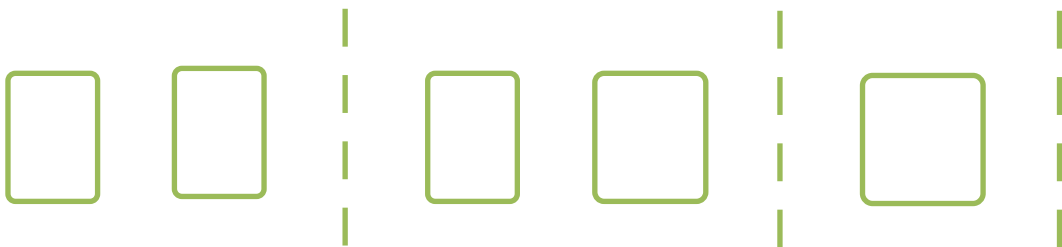
$$P_0 = \frac{1}{1 + \frac{\rho}{1!} + \dots + \frac{\rho^n}{n!} \left( \frac{\rho}{n} + \frac{\rho^2}{n^2} + \dots + \frac{\rho^m}{n^m} \right)}$$

Умовою існування обмеженого мат.сподівання є  $\frac{\rho}{n} < 1$ .

Характеристики  $MT, MR$  знайти самостійно, для випадку  $\frac{\rho}{n} < 1$

#### 5.11.7 Багатоканальна СМО з обмеженим часом перебування заявки в черзі

Багатоканальна СМО з обмеженим часом перебування заявки в черзі працює як і багатоканальна СМО з необмеженою чергою з додаванням умови: час перебування будь-якої заявки в черзі має експоненційний розподіл за параметром  $\nu$  і заявки залишають чергу не залежно одна від одної.



$$P_1 = \frac{\rho}{1!} P_0$$

$$P_n = \frac{\rho^n}{n!} P_0$$

$$P_{n+1} = \frac{\rho^n \lambda}{n!(n\mu + \nu)} P_0$$

$$P_{n+r} = \frac{\rho^n \lambda^r}{n!(n\mu + \nu)(n\mu + 2\nu) \dots (n\mu + r\nu)} P_0$$

$$P_0 = \frac{1}{1 + \frac{\rho}{1!} + \dots + \frac{\rho^n}{n!} \left( 1 + \frac{\lambda}{n\mu + \nu} + \dots + \frac{\lambda^r}{(n\mu + \nu)(n\mu + 2\nu) \dots (n\mu + r\nu)} \right) + \dots}$$

$P_0 > 0$  при будь-яких  $\lambda, \nu, \mu > 0$ .

Знаменник для  $P_0$  при будь-яких  $\lambda, \nu, \mu$  є обмеженим числом (це результат строго доводиться). Проведемо якісне пояснення цього результату.

Знайдемо  $A$  – абсолютну пропускну спроможність системи. Можна строго довести, що  $A = \lambda - MR \cdot \nu$ , де  $MR$  – математичне сподівання довжини черги. Дійсно,  $\lambda$  – математичне сподівання кількості заявок, що поступають за одиницю часу, а математичне сподівання довжини черги можна інтерпретувати, як стала черга на обслуговування заявок з довжиною  $MR$ , кожна з яких генерує потік уходу заявок з математичним сподіванням  $\nu$ . Тоді  $MR \cdot \nu$  – математичне сподівання загальної кількості заявок, що залишили систему в одиницю часу на обслуговування.

Математичне сподівання числа зайнятих каналів:

$$MZ = 1 P_1 + 2 P_2 + \dots + n \left( 1 - \sum_{i=0}^{n-1} P_i \right)$$

$$MZ = \frac{A}{\mu}$$

Для знаходження  $MR$  можна спочатку знайти  $MZ$ , потім знайти  $A = MZ \cdot \mu$ ,

і з формули  $A = \lambda - MR \cdot \nu$  отримати, що  $MR = \frac{\lambda - A}{\nu}$ .

#### 5.11.8 Замкнена СМО

В більшості реальних задач параметр вхідного потоку  $\lambda$  не залежить від стану системи. Такі СМО звуться відкритими. Але в деяких прикладних задачах параметр вхідного потоку принципово залежить від стану, в якому знаходиться система. Такі СМО звуться закритими.

### *Приклад*

Маємо  $n$  верстатів, які незалежно один від одного виходять з ладу. Час роботи кожного верстата до виходу з ладу має експоненціальний розподіл з параметром  $\lambda$ . Існує робітник, який лагодить верстат. Час ремонту має експоненціальний розподіл з параметром  $\mu$ .

Схема роботи СМО наступна:

- 1) Якщо станок вийшов з ладу і робітник вільний, станок стає на обслуговування (ремонтується).
- 2) Якщо станок вийшов з ладу і робітник зайнятий, станок стає в чергу на обслуговування (ремонт).

Показати самим, що в системі функціонує регулярний однорідний марківський процес з наступним графом переходів:



Стани системи:

$A_0$  – всі верстати працюють;

$A_1$  – один верстат несправний, черга дорівнює нулю;

...

$A_n$  – всі верстати несправні, черга дорівнює  $n-1$



$$\rho = \frac{\lambda}{\mu}$$

Граничні ймовірності:

$$p_1 = n\rho p_0$$

$$p_2 = n(n-1)\rho^2 p_0$$

...

$$p_n = n! \rho^n p_0$$

$$p_0 = \frac{1}{1 + n\rho + n(n-1)\rho^2 + \dots + n! \rho^n}$$

Математичне сподівання довжини черги:

$$MR = 1 p_2 + 2 p_3 + \dots + (n-1) p_n$$

## 5.12 Метод псевдостанів

Якщо вхідний потік заявок не є пуассонівським, а розподіл часу обслуговування заявки не є експоненціальним, в системі не функціонує регулярний однорідний марківський процес. Існує лише дуже обмежена кількість випадків, в яких аналітично знаходяться характеристики ефективності немарківських СМО.

Універсальним засобом аналізу загальних СМО є імітаційне моделювання, де використовуються відповідні спеціалізовані мови програмування. Імітуються роботи СМО з будь-якими вхідними потоками, законами обслуговування заявок, формування черги. Відповідні характеристики роботи системи знаходяться статистичними методами як середнє значення на достатньо довгому відрізку часу моделювання СМО в стаціонарному режимі.

Недоліки імітаційного моделювання:

- великі витрати інформаційних ресурсів
- будь-які зміни будь-якого параметру системи призводять до необхідності повністю перемодельовувати її роботу

Тому так високо ціняться будь-які способи аналітичного знаходження характеристик систем.

Один з таких аналітичних методів – є метод псевдостанів. Використовується лише, коли вхідний потік заявок на обслуговування є стаціонарним. Тоді він апроксимується потоком Ерланга відповідного порядку.

Нехай час обслуговування заявки не є експоненціальним, тоді він апроксимується (замінюється) розподілом Ерланга з відповідними параметрами (під розподілом Ерланга розуміється розподіл інтервалів часу між сусідніми подіями потоку Ерланга відповідного порядку).

Вводяться додаткові стани системи (реально неіснуючі), які формально немарківську СМО переводять в марківську, які і досліджується відомими аналітичними методами.

### *Приклад*

#### Постановка задачі

Вхідний потік заявок є пуассонівським з параметром  $\lambda$ . Одноканальна СМО. Нескінченна черга.

#### Рішення

Час обслуговування заявки має розподіл Ерланга другого порядку з параметром  $\Lambda = \mu$ . Тоді час обслуговування заявки каналом  $T$  є

$$T = T_1 + T_2,$$

де  $T_i, i=1,2$  – мають експоненціальний розподіл з параметром  $2\mu$ .

Вводимо стани системи:

$A_0$  – канал вільний, черги немає (реальний);

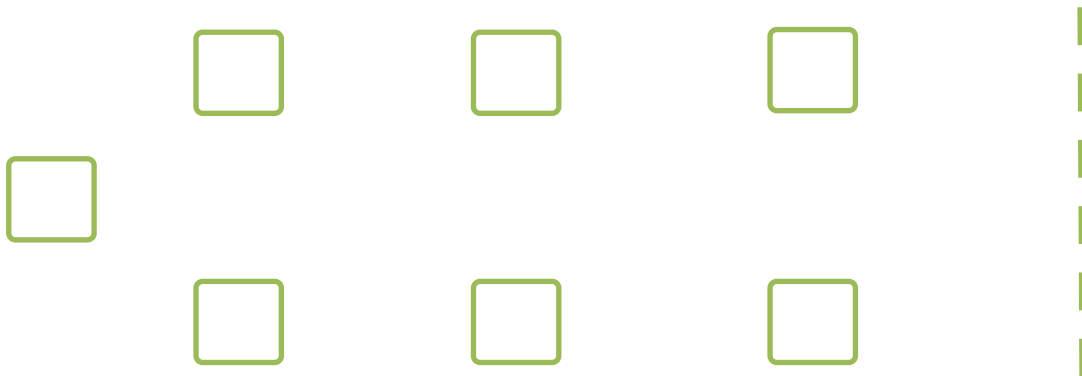
$A_{11}$  – канал зайнятий (обслуговує заявку в першій фазі -  $T_1$ ), черги немає.

$A_{12}$  - канал зайнятий (обслуговує заявку в другій фазі -  $T_1$ ), черги немає.

$A_{k1}$  – канал зайнятий (обслуговує заявку в першій фазі -  $T_1$ ), довжина черги дорівнює  $k-1$ .

$A_{k2}$  - канал зайнятий (обслуговує заявку в другій фазі -  $T_1$ ), довжина черги дорівнює  $k-1$ .

Показати, що в системі функціонує однорідний регулярний марківський процес з наступним графом переходів:



Цей граф задовольняє умовам граничної теореми.

Знаходимо граничні ймовірності:

$$\lambda p_0 = 2\mu p_{12}$$

$$p_{11}(\lambda + 2\mu) = p_0\lambda + p_{22}2\mu$$

$$p_{12}(\lambda + 2\mu) = p_{11}2\mu$$

$$p_{21}(\lambda + 2\mu) = p_{11}\lambda + p_{32}2\mu$$

$$p_{22}(\lambda + 2\mu) = p_{12}\lambda + p_{21}2\mu$$

...

$$p_{k1}(\lambda + 2\mu) = p_{(k-1)1}\lambda + p_{(k+1)2}2\mu$$

$$p_{k2}(\lambda + 2\mu) = p_{(k-1)2}\lambda + p_{k1}2\mu$$

...

Отримали нескінченну СЛАР зі сталими коефіцієнтами. Емпіричний спосіб її розв'язку: з загальних міркувань обмежується числом довжина черги. Отримуємо обмежену СЛАР зі сталими коефіцієнтами, яку розв'язуємо чисельними методами.

### 5.13 Випадкові процеси(стохастичні процеси)

Маємо ймовірнісний простір  $(\Omega, \sigma, P)$ .

Випадковим процесом зветься числова скалярна функція двох аргументів  $\varphi(\omega, t)$ , де  $\omega \in \Omega$ , а  $t$  – дійсний аргумент (в прикладних задачах інтерпретується як час).

При фіксованому значенні  $\omega_\phi$  функція  $\varphi(\omega_\phi, t)$  є дійсною функцією дійсного аргументу  $t$  і зветься реалізацією випадкового процесу в наслідок випробування в якому настала елементарна подія  $\omega_\phi$ .

Таким чином випадковий процес переводить  $\Omega$  в множину всіх дійсних функцій і є узагальненням випадкової величини: а саме реалізацією випадкової величини в випробуванні є число  $\varphi(\omega_\phi)$ , якщо в випробуванні настала елементарна подія  $\omega_\phi$ .

При фіксованому часі  $t_\phi$  випадковий процес  $\varphi(\omega, t)$  перетворюється в випадкову величину  $\varphi(\omega, t_\phi)$  вимірну відносно  $\sigma$ -алгебри.

Ці функції в інженерній практиці ніхто не будує. Вони використовуються для загальної теорії ймовірності, випадкова величина і випадковий процес задані на одному ймовірнісному просторі  $(\Omega, \sigma, P)$

При записі випадкового процесу будемо писати  $\varphi(t), X(t), Y(t), Z(t)$ , розуміючи, що при фіксованому часі  $t_\phi$  будемо мати випадкову величину.

Прикладом випадкових процесів є вивчений нами регулярний однорідний Марковський процес.

З означення випадкового процесу випливає існування наступної функції. Фіксуємо довільне натуральне число  $n$  і дійсні числа  $x_1, \dots, x_n$  та

$t_1, \dots, t_n$ . Розглянемо функцію  $F(x_1, \dots, x_n, t_1, \dots, t_n) = P(\omega: \varphi(\omega, t_1) < x_1, \dots, \varphi(\omega, t_n) < x_n)$

.

Ця функція існує.

Дійсно, розглянемо події  $A_i = \{\omega: \varphi(\omega, t_i) < x_i\}$ ,  $i = \overline{1, n}$  подію  $A = \{\omega: \varphi(\omega, t_1) < x_1, \dots, \varphi(\omega, t_n) < x_n\}$ .

$$A = \bigcap_{i=1}^n A_i$$

$A_i \in \sigma$ , тоді за означенням  $\sigma$ -алгебри  $A \in \sigma$ .

Значить для будь-яких  $x_1, \dots, x_n$  та  $t_1, \dots, t_n$  існує  $P(A)$ , а значить функція  $F(x_1, \dots, x_n, t_1, \dots, t_n)$  існує для будь-яких значень своїх аргументів і зветься багатовимірною функцією розподілу випадкового процесу. Ця функція обов'язково задовольняє наступним двом властивостям:

1) Для будь-яких натуральних  $m < n$  виконується

$$F(x_1, \dots, x_m, \infty, \dots, \infty, t_1, \dots, t_n) = F(x_1, \dots, x_m, t_1, \dots, t_m)$$

2) Беремо довільну перестановку чисел  $1, \dots, n$ , як  $i_1, \dots, i_n$ , тоді

$$F(x_1, \dots, x_n, t_1, \dots, t_n) = F(x_{i_1}, \dots, x_{i_n}, t_{i_1}, \dots, t_{i_n})$$

Обґрунтувати самим.

Випадковий процес називається стаціонарним у вузькому розумінні слова, якщо  $\forall n \in \mathbb{N}$  і будь-яких дійсних  $u$ ,  $x_1, \dots, x_n, t_1, \dots, t_n$  виконується рівність

$$F(x_1, \dots, x_n, t_1, \dots, t_n) = F(x_1, \dots, x_n, t_1 + u, \dots, t_n + u)$$

Еквівалентне означення: випадкові величини  $\varphi(t_1), \dots, \varphi(t_n)$  та

$\varphi(t_1 + u), \dots, \varphi(t_n + u)$  мають однакову  $n$ -мірну функцію розподілу.

Властивості стаціонарного в вузькому розумінні слова випадкового процесу

Нехай розглядається випадковий процес  $\varphi(t)$ , тоді  $M\varphi(t)$  - детермінована числова скалярна функція дійсного аргументу  $t$ , яка при фіксованому  $t_\phi \in \mathbb{R}$  є математичним сподіванням випадкової величини  $\varphi(t_\phi)$ .

Так само задається і дисперсія випадкового процесу  $D\varphi(t)$ , яка при фіксованому  $t_\phi$  дорівнює дисперсії випадкової величини  $\varphi(t_\phi)$ .

Математичне сподівання та дисперсія для стаціонарного випадкового процесу в вузькому розумінні слова мають властивість:

$$M\varphi(t) = \text{const}_1$$

$$D\varphi(t) = \text{const}_2$$

*Доведення:*

Розглянемо два довільних числа  $t_1, t_2$  і  $\varphi(t_1), \varphi(t_2)$ .

$$t_2 = t_1 + (t_2 - t_1)$$

Маємо функції розподілу  $\varphi(t_1)$  і  $\varphi(t_1 + (t_2 - t_1))$ .

З означення стаціонарності у вузькому розумінні слова, взявши  $n=1$  та  $u=t_2-t_1$  отримаємо рівність розподілів  $\varphi(t_1)$  та  $\varphi(t_2)$ , а отже і рівність математичних сподівань і дисперсій для різних значень  $t$ . ■

Функція коваріації випадкового процесу  $\varphi(t)$  зветься детермінована числова скалярна функція двох дійсних аргументів, яка при фіксованих значеннях двох своїх аргументів дорівнює:

$$B(t_1, t_2) = \text{cov}(\varphi(t_1), \varphi(t_2))$$

Функція коваріації для стаціонарного випадкового процесу має наступну властивість:

$$B(t_1, t_2) = B(0, t_2 - t_1) = B(u), u = t_2 - t_1$$

*Доведення:*

Розглянемо два довільних числа  $t_1, t_2$  і  $\varphi(t_1), \varphi(t_2)$ .

$$t_2 = t_1 + (t_2 - t_1)$$

Маємо функції розподілу  $\varphi(t_1)$  і  $\varphi(t_1 + (t_2 - t_1))$ .

З умови стаціонарності випадкового процесу, взявши  $n=2$  та  $u=-t_1$  отримаємо рівність двох двовимірних розподілів  $\varphi(t_1), \varphi(t_2)$  та

$\varphi(0), \varphi(t_2 - t_1)$ . З цього випливає рівність коефіцієнтів коваріації для цих двох двовимірних розподілів. ■

Випадковий процес зветься стаціонарним у широкому розумінні слова, якщо у нього математичне сподівання та дисперсія сталі, а функція коваріації є функцією від одного аргументу  $u = t_2 - t_1$ .

*Наслідок:*

Стаціонарний процес у вузькому розумінні слова є одночасно і стаціонарним процесом у широкому розумінні слова. Зворотне не є вірним.

Стаціонарний процес у широкому розумінні слова (далі просто стаціонарний процес) вводиться з таких причин:

1. Стаціонарні випадкові процеси мають цікаві властивості, які ми розглянемо далі.

2. Стаціонарність в широкому розумінні слова можна перевірити статистичними методами, що дозволяє оцінити математичне сподівання, дисперсію та функцію коваріації.

Випадковий процес зветься неперервним в середньоквадратичному, якщо для будь-якого  $t$  існує

$$\lim_{\Delta t \rightarrow 0} M(\varphi(t + \Delta t) - \varphi(t))^2 = 0.$$

Використовуючи нерівність Чебишева, отримаємо:

$$\exists \lim_{\Delta t \rightarrow 0} P(|\varphi(t + \Delta t) - \varphi(t)| < \varepsilon) = 0, \forall \varepsilon > 0$$

З цього виразу можна б було зробити висновок, що у неперервному в середньоквадратичному випадковому процесі всі його реалізації з ймовірністю 1 є неперервними функціями, але це не вірно. Дійсно, розглянемо регулярний однорідний марківський процес

$$P_{ij}(\Delta t) = \lambda_{ij} \Delta t + o(\Delta t)$$

З цього випливає, що для будь-якого  $t$  ймовірності переходу з одного стану в інший дорівнюють нулю, але реалізація його не є неперервною функцією.

Властивості функції коваріації для неперервного у середньоквадратичному випадкового процесу:

1. З означення маємо:

$$B(t_1, t_2) = B(t_2 - t_1) = B(\varphi(0), \varphi(t_2 - t_1))$$

2. Нехай  $u = t_2 - t_1$ , тоді маємо:

$$B(u) = B(-u)$$

$$3. B(0) \geq B(u)$$

$$B(0) = \text{cov}(\varphi(t), \varphi(t)) = D\varphi(t), \forall t$$

*Доведення:*

$$M[(\xi(0) - M\xi(0)) \pm (\xi(U) - M\xi(U))]^2 \geq 0$$

$$2B(0) \pm 2B(u) \geq 0$$

$$B(0) \geq B(u)$$

4. Якщо стаціонарний процес неперервний у середньоквадратичному, то функція коваріації центрованого випадкового процесу дорівнює функції коваріації не центрованого випадкового процесу:

$$B_{\varphi(t) - M\varphi(t)}(u) = B_{\varphi(t)}(u)$$

5. Функція коваріації  $B(u)$  є неперервною функцією. Для спрощення вважаємо, що  $M\varphi(t) = 0$ .

*Доведення:*

$$|B(u + \Delta u) - B(u)| = |M(\varphi(0)\varphi(u + \Delta u)) - M(\varphi(0)\varphi(u))| = |M(\varphi(0)(\varphi(u + \Delta u) - \varphi(u)))|$$

Вважаємо, що  $\varphi(0)$  та  $\eta = \varphi(u + \Delta u) - \varphi(u)$  абсолютно неперервні і мають двовимірні функції щільності (для 90% стаціонарних неперервних у



середньоквадратичному випадкових процесів  $\varphi(0)$  та  $\eta$  абсолютно неперервні).

Нехай  $f(x, y)$  двовимірною функцією щільності  $\varphi(0)$ ,  $\eta$ . Використаємо нерівність Коші-Буняковського-Шварца:

$$\left| \iint_{-\infty}^{\infty} g_1(x, y) g_2(x, y) dx dy \right| \leq \sqrt{\iint_{-\infty}^{\infty} g_1^2(x, y) dx dy \cdot \iint_{-\infty}^{\infty} g_2^2(x, y) dx dy}$$

$$\left| M(\varphi(0)(\varphi(u+\Delta u) - \varphi(u))) \right| = \left| \iint_{-\infty}^{\infty} xy \cdot f(x, y) dx dy \right|$$

В нашому випадку візьмемо  $g_1(x, y) = x \sqrt{f(x, y)}$ , а  $g_2(x, y) = y \sqrt{f(x, y)}$

$$\left| \iint_{-\infty}^{\infty} xy \cdot f(x, y) dx dy \right| \leq \sqrt{\iint_{-\infty}^{\infty} x^2 f(x, y) dx dy \cdot \iint_{-\infty}^{\infty} y^2 f(x, y) dx dy} = \sqrt{M\varphi^2(0) M\eta^2} = \sqrt{M\varphi^2(0) M(\varphi(u+\Delta u) - \varphi(u))^2}$$

Отже

$$\left| B(u+\Delta u) - B(u) \right| \leq \sqrt{M\varphi^2(0) M(\varphi(u+\Delta u) - \varphi(u))^2}$$

$$\lim_{\Delta u \rightarrow 0} \left| B(u+\Delta u) - B(u) \right| \leq \lim_{\Delta u \rightarrow 0} \sqrt{M\varphi^2(0) M(\varphi(u+\Delta u) - \varphi(u))^2} = 0$$

$$\lim_{\Delta u \rightarrow 0} \left| B(u+\Delta u) - B(u) \right| = 0$$

6. Емпірична (інженерна) властивість

$$\lim_{u \rightarrow \infty} B(u) = 0$$

Якщо різниця  $u = t_2 - t_1$  необмежено велика, то вважають, що випадкові величини  $\varphi(t_1)$  та  $\varphi(t_2)$  незалежні і  $\text{cov}(\varphi(t_1), \varphi(t_2)) = 0$ .

7. В інженерній практиці функцію коваріації шукають в одному з чотирьох виглядів:

$$B(u) = \begin{cases} \sigma^2 e^{-\alpha|u|} \\ \sigma^2 e^{-\alpha|u|} \cos \beta u, \alpha, \beta > 0 \\ \sigma^2 e^{-\alpha u^2} \\ \sigma^2 e^{-\alpha u^2} \cos \beta u \end{cases}$$

Далі буде показано, що якщо випадковий процес має функцію (перший чи другий вираз), то він не має похідної, а якщо третій чи четвертий вираз – то має похідну.

#### 5.14 Стохастичний інтеграл від випадкового процесу

Розглянемо числовий інтервал  $(a, b)$  на якому задана довільна інтегрована детермінована функція  $f(t)$  та стаціонарний випадковий процес  $\varphi(t)$ . Відрізок  $[a, b]$  поділимо на  $n$  частин довжиною  $\Delta t_i = \frac{b-a}{n}$ .  $t_i$  — лівий кінець  $i$ -того відрізка.

$$I_n = \sum_{i=1}^n f(t_i) \varphi(t_i) \Delta t_i$$

Послідовність випадкових величин  $I_n$  при  $n \rightarrow \infty$  має границю  $I$  в середньоквадратичному, якщо існує:

$$\lim_{n \rightarrow \infty} M(I_n - I)^2 = 0$$

Використовуючи нерівність Чебишева можна показати, що з цього випливає збіжність послідовності випадкових величин  $I_n$  до випадкової величини  $I$  по ймовірності:

$$\lim_{n \rightarrow \infty} P(|I_n - I| > \varepsilon) = 0, \forall \varepsilon > 0$$

Якщо границя існує, то вона зветься стохастичним інтегралом і має запис:

$$I = \int_a^b f(t) \varphi(t) dt$$

Існує альтернативне означення стохастичного інтегралу, а саме:  $I$  існує, якщо існує  $\forall m \neq n \lim_{n, m \rightarrow \infty} M(I_n - I_m)^2 = 0$ , а з нерівності Чебишева маємо:

$$\exists \lim_{n, m \rightarrow \infty} P(|I_n - I_m| > \varepsilon) = 0, \forall \varepsilon > 0$$

Має місце наступна теорема:

Стохастичний інтеграл існує, якщо існує звичайний двократний інтеграл

$$R = \int_a^b \int_a^b f(t) f(s) B_{\varphi(t)}(t-s) dt ds$$

Доведення:

Будемо вважати, що  $M\varphi(t)=0$ . Покажемо, що це не обмежує доведення теореми.

$$B_{\varphi(t)-M\varphi(t)}(u)=B_{\varphi(t)}(u)$$

$$\text{Нехай існує } \int_a^b f(t)(\varphi(t)-M\varphi(t))dt = \int_a^b f(t)\varphi(t)dt - \int_a^b f(t)M\varphi(t)dt$$

Якщо  $I = \int_a^b f(t)(\varphi(t)-M\varphi(t))dt$  існує, то його можна представити у вигляді

$$I_1 = \int_a^b f(t)\varphi(t)dt \quad \text{та} \quad I_2 = \int_a^b f(t)M\varphi(t)dt \quad (\text{що існує}).$$

Тоді, якщо ми доведемо, що при  $n \rightarrow \infty$  інтеграл  $I$  існує, тоді ми доведемо що існує  $I_1 = I + I_2$ .

$$M(I_n - I_m)^2 = M\left(\sum_{i=1}^n f(t_i)\varphi(t_i)\Delta t_i - \sum_{j=1}^m f(s_j)\varphi(s_j)\Delta s_j\right)^2 = M\left(\sum_{i=1}^n \sum_{j=1}^n f(t_i)f(t_j)\varphi(t_i)\varphi(t_j)\Delta t_i\Delta t_j - 2\sum_{i=1}^n \sum_{j=1}^m f(t_i)f(s_j)\varphi(t_i)\varphi(s_j)\Delta t_i\Delta s_j + \sum_{j=1}^m \sum_{k=1}^m f(s_j)f(s_k)\varphi(s_j)\varphi(s_k)\Delta s_j\Delta s_k\right)$$

$$M(\varphi(t_i)\varphi(t_j)) = B_{\varphi(t)}(t_i - t_j)$$

$$M(\varphi(t_i)\varphi(s_j)) = B_{\varphi(t)}(t_i - s_j)$$

$$M(\varphi(s_i)\varphi(s_j)) = B_{\varphi(t)}(s_i - s_j)$$

Оскільки існує інтеграл  $R$ , тоді маємо:

$$\lim_{n,m \rightarrow \infty} M(I_n - I_m)^2 = R - 2R + R = 0$$

Якщо стохастичний інтеграл існує, то в інженерних задачах для знаходження реалізації випадкової величини  $I$  треба в звичайний інтеграл

$$\int_a^b f(t)\varphi(t)dt$$

підставити реалізацію випадкового процесу  $\varphi(t)$  на відрізку  $[a, b]$  і чисельними методами знайти його.

Ергодична теорема Біргофа-Хінчіна

Нехай  $\varphi(t)$  – стаціонарний у вузькому розумінні слова, тоді:

$$1. \exists \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \varphi(t)dt = M\varphi(t) = \text{const}$$

$$2. \quad \forall u \exists \lim_{T \rightarrow \infty} \frac{1}{T-u} \int_0^{T-u} \varphi(t) \varphi(t+u) dt = B_{\varphi(t)}(u)$$

Інженерна цінність теореми в тому, що математичне сподівання, дисперсія і функцію коваріації стаціонарного випадкового процесу у вузькому розумінні слова можна одержати лише маючи одну його довгу реалізацію.

Означення. Випадковий процес  $\varphi(t)$  є стохастичним інтегралом від випадкового процесу  $X(t)$  якщо  $\varphi(t) = \int_0^t X(u) du$

Довести самим, що випадковий процес  $\varphi(t)$  існує, якщо існує звичайний інтеграл

$$\int_0^t \int_0^t B_{\varphi(t)}(u_1, u_2) du_1 du_2$$

Примітка:

Існує інтеграл  $\int_0^t M\varphi(u) du$  та  $B_{\varphi(t)}(t_1, t_2) = B_{X(t)-MX(t)}(t_1, t_2)$

Знайдемо математичне сподівання і функцію коваріації випадкового процесу  $Y(t)$ , що є стохастичним інтегралом від випадкового процесу  $X(t)$ .

Розіб'ємо відрізок  $[0, t]$  на відрізки довжиною  $\Delta t$ :

$$MY(t) = M \lim_{\Delta t \rightarrow 0} \sum_{i=1}^n X(t_i) \Delta t_i$$

Доводиться, що якщо існує границя в середньоквадратичному інтегральній суми, то операцію математичного сподівання і границі можна поміняти місцями:

$$\lim_{\Delta t \rightarrow 0} \sum_{i=1}^n MX(t_i) \Delta t_i = \int_0^t MX(u) du$$

Знайдемо функцію коваріації. За означенням:

$$\bar{Y}(t) = Y(t) - MY(t)$$

$$\bar{X}(t) = X(t) - MX(t)$$

Можна показати, що :

$$\bar{Y}(t) = \int_0^t \bar{X}(u) du$$

Тоді:

$$B_{Y(t)}(t_1, t_2) = M(\bar{Y}(t_1)\bar{Y}(t_2)) = M\left(\int_0^{t_1} \bar{X}(u) du \int_0^{t_2} \bar{X}(s) ds\right) = M\left(\int_0^{t_1} \int_0^{t_2} \bar{X}(u)\bar{X}(s) duds\right)$$

Оскільки існує границя в середньоквадратичному для  $\int_0^{t_1} \bar{X}(u) du$  і  $\int_0^{t_2} \bar{X}(s) ds$

, то для  $\int_0^{t_1} \int_0^{t_2} \bar{X}(u)\bar{X}(s) duds$  теж існує границя в середньоквадратичному. Отже операцію інтегрування і математичного сподівання можемо поміняти місцями:

$$\int_0^{t_1} \int_0^{t_2} M(\bar{X}(u)\bar{X}(s)) duds = \int_0^{t_1} \int_0^{t_2} B_{X(t)}(u, s) duds$$

Отримали, що функція коваріації інтеграла випадкового процесу дорівнює подвійному інтегралу від функції коваріації цього випадкового процесу.

Примітка:

Нехай  $X(t)$  – стаціонарний випадковий процес. В нашому випадку вираз для  $MY(t)$  не змінився, а для функції коваріації отримали б:

$$B_{Y(t)}(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} B_{X(t)}(u-s) du ds$$

Показати самим, що інтеграл від стаціонарного процесу не є стаціонарним процесом.

## 5.15 Похідна від випадкового процесу

Розглянемо довільний випадковий процес  $X(t)$  і випадкову величину  $I_{\Delta_1}$

$$I_{\Delta_1} = \frac{X(t + \Delta t) - X(t)}{\Delta_1}, \quad \Delta_1 \rightarrow 0$$

По аналогії з стохастичним інтегралом: якщо  $\exists \lim_{\substack{\forall \Delta_1 \rightarrow 0 \\ \Delta_2 \rightarrow 0}} (I_{\Delta_1} - I_{\Delta_2})^2 = 0$ , то

існує границя в середньому квадратичному для  $I_{\Delta}$  (послідовність випадкових величин) при  $\Delta \rightarrow 0$ , яка зветься похідною від випадкового

процесу  $\frac{dX(t)}{dt}$ .

Примітка 1. З нерівності Чебишева випливає

$$\lim_{\substack{\Delta_1 \rightarrow 0 \\ \Delta_2 \rightarrow 0}} P(|I_{\Delta_1} - I_{\Delta_2}| > \varepsilon) = 0, \quad \forall \varepsilon > 0.$$

Примітка 2. Якщо існує звичайна похідна  $\frac{dMX(t)}{dt}$ , то при доведенні умови існування похідної від випадкового процесу можна покласти  $MX(t) = 0$ .

Знайдемо умову існування похідної від випадкового процесу.

Доведення самостійно.

1) Використати той самий спосіб, як доведення умови існування інтегралу (частину, де поклали  $MX(t) = 0$ ).

2)  $B_{X(t)}(t_1, t_1) = B_{X(t) - MX(t)}(t_1, t_1)$  - функція коваріації центрованого випадкового процесу.

$$\frac{d(X(t) - MX(t))}{dt} = \frac{dX(t)}{dt} - \frac{dMX(t)}{dt} \text{ - і дов.існує.}$$

Нагадаємо формулу:

$$\frac{\partial^2 f(x, y)}{\partial x \partial y} = \lim_{\substack{\forall \Delta_1 \rightarrow 0 \\ \Delta_2 \rightarrow 0}} \frac{f(x + \Delta_1, y + \Delta_2) - f(x, y) - f(x + \Delta_1, y) - f(x, y + \Delta_2)}{\partial x \partial y}$$

$$M(I_{\Delta_1} - I_{\Delta_2})^2 = M\left(\frac{X(t+\Delta_1) - X(t)}{\Delta_1} - \frac{X(t+\Delta_2) - X(t)}{\Delta_2}\right)^2 =$$

$$= \frac{B(t+\Delta_1, t+\Delta_1) - B(t, t) - B(t+\Delta_1, t) - B(t, t+\Delta_1)}{\Delta_1 \Delta_1} +$$

Примітка 1. Розкриваємо квадрат, знаходимо мат. сподівання доданків

Примітка 2.  $B(t_1, t_2) = B(t_2, t_1)$ . Тому замість подвоєного добутку пишуться два

добутки, в якому в другому змін.порядок співмножників. Аналогічно записані наступні три дроби.

$$\frac{B(t+\Delta_2, t+\Delta_2) + B(t, t) - B(t+\Delta_2, t) - B(t, t+\Delta_2)}{\Delta_2 \Delta_2} -$$

$$- 2 \frac{B(t+\Delta_1, t+\Delta_2) + B(t, t) - B(t+\Delta_1, t) - B(t, t+\Delta_2)}{\Delta_1 \Delta_2}$$

Якщо існує  $\left. \frac{\partial^2 B_{X(t)}(t_1, t_2)}{\partial t_1 \partial t_2} \right|_{t_1=t_2=t}$ , то

$$\lim_{\substack{\Delta_1 \rightarrow \infty \\ \Delta_2 \rightarrow \infty}} M(I_{\Delta_1} - I_{\Delta_2})^2 = \frac{\partial^2 B(t, t)}{\partial t^2} + \frac{\partial^2 B(t, t)}{\partial t^2} - 2 \frac{\partial^2 B(t, t)}{\partial t^2} = 0$$

Значить ліміт існує і дорівнює нулю.

Це і є умовою існування похідної від випадкового процесу.

Примітка 1. Якщо  $X(t)$  - стаціонарний процес, то показати самим, що

умова існування похідної від нього є  $\left. \frac{\partial^2 B_{X(t)}(u)}{\partial u^2} \right|_{u=0}$ . З цього випливає, що з чотирьох приведених виразів(форм) запису функції  $\text{cov}$  стаціонарного процесу, похідну не має стаціонарний процес для 1, 2 випадків(6-та властивість функції  $\text{cov}$  стаціонарного випадкового процесу).

Знайдемо мат.сподівання функції  $\text{cov}$  від похідної випадкового процесу

$$Y(t) = M \lim_{\Delta \rightarrow 0} \frac{X(t + \Delta) - X(t)}{\Delta}$$

можна показати, що границя в середньому квадратичному існує, то мат.сподівання і границю можна поміняти місцями.

$$\lim_{\Delta \rightarrow 0} M \frac{X(t + \Delta) - X(t)}{\Delta} = \frac{dMX(t)}{dt}.$$

Знаходимо функцію cov

$$\overset{0}{Y}(t) = Y(t) - MY(t)$$

$$\overset{0}{X}(t) = X(t) - MX(t)$$

$$\overset{0}{Y}(t) = \frac{d\overset{0}{X}(t)}{dt}$$

$$MY(t) = \frac{dMX(t)}{dt}$$

$$B_{Y(t)}(t_1, t_2) = M \left( \overset{0}{Y}(t_1) \cdot \overset{0}{Y}(t_2) \right) = M \left[ \frac{d \overset{0}{X}(t_1)}{dt_1} - \frac{d \overset{0}{X}(t_2)}{dt_2} \right] = M \left( \frac{\partial^2 (\overset{0}{X}(t_1) \overset{0}{X}(t_2))}{\partial t_1 \partial t_2} \right) \quad (\text{друга}$$

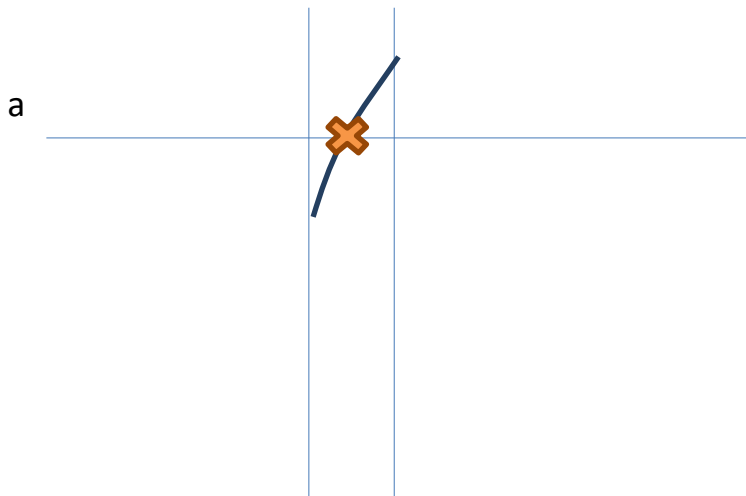
змішана часткова похідна від випадкового процесу і вона існує, коли існує відповідна границя в середньому квадратичному  $I_{\Delta_1 \Delta_2}$ )

## 5.16 Викиди випадкових процесів

Розглянемо випадковий процес  $X(t)$  і випадкову величину  $V(t) = \frac{dX(t)}{dt}$ .

Викидом випадкового процесу  $X(t)$  за рівень  $a$  на інтервалі довжини  $\Delta t$  є перетин випадковим процесом  $X(t)$  на цьому інтервалі рівня  $a$  знизу вгору.





В цьому розділі буде вивчатися ймовірнісний характер викидів випадкових процесів вище рівня  $a$ .

Ці моделі в інженерній практиці актуальні там де об'єкт (будь-якої фізичної природи) функціонує під впливом випадкового навантаження, що описується випадковим процесом  $X(t)$ . В рамках цього підходу розглянемо наступні 4 задачі.

#### Задача 1.

Знайдемо ймовірність настання викиду випадкового процесу вище рівня  $a$ .  $\Delta t$  - таке мале число, таке, що якщо на цьому інтервалі є викид, то він тільки 1.

Тоді викид настане, якщо:  $X(t) < a$

$$X(t + \Delta t) > a$$

З точністю  $o(\Delta t)$  можна записати, що  $X(t + \Delta t) = X(t) + V(t)\Delta t$ .

Таким чином, викид настане, якщо виконується нерівність

$$a - V(t)\Delta t < X(t) < a$$

Нехай  $X(t), V(t) \forall t$  - неперервні (абсолютно неперервні випадкові величини). Тоді існує двовимірна функція щільності  $f(x, v/t)$ ,

де  $x$  - аргумент випадкової величини  $X(t)$ ,

$v$  - аргумент випадкової величини  $V(t)$ .

Примітка. ">" можна замінити на " $\geq$ " і "<" на " $\leq$ " тому що величини неперервні.

Тоді ймовірність настання цієї події

$$P(a - V(t)\Delta t < X(t) < a) = \int_0^{\infty} \int_{a-v(t)}^a f(x, v/t) dx dv$$

$$P(XY \in G) = \iint_G f(xy) dx dy$$

Використана формула:

Примітка. Область  $G$  така, що подвійний інтеграл зводиться до двократного.

За теоремою про середнє отримаємо наближену формулу

$$P(a - V(t)\Delta t < X(t) < a) \approx \Delta t \int_0^{\infty} f(a, v/t) v dv$$

Чим менше  $\Delta t$  - тим більше точність наближеної формули.

Примітка. Використовується в задачах, де  $\Delta t \rightarrow 0$ .

Задача 2.

Знайти мат.сподівання часу перебування випадкового процесу вище рівня  $a$  на часовому інтервалі довжини  $T$ .

Для цього розіб'ємо інтервал довжини  $T$  на  $n$ -частин,  $n$  - достатньо велике число, таке, що можна припустити: на кожному  $i$ -ому інтервалі, де  $t_i$  - його лівий кінець, випадковий процес повністю вище рівня  $a$ , чи нижче.

Це припущення стає вірним, коли  $n$  прямує до нескінченності.

Тоді мат.сподівання дорівнює  $M \sum_{i=1}^n \varepsilon_i$

Де  $\varepsilon_i$  приймає два значення: 0 – якщо випадковий процес нижче рівня  $a$  на  $i$ -тому відрізку,  $\Delta t$  якщо випадковий процес вище рівня  $a$  на  $i$ -ому відрізку.

$$M \sum_{i=1}^n \varepsilon_i = \sum_{i=1}^n \Delta t \int_0^{\infty} f(x/t_i) dx,$$

де  $f(x/t)$  - функція щільності випадкової величини  $X(t)$ .

Спрямуємо  $\Delta t$  до 0:

$$\lim_{\Delta t \rightarrow 0} \sum_{i=1}^n \left( \Delta t \int_0^{\infty} f(x/t_i) dx \right) = \int_t^{t+T} \left[ \int_a^{\infty} (x/u) dx \right] du$$

Задача 3.

Знайдемо мат. сподівання кількості викидів вище рівня  $a$  випадкового процесу на інтервалі довжини  $T$ .

Розбиваємо відрізок на  $n$ -частин  $\varepsilon_i = \{0, 1\}$  (0 – викид не настався на  $i$ -ому відрізку, 1 – настав викид)

$$\sim M \sum_{i=1}^n \varepsilon_i$$

Точне значення  $M$  отримаємо, коли розглянемо  $\lim$  при  $\Delta t \rightarrow 0$

$$\lim_{\Delta t \rightarrow 0} M \sum_{i=1}^n \varepsilon_i = \lim_{\Delta t \rightarrow 0} \sum_{i=1}^n \left( \Delta t \int_0^{\infty} \varphi(a, v/t_i) v dv \right) = \int_t^{t+T} \left[ \int_0^{\infty} f(a, v/u) v dv \right] du$$

Задача 4.

Знайдемо мат. сподівання часу перебування випадкового процесу вище рівня  $a$  за 1 викид.

Розв'яжемо для спрощеного випадку.

$$f(x, v/t) = f(x, v)$$

$f(x/t) = f(x)$  - функція щільності від  $t$  не залежить.

В цьому випадку мат.сподівання часу перебування випадкового процесу вище рівня  $a$  за 1 викид дорівнює

$$\frac{\int_t^{t+T} \left[ \int_a^\infty f(x) dx \right] du}{\int_t^{t+T} \left[ \int_0^\infty f(a, v) v dv \right] du} = \frac{\int_a^\infty f(x) dx}{\int_0^\infty f(a, v) v dv}$$

## 6.

### 6.4 Випадкові послідовності

Розглянемо нескінченну випадкову послідовність випадкових величин

$$..., Z_{-n}, Z_{-n+1}, ..., Z_0, Z_1, ..., Z_n, ...$$

Означення. Випадкова послідовність зветься стаціонарною, якщо для неї виконуються:

$$\forall n, M Z_n = \text{const}_1, D Z_n = \text{const}_2,$$

$$\text{cov}(z_i, z_j) = \gamma_{|i-j|} = \text{cov}(z_0, z_{i-j}).$$

Математичний апарат побудови математичної моделі, що описує стаціонарні випадкові послідовності – це апарат кінцево-різницевого рівнянь.

Розглянемо наступні оператори:

1. Одиничний оператор

$$1z_t = z_t$$

2. Оператор зсуву на одиницю вліво

$$Bz_t = z_{t-1}$$

3. Оператор зсуву на j-одиниць

$$Bz_t = z_{t-j}$$

4. Перша різниця  $\Delta$

$$\Delta z_t = z_t - z_{t-1} = (1 - B)z_t$$

5. Оператор різниці II-го порядку

$$\Delta^2 = \Delta \cdot \Delta \text{ (послідовна дія оператора на оператор)}$$

$$\Delta^2 z_t = \Delta \cdot \Delta z_t = \Delta(z_t - z_{t-1}) = (1-B)(z_t - z_{t-1}) =$$

$$z_t - z_{t-1} - z_{t-1} + z_{t-2} = z_t - 2z_{t-1} + z_{t-2} = (1-2B+B^2)z_t = (1-B)^2 z_t$$

$$6. \Delta^d = \Delta \cdot \Delta^{d-1} = (1-B)^d$$

Примітка. Всі дужки розкриваються за правилами алгебри.

## 7. Оператор різниці I-го порядку

$$\Delta^{-1} \cdot \Delta = \Delta \cdot \Delta^{-1} = 1$$

$$\Delta = 1 - B$$

$$\Delta^{-1} = \frac{1}{1-B}, \left( \frac{1}{1-B} - \text{умовна запис оператора } \Delta^{-1} \right)$$

$$\Delta^{-1} \Delta z_t = \Delta \cdot \Delta^{-1} z_t = z_t$$

Покажемо, що  $\Delta^{-1} = 1 + B + B^2 + B^3 + \dots$ .

Дійсно

$$\Delta \Delta^{-1} z_t = \Delta (1 + B + B^2 + B^3 + \dots) z_t = (1-B)(z_t + z_{t-1} + \dots) z_t = z_t + z_{t-1} - z_{t-1} + \dots = z_t = 1 \cdot z_t$$

Головна задача цього розділу: знаходження більш компактних рекурентних моделей, що задають всевдостационарні послідовності.

Модель 1.

$$\tilde{z}_t = a_t + \psi_1 B a_t + \psi_2 B^2 a_t + \psi_3 B^3 a_t + \dots (1)$$

$$\tilde{z}_t = z_t - \mu, \mu = M Z_t \text{ для } \forall t$$

$\forall i \psi_i$  - вагові коефіцієнти.

$a_t$  - незалежні випадкові величини для всіх  $t$  з нульовим мат.сподіванням і невідомою, але обмеженою дисперсією. Всі  $a_t$  розподілені за нормальним законом. Послідовність випадкових величин  $a_t$  зветься білим шумом.

Можна (1) представити в операторній формі у вигляді

$$z_t = \psi(B) a_t \quad (2)$$

$$\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$$

$$\psi(0) = 1$$

Має місце наступна теорема (дається без доведення)

Модель 1 задає стаціонарну випадкову послідовність, якщо вважаючи  $B$  не оператором зсуву а числом

$$\forall B: |B| \leq 1$$

$$|\psi(B)| < \infty$$

Модель 2.

Авторегресії порядку  $p$

$$\tilde{z}_t = \Phi_1 \tilde{z}_{t-1} + \dots + \Phi_p \tilde{z}_{t-p} + a_t \quad (3)$$

де  $\Phi_i$  - раціональні числа, вагові коефіцієнти.

Примітка. Модель має таку назву, тому що  $\tilde{z}_t$  регресує на свої попередні значення.

Введемо оператор  $\Phi(B) = 1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p$

Тоді модель (3) задається в операторній формі скороченим чином:

$$\Phi(B) \tilde{z}_t = a_t \quad (4)$$

Знайдемо оператор, обернений до  $\Phi(B)$

$$\Phi^{-1}(B) \cdot \Phi(B) = 1$$

Будемо його шукати для наступного часткового випадку, а саме

$$\Phi(B) = \prod_{i=1}^p (1 - G_i B), \text{ де } G_i - \text{дійсні числа.}$$

Покажемо, що в цьому випадку

$$\Phi^{-1}(B) = \sum_{i=1}^p \frac{K_i}{1 - G_i B}, \quad \frac{1}{1 - G_i B} = (1 - G_i B)^{-1} = 1 + G_i B + G_i^2 B^2 + \dots$$

Примітка. Доведення абсолютно аналогічне тому, яке наведено для оператора  $\Delta^{-1}$ .

Коефіцієнти  $K_i$  задовольняють умові  $\sum_{i=1}^p K_i = 1$ .

Доведення достатньо привести для прикладу, коли  $p=2$ . З цього прикладу однозначно випливає виведення загального результату.

$$1 + \Phi_1 B + \Phi_2 B^2 = (1 - G_1 B)(1 - G_2 B)$$

$$\sum_{i=1}^2 \frac{K_i}{(1 - G_i B)} \prod_{j=1}^2 (1 - G_j B)$$

$$K_1(1 - G_2 B) + K_2(1 - G_1 B) = 1$$

Примітка. У другому доданку поміняти порядок дужок, тому що виконується умова комутативності операторів

Отримали систему рівнянь

$$K_1 + K_2 = 1$$

$$-G_2 K_1 - G_1 K_2 = 0$$

Примітка. Самостійно задати структуру системи лінійних рівнянь для будь-якого  $p$ .

Таким чином, обернений оператор можна записати у вигляді  $\Phi^{-1}(B)$ .

$$\tilde{z}_t = \Phi^{-1}(B) a_t \quad (5) \text{ впливає з (4).}$$

Примітка. Для цього зліва і справа потрібно подіяти оператором  $\Phi^{-1}(B)$ .

$$\tilde{z}_t = \sum_{i=1}^p K_i (a_t + G_i a_{t-1} + G_i^2 a_{t-2} + \dots) \quad (6)$$

Тоді необхідною і достатньою умовою того, що модель 2(авторегресії порядку  $p$ ) задає стаціонарну випадкову послідовність, є  $\forall_i |G_i| < 1$ .

Показати самим.

Модель 3.

### Модель ковзного-середнього порядку q

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (7)$$

$\theta_i$  - дійсні числа

$\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ . Модель (7) має вигляд

$$\tilde{z}_t = \theta(B) a_t$$

Всі  $\theta_i$  обмежені по модулю. Показати самим, що модель задає стаціонарну випадкову послідовність.

Модель 4.

Модель авторегресії та ковзаного-середнього порядку pq

$$\tilde{z}_t = \Phi_1 \tilde{z}_{t-1} + \Phi_2 \tilde{z}_{t-2} + \dots + \Phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (8)$$

Використовуємо вже введені нами оператори

$$\Phi(B) \tilde{z}_t = \Theta(B) a_t \quad (9)$$

$$\tilde{z}_t = \Phi^{-1}(B) \Theta(B) a_t \quad (10)$$

Умова стаціонарності, це об'єднання двох умов: всі  $\theta_i$  обмежені по модулю,  $\forall_i |G_i| < 1$ .

Модель 5.

Частковий випадок нестационарної послідовності. Модель 1 не може задавати нестационарну випадкову послідовність. Для того, щоб модель 1 задавала нестационарну випадкову послідовність, необхідно замість  $\psi_i$  підставити  $\psi_i(t)$ , тобто ф-цію від часу.

Але одним частковим випадком, задається моделлю (1), а саме

$$\Phi(B)(1-B)^d z_t = \theta(B) a_t \quad (11)$$



Тобто, модель 5 є стаціонарною відносно  $\Delta^d z_t$ , але не стаціонарною відносно  $z_t$ .

### *Додаткові властивості стаціонарних послідовностей*

Коваріаційна та кореляційна матриці випадкової стаціонарної послідовності.

Розглянемо випадкову величину  $\tilde{z}_1, \tilde{z}_2 \dots \tilde{z}_n$  (стаціонарна послідовність)

Побудуємо їх коваріаційну матрицю

За означенням стаціонарної послідовності її  $(ij)$  елемент має вигляд

$$(ij) = \text{cov}(\tilde{z}_i, \tilde{z}_j) = M(\tilde{z}_i, \tilde{z}_j) = j_{|j-i|}$$

Таким чином коваріаційна матриця :

$$\begin{pmatrix} j_0 & j_1 & j_2 & \dots & j_{n-1} \\ j_1 & j_0 & j_1 & \dots & j_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ j_{n-1} & j_{n-2} & \dots & j_1 & j_0 \end{pmatrix} = j_0 \begin{pmatrix} 1 & \rho_1 & \dots & \dots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{n-1} & \dots & \dots & \dots & 1 \end{pmatrix}, \text{ це}$$

кореляційна матриця

Примітки

$$j_0 = D\tilde{z}_i = z_z^2$$

Нагадаємо, що

$$\rho_{xy} = \frac{\text{cov}(xy)}{\delta x \delta y}$$

Звідки

$$\rho_{|i-j|} = \frac{j_{|i-j|}}{j_{|0|}}$$

Розглянемо випадкові величини:

$$\tilde{z}_t, \tilde{z}_{t-1}, \dots, \tilde{z}_{t-n+1}$$

Беремо довільні дійсних чисел:

$$l_1, l_2, \dots, l_n, \exists l_j \neq 0$$

Беремо випадкову величину:

$$\tilde{L}_t = \sum_{i=1}^n l_i \tilde{z}_{t-i+1}$$

Так, як  $\tilde{z}_{t-i+1}$  мають нульове мат. сподівання, то

$$D\tilde{L}_t = M\tilde{L}_t^2 = M \sum_{i=1}^n \sum_{j=1}^n l_i l_j \cdot \tilde{z}_{t-i+1} \tilde{z}_{t-j+1} = \sum_{i=1}^n \sum_{j=1}^n l_i l_j M(\tilde{z}_{t-i+1} \cdot \tilde{z}_{t-j+1}) = \sum_{i=1}^n \sum_{j=1}^n l_i l_j \cdot j_{|i-j|} > 0$$

Таким чином ми показали, що коваріаційна (або кореляційна) матриця додатньо означена, тобто детермінанти всіх її головних мінорів строго більше нуля.

Проблема оберненості моделі.

Головна задача цього розділу – знаходження найбільш компактних рекурентних моделей, що задають стаціонарні послідовності.

$$\tilde{z}_t = a_t + \psi_1 \cdot a_{t-1} + \psi_2 \cdot a_{t-2} + \dots = \psi(B) \cdot a_t(1)$$

$$\psi_0 = 1$$

$$\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j(2)$$

$$\text{Ця модель має еквівалентний вигляд } \psi^{-1}(B) \cdot \tilde{z}_t = a_t$$

Зліва і зправа подіяли оберненим оператором.

*Приклад 1*

$$\tilde{z}_t = a_t - \theta \cdot a_{t-1} = (1 - \theta \cdot B) \cdot a_t(3)$$

$$(1 - \theta \cdot B)^{-1} = 1 + \theta \cdot B + \theta^2 \cdot B^2 + \dots$$

Тоді для прикладу, модель (2) має вигляд.

$$\tilde{z}_t = \theta \widetilde{z_{t-1}} - \theta^2 \widetilde{z_{t-2}} - \theta^3 \widetilde{z_{t-3}}$$

3) Задає стац. послідовність для всіх  $|\theta| < \infty$  (обмеженість по модулю, дивись попередню лекцію).

Таким чином сформулювали проблему оберненості. Якщо  $|\theta| > 1$ , тоді обернена модель (3) не має жодного практичного змісту, бо чим більше  $j$ , тим більше вклад  $\widetilde{z_{t-j}}$  в значення  $\tilde{z}_t$ . Проблема оберненого розв'язку моделі (3) відповідає здоровому глузду, якщо  $|\theta| < 1$ .

*Приклад 2*

Модель ковзкого середнього порядку  $p$  :

$$\tilde{z}_t = \theta(B)$$

$$\theta(B) = \prod_{j=1}^q (1 - M_j \cdot B)$$

Де  $M_j$  – дійсні числа. Модель буде практичною, коли всі  $M_j < 1$ .

Показати самим.

Примітка. Подивитися вигляд оберненого оператора авторегресії порядку  $p$ .

Для першої моделі знайдемо аналітичний вираз для  $\gamma_k$

$$\tilde{z}_t = \sum_{j=0}^{\infty} \psi_j \cdot a_{t-j}$$

$$\gamma_k = M(\tilde{z}_t \cdot \tilde{z}_{t-k})$$

$$\tilde{z}_t = \sum_{j=0}^{\infty} \psi_j \cdot a_{t-j}$$

$$\tilde{z}_{t-k} = \sum_{h=0}^{\infty} \psi_h \cdot a_{t-k-h}$$

$$\gamma_k = \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} \psi_j \psi_h \cdot M(a_{t-j} \cdot a_{t-k-h})$$

Випадкові величини  $a_t$  є незалежними, розподілені однаково по нормальному закону з нульовим математичним сподіванням та дисперсією  $\sigma_a^2$ , тоді в нашій подвійній сумі всі добутки для яких  $t-j \neq t-k-h$  дорівнюють нулю.

Таким чином залишаються лише члени з індексами, де  $j = k+h$ .

Звідки

$$j_k = \sigma_a^2 \cdot \sum_{h=0}^{\infty} \psi_{h+k} \psi_h$$

$$j_0 = \sigma_a^2 = \sum_{h=0}^{\infty} \psi_h^2$$

### 6.5.1 Модель авто регресії порядку $p$

$$\tilde{Z}_t = \Phi_1 \tilde{Z}_{t-1} + \dots + \Phi_p \tilde{Z}_{t-p} + a_t \quad (1)$$

$$\tilde{Z}_t = \Phi^{-1}(B) a_t = a_t + \psi_1 a_{t-1} + \dots, \quad (2)$$

Показати самим.

Примітка: коефіцієнт при  $a_t$  дорівнює

$$\sum_{i=1}^p k_i = 1.$$

Тоді

$$\gamma_k = M(\tilde{Z}_t \tilde{Z}_{t-k}) = \left[ \text{Примітка: 1} \right] \text{рівність один зліва і справа множимо на } \tilde{Z}_{t-k} \text{ і беремо математичне сподівання}$$

$$\tilde{Z}_t \Phi_1 \gamma_{|k-1|} + \Phi_2 \gamma_{|k-2|} + \dots + \Phi_p \gamma_{|k-p|} \quad (3)$$

$$\gamma_0 = \Phi_1 \gamma_1 + \Phi_2 \gamma_2 + \dots + \Phi_p \gamma_p + \sigma_a^2$$

Поділимо вираз на  $\gamma_0$  ( $\gamma_0 = \sigma_{\tilde{Z}_t}^2$  для  $\forall t$ ):

$$1 = \Phi_1 \rho_1 + \Phi_2 \rho_2 + \dots + \Phi_p \rho_p + \frac{\sigma_a^2}{\sigma_{\tilde{Z}_t}^2}$$

$$\sigma_{\tilde{Z}_t}^2 = \frac{\sigma_a^2}{1 - \Phi_1 \rho_1 - \Phi_2 \rho_2 - \dots - \Phi_p \rho_p}$$

Рівняння Юла-Уокера

Рівняння (3) запишемо для коефіцієнтів кореляції (а не для коваріації)

для усіх  $k=1, \dots, p$ , які зветься рівнянням Юла-Уокера

$$\rho_k = \Phi_1 \rho_{|k-1|} + \Phi_2 \rho_{|k-2|} + \dots + \Phi_p \rho_{|k-p|}, k=1, \dots, p \quad (4)$$

$$\begin{cases} \rho_1 = \Phi_1 + \Phi_2 \rho_1 + \dots + \Phi_p \rho_{p-1} \\ \rho_2 = \Phi_1 + \Phi_2 \rho_1 + \dots + \Phi_p \rho_{p-2} \\ \dots \\ \rho_p = \Phi_1 + \Phi_2 \rho_1 + \dots + \Phi_{p-1} \rho_p + \Phi_p \end{cases}$$

Визначення: Якщо відомі коефіцієнти  $\Phi_1 \dots \Phi_p$ , то по системі рівнянь (4)

можна знайти  $\rho_1 \dots \rho_p$ . А по рівнянню (4) можна знайти усі інші  $\rho_k$  для  $k > p$ .

### 6.5.2 Модель ковзуного середнього порядку $q$

$$\tilde{Z}_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

$$\gamma_k = M \left[ (a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}) (a_{t-k} - \theta_1 a_{t-k-1} - \dots - \theta_q a_{t-k-q}) \right]$$

$$\gamma_0 = \sigma_a^2 (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)$$

$$\gamma_k = 0, k \geq q+1$$

$$\gamma_k = \sigma_a^2 (-\theta_k + \theta_{k+1} \theta_1 + \dots + \theta_q \theta_{q-k}), k < q$$

$$\rho_k = \frac{-\theta_k + \theta_{k+1} \theta_1 + \dots + \theta_q \theta_{q-k}}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2}$$

### 6.5.3 Модель авторегресії та ковзного середнього порядку $pq$

$$\Phi(B) \tilde{z}_t = \theta(B) a_t$$

$$\tilde{z}_t = \Phi_1 \tilde{z}_{t-1} + \dots + \Phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} (1)$$

$$\tilde{z}_t = \Phi^{-1}(B) \Theta(B) a_t = a_t + \psi a_{t-1} + \dots$$

Показати самим

$$\gamma_k = M(\tilde{z}_t \tilde{z}_{t-k})$$

Прим.: зліва та справа (1) домножимо на  $\tilde{z}_{t-k}$  та візьмемо мат. сподівання.

$$\gamma_k = \Phi_1 \gamma_{|k-1|} + \dots + \Phi_p \gamma_{|k-p|} + \gamma_{za}(k) - \Theta_1 \gamma_{za}(k-1) - \dots - \Theta_q \gamma_{za}(k-q),$$

$$\gamma_0 = \Phi_1 \gamma_1 + \dots + \Phi_p \gamma_p + \gamma_{za}(0) - \Theta_1 \gamma_{za}(-1) - \dots - \Theta_q \gamma_{za}(-q)$$

де

$$\gamma_{za}(k-j) = M(\tilde{z}_{t-k} a_{t-j})$$

$$\gamma_{za}(u) = 0, \forall u > 0$$

### 6.5.4 Модель авторегресії та інтегрованого ковзного середнього порядку $dprq$

$$\Phi(B) \Delta^d z_t = \Theta(B) a_t (1)$$

$$\Phi(B) (1-B)^d z_t = \Theta(B) a_t$$

Рівняння (1) шукають у вигляді  $z_t = \psi(B) a_t$ . Перепишемо (1) у вигляді:

$$\varphi(B) z_t = \varphi(B) \psi(B) a_t, \varphi(B) = \Phi(B) (1-B)^d$$

$\psi(B)$  – має наступний вигляд:  $\sum_{j=0}^{\infty} \psi_j(B)$ , тому що  $\psi_0=1$ .

$$\varphi(B)\psi(B)=\Theta(B), \text{ тобто}$$

$$\varphi(B)z_t=\Theta(B)a_t$$

Далі необхідно коефіцієнти при  $B$  с однаковими степенями зліва та справа прирівняти та розв'язати відповідну систему лінійних рівнянь.

## 7. Елементи теорії нечітких множин

Розглянемо одне з визначень звичайної множини:

Є певна властивість і всі елементи множини  $A$  задовольняють цій властивості

Зде перший показав, що існують такі властивості, відносно яких не завжди можна точно відповісти, задовольняє елемент цій властивості чи ні.

Наприклад: всі предмети, які мають червоний колір.

Нечіткою множиною  $A$ , заданою на звичайній множині  $X$  є пара:

$$\{\mu_A(x), X\}, \forall x \in X$$

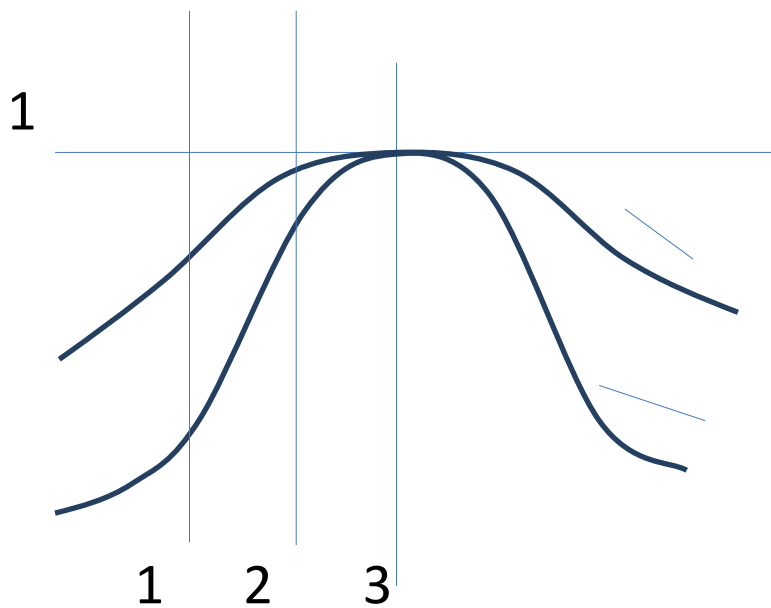
$0 \leq \mu_A(x) \leq 1$  - функція належності нечіткої множині  $A$ .

Нечітка множина називається правильною, якщо  $\exists \max_x \mu_A(x) = 1$ .

*Визначення:*  $A \subseteq B$ , якщо  $\mu_A(x) \leq \mu_B(x)$

Прим.: має місце зміст тільки тоді, коли функції приналежності характеризують одну властивість.

Приклад: множина  $B$  – множина усіх чисел, за модулем близьких до 3,  $A$  – усі числа, котрі дуже близько до 3.



Дії над нечіткими множинами:

Об'єднання нечітких множин  $A \cup B$ :

$$1. \mu_{A \cup B}(x) = \begin{cases} 1, & \mu_A(x) + \mu_B(x) > 1 \\ \mu_A(x) + \mu_B(x), & \mu_A(x) + \mu_B(x) < 1 \end{cases}$$

$$2. \mu_{A \cup B}(x) = \max\{\mu_A(x), \mu_B(x)\}$$

Прим.: можуть використовуватися, коли нечіткі множини задані різними властивостями.

Прим.: Дії над нечіткими множинами неоднозначні, як видно із визначення. Але обов'язково повинні виконуватися визначення операції об'єднання для звичайних множин:

$$\mu_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

Перше та друге визначення задовольняють визначенню об'єднання.

Перетин нечітких множин  $A \cap B$

$$1) \mu_{A \cap B}(x) = \mu_A(x) \mu_B(x)$$

$$2) \mu_{A \cap B}(x) = \min_{\mu} \{\mu_A(x), \mu_B(x)\}$$

Може задаватися, коли  $A$  та  $B$  задані різними властивостями.

Прим.: В теорії множин найбільш розповсюджені наступна дія (операція об'єднання):

$$\mu_{A \cup B} = \max_{\mu} \{ \mu_A(x), \mu_B(x) \}$$

інтерпретується як найбільша нечітка множина, що характеризує властивості або нечіткої множини  $A$ , або нечіткої множини  $B$

Приклад:  $A$  – множина чисел за модулем близьких до 5 (або нечітке число 5).  $B$  – нечітке число 10.

Тоді  $A \cup B$ -множина усіх чисел близьких або до 5 або до 10.

$$\mu_{A \cap B} = \min_{\mu} \{ \mu_A(x), \mu_B(x) \}$$

Інтерпретується як найменша множина, що характеризує міру властивості як властивості  $A$  так і властивості  $B$ . У нашому прикладі це нечітка множина  $A \cap B$  – усі числа, котрі за модулем близькі як к 5, так і к 10 одночасно.

Визначення: Нечітка множина  $\bar{A}$  задається функцією приналежності як  $1 - \mu_A(x)$  та задає кількісну міру протилежної властивості.

Визначення: Нечітка множина  $A \dot{-} B$  задається як:

$$\mu_{A \dot{-} B} = \begin{cases} 0, & \mu_A - \mu_B < 0 \\ \mu_A - \mu_B, & \mu_A - \mu_B > 0 \end{cases}$$

$A$  та  $B$  повинні характеризувати одну й ту ж саму властивість.

Визначення:

Є система  $\{A_1 X_1\} \dots \{A_n X_n\}$   $X_1 \dots X_n$  – звичайні множини, на кожній з яких задана нечітка множина  $A_1 \dots A_n$ .

Тоді нечітка множина  $A = A_1 \oplus \dots \oplus A_n$ , що є декартовим добутком, що задана на звичайній множині  $X = X_1 \dots \oplus X_n$

$$\text{Якщо } x = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} \in X_i, x_i \in X_i$$

$$\mu_A(x) = \min_{\square} \{ \mu_{A_1}(x_1), \mu_{A_2}(x_2) \dots \mu_{A_n}(x_n) \}$$



Визначення.

Звичайна множина  $X_\alpha = \{x : \mu_A(x) \geq \alpha\}$

Самим довести, що

$$\mu_A(x) = \min_{\alpha} \{x \in X_\alpha\}$$

Таким чином теорія нечітких множин є математичним апаратом, який як і теорія ймовірностей досліджує об'єкти за властивостями.

## 8. Література

1. Гихман И.И., Скороход А. В., ядренко М.И. Теория вероятностей и математическая статистика. –К.: Вища школа, 1979. -408с.
2. Розанов Ю.А. Теория вероятностей, случайные процессы и математическая статистика. –М.: Наука, 1985. -320с.
3. Смирнов Н.В., Дунин – Барковский И.В. Курс теории вероятностей и математической статистики для технических приложений. –М.: Наука, 1965. -512с.
4. Худсон Д. Статистика для физиков. –М.: Мир, 1967. -242с.
5. Крамер Г. Математические методы статистики. –М.: Мир, 1975. -648с.
6. Андерсон Т. Введение в многомерный статистический анализ. – М.: Физматгиз, 1963. -500с.
7. Коваленко И.Н., Филиппова А.А. Теория вероятностей и математическая статистика. –М.: Высшая школа, 1982. -256с.
8. Вентцель Е.С., Овчаров Л.А. Теория случайных процессов и её инженерные приложения. –М.: Наука, 1991. -384с.
9. Свешников А.А. Прикладные методы теории случайных функция. –М.:наука, 1968. -463с.
- 10.Анисимов В.В., Закусило О.К., Донченко В.С. Элементы теории саммового обслуживания и асимптотического анализа систем. – К.: Вища школа, 1987. -248с.

- 11.Боровиков В.П., Ивченко Г.И. Прогнозирование в системе STATISTICA в среде WINDOWS. –М.: Финансы и статистика, 1999. -384с.
- 12.Коньшева Л.К., Назаров Д.М. Основы теории нечётких множеств. –СПб.: Питер, 2011. -192с.