Довірчі інтервали для параметрів нормального розподілу.

1 Теоретичні відомості

Що таке довірчий інтервал?

Нехай $\theta \in \Theta \subset \mathbb{R}$ – невідомий параметр, $p \in (0, 1)$. Довірчим інтервалом для параметра θ довірчого рівня p (або рівня значущості $1 - p$) називають таку пару статистик $(\hat{\theta}_-, \hat{\theta}_+)$, що $\hat{\theta}_- < \hat{\theta}_+$ м.н. та

$$\mathbf{P}_\theta \left(\theta \in [\hat{\theta}_-, \hat{\theta}_+] \right) = p, \; \theta \in \Theta.$$

Зауважте, що випадковість лише є у побудованих статистиках $\hat{\theta}_\pm$, бо вони є функціями від спостережуваної вибірки X.

Тобто довірчий інтервал – це такий інтервал, побудований за даними X, який покриває невідомий параметр θ із імовірністю p.

Розподіли, пов'язані з нормальним

Нехай $\{\xi_j\}_{j=1}^n$ – послідовність н.о.р. випадкових величин зі стандартним нормальним розподілом, тобто $X_1 \sim N(0, 1)$.

Випадкова величина, яка є сумою квадратів ξ_j:

$$\eta_n = \sum_{j=1}^n \xi_j^2$$

має розподіл хі-квадрат з n ступенями вільності (тобто $\eta_n \sim \chi_n^2$).

Нехай $\xi \sim N(0, 1)$ – ще одна випадкова величина, яка не залежить від η_n. Відношення ξ до квадратного кореня з η_n/n

$$\tau_n = \frac{\xi}{\sqrt{\frac{1}{n} \cdot \eta_n}}$$

має розподіл Стьюдента (або T-розподіл) з n ступенями вільності (тобто $\tau_n \sim T(n)$).

Нехай $\eta_n \sim \chi_n^2$ та $\eta_d \sim \chi_d^2$, причому η_n та η_d є незалежними. Відношення вигляду

$$f_{n,d} = \frac{\frac{1}{n} \cdot \eta_n}{\frac{1}{d} \cdot \eta_d}$$

має розподіл Фішера (або F-розподіл) з n ступенями вільності чисельника та d ступенями вільності знаменника (тобто $f_{n,d} \sim F(n, d)$).

Побудова довірчих інтервалів для середнього та дисперсії

Розглянемо кратну вибірку $X = (X_1, \dots, X_n)$ з нормальним розподілом спостережень, тобто $X_j \sim N(\mu, \sigma^2)$.

Побудуємо довірчий інтервал для μ в залежності від того, знаємо ми σ чи ні.

1. σ відома. Розглянемо статистику $\overline{X}_n = \frac{1}{n}\sum_{j=1}^n X_j$ – вибіркове середнє. Неважко побачити (в силу властивості розподілу суми нормальних величин), що

$$\overline{X}_n \sim N(\mu, \sigma^2/n).$$

   Значить,

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \sim N(0,1).$$

   Нехай x_p – таке число, щоб виконувалася рівність

$$\mathbf{P}_\mu\left(\left|\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma}\right| \le x_p\right) = p.$$

   Тобто $x_p = Q^{N(0,1)}((1+p)/2)$ є квантилем стандартного нормального розподілу рівня p. Розв'яжемо нерівність в останній імовірності:

$$\begin{aligned}
&\left|\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma}\right| \le x_p \\
&\Leftrightarrow -x_p \le \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \le x_p \\
&\Leftrightarrow -x_p\frac{\sigma}{\sqrt{n}} \le \overline{X}_n - \mu \le x_p\frac{\sigma}{\sqrt{n}} \\
&\Leftrightarrow -\overline{X}_n - x_p\frac{\sigma}{\sqrt{n}} \le -\mu \le -\overline{X}_n + x_p\frac{\sigma}{\sqrt{n}} \\
&\Leftrightarrow \overline{X}_n - x_p\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X}_n + x_p\frac{\sigma}{\sqrt{n}}
\end{aligned}$$

   де остання нерівність дає кінці довірчого інтервалу для μ рівня p у випадку відомої дисперсії.

2. σ невідома. В цьому разі звести до стандартного нормального розподілу не можемо, бо треба знати σ. З іншого боку, (виправлена) вибіркова дисперсія $\hat{s}_n^2 = \frac{1}{n-1}\sum_{j=1}^n (X_j - \overline{X}_n)^2$ має розподіл хі квадрат, тобто

$$\frac{(n-1)}{\sigma^2}\hat{s}_n^2 \sim \chi^2_{n-1}$$

   Значить, величина вигляду

$$\hat{\tau}_n = \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sqrt{\hat{s}_n^2}} = \frac{\frac{\sqrt{n}}{\sigma}(\overline{X}_n - \mu)}{\sqrt{\frac{1}{n-1}\left(\frac{(n-1)}{\sigma^2}\hat{s}_n^2\right)}} \sim T(n-1)$$

   має розподіл Стьюдента з $n-1$ ступенями вільності. Для побудови довірчого інтервалу для μ можна скористатися попередніми міркуваннями, оскільки розподіл Стьюдента є симетричним розподілом.

Дійсно, нехай t_p – таке число, щоб виконувалася рівність

$$\mathbf{P}_\theta\left(|\hat{\tau}_n| \le t_p\right) = p, \text{ де } \theta = (\mu, \sigma).$$

Тобто $t_p = Q^{T(n-1)}((1+p)/2)$ є квантилем розподілу Стьюдента з $n-1$ ступенями вільності рівня p. Розв'яжемо нерівність в останній імовірності:

$$\begin{aligned} |\hat{\tau}_n| = \left|\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma}\right| &\le t_p \\ \Leftrightarrow -t_p \le \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sqrt{\hat{s}_n^2}} &\le t_p \\ \Leftrightarrow \overline{X}_n - t_p\sqrt{\frac{\hat{s}_n^2}{n}} \le \mu &\le \overline{X}_n + t_p\sqrt{\frac{\hat{s}_n^2}{n}} \end{aligned}$$

де остання нерівність дає кінці довірчого інтервалу для μ рівня p у випадку невідомої дисперсії.

Побудуємо довірчий інтервал для σ^2 в залежності від того, знаємо ми μ чи ні.

1. μ відоме. Розглянемо статистику $\hat{\sigma}_n^2 = \frac{1}{n}\sum_{j=1}^n (X_j - \mu)^2$. Неважко побачити, що $\hat{\sigma}_n^2$ має розподіл
$$\frac{n}{\sigma^2}\hat{\sigma}_n^2 \sim \chi_n^2.$$

   Нехай h_- та h_+ – такі числа, що

$$\mathbf{P}_\mu\left(h_- \le \frac{n}{\sigma^2}\hat{\sigma}_n^2 \le h_+\right) = p. \tag{1}$$

   Можуть бути різні підходи до вибору $h_\pm$. Ми обмежимося наступним: оберемо $h_\pm$ так, щоб імовірність вийти за межі відрізка в ліву або в праву сторону була рівною, тобто

$$\mathbf{P}_\mu\left(\frac{n}{\sigma^2}\hat{\sigma}_n^2 < h_-\right) = \mathbf{P}_\mu\left(\frac{n}{\sigma^2}\hat{\sigma}_n^2 > h_+\right) = \frac{1-p}{2}.$$

   Можна побачити, що $h_- = Q^{\chi_n^2}((1-p)/2)$ та $h_+ = Q^{\chi_n^2}(1-(1-p)/2)$ є квантилями розподілу хі-квадрат з n ступенями вільності рівня $(1-p)/2$ та $1-(1-p)/2 = (1+p)/2$ відповідно.

   Розв'яжемо нерівність в імовірності (1) відносно σ^2:

$$\begin{aligned} h_- \le \frac{n}{\sigma^2}\hat{\sigma}_n^2 &\le h_+ \\ \Leftrightarrow \frac{h_-}{n\hat{\sigma}_n^2} \le \frac{1}{\sigma^2} &\le \frac{h_+}{n\hat{\sigma}_n^2} \\ \Leftrightarrow \frac{n\hat{\sigma}_n^2}{h_+} \le \sigma^2 &\le \frac{n\hat{\sigma}_n^2}{h_-} \end{aligned}$$

   Остання нерівність дає довірчий інтервал для дисперсії рівня p у випадку відомого середнього.

2. μ невідоме. Можна звести міркування до попередніх, якщо взяти до уваги виправлену вибіркову дисперсію $\hat{s}_n^2$. Тоді

$$\frac{(n-1)\hat{s}_n^2}{h_+} \leq \sigma^2 \leq \frac{(n-1)\hat{s}_n^2}{h_-}$$

задає довірчий інтервал для дисперсії коли середнє невідоме, де $h_- = Q^{\chi^2_{n-1}}((1-p)/2)$ та $h_+ = Q^{\chi^2_{n-1}}(1-(1-p)/2)$ є квантилями розподілу хі-квадрат з $n-1$ ступенями вільності рівня $(1-p)/2$ та $1-(1-p)/2 = (1+p)/2$ відповідно.

Довірчі інтервали для функцій від невідомих параметрів

Розглянемо дві незалежні кратні вибірки $X = (X_1, \ldots, X_{n_X})$ та $Y = (Y_1, \ldots, Y_{n_Y})$, де спостереження мають розподіл $X_j \sim N(\mu_X, \sigma_X^2)$ та $Y_j \sim N(\mu_Y, \sigma_Y^2)$.

Дослідимо можливості побудови довірчого інтервалу для різниці між середніми, тобто для $\mu_X - \mu_Y$ за припущення, що $\sigma_X = \sigma_Y = \sigma > 0$. Введемо статистики $\overline{X} = \frac{1}{n_X}\sum_{j=1}^{n_X} X_j$ та $\overline{Y} = \frac{1}{n_Y}\sum_{j=1}^{n_Y} Y_j$. Зауважимо:

$$\overline{X} - \overline{Y} \sim N\left(\mu_X - \mu_Y, \sigma^2\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)\right) \Rightarrow$$
$$\Rightarrow \sqrt{\frac{n_X n_Y}{n_X + n_Y}} \cdot \frac{(\overline{X} - \overline{Y}) - (\mu_X - \mu_Y)}{\sigma} \sim N(0,1).$$

Остання величина вже дає нам можливість побудувати довірчий інтервал для різниці середніх у випадку відомої дисперсії. Але коли дисперсія невідома, то треба добивати далі, до розподіла Стьюдента.

Отже продовжимо міркування коли $\sigma > 0$ невідоме. Введемо виправлені вибіркові дисперсії $\hat{s}_X^2 = \frac{1}{n_X-1}\sum_{j=1}^{n_X}(X_j - \overline{X})^2$ та $\hat{s}_Y^2 = \frac{1}{n_Y-1}\sum_{j=1}^{n_Y}(Y_j - \overline{Y})^2$. В силу незалежності статистик та їхнього розподілу, маємо

$$\frac{n_X-1}{\sigma^2}\hat{s}_X^2 + \frac{n_Y-1}{\sigma^2}\hat{s}_Y^2 \sim \chi^2_{n_X+n_Y-2}$$

Отже, статистика вигляду

$$\hat{\tau}_{n_X,n_Y} = \sqrt{\frac{n_X n_Y}{n_X + n_Y}} \cdot \frac{(\overline{X} - \overline{Y}) - (\mu_X - \mu_Y)}{\sqrt{((n_X-1)\hat{s}_X^2 + (n_Y-1)\hat{s}_Y^2)/(n_X+n_Y-2)}} \sim T(n_X + n_Y - 2)$$

має розподіл Стьюдента з $n_X + n_Y - 2$ ступенями вільності, з якої вже можна виводити довірчий інтервал для різниці середніх схожими міркуваннями, як у випадку однієї вибірки.

Наостанок розглянемо ситуацію із побудовою довірчих інтервалів для відношення дисперсій σ_X^2/σ_Y^2 (тут вважаємо що дисперсії у двох вибірках необов'язково рівні). Введемо величину

$$\hat{f}_{n_X,n_Y} = \frac{\frac{1}{\sigma_X^2}}{\frac{1}{\sigma_Y^2}} \cdot \frac{\hat{s}_X^2}{\hat{s}_Y^2} = \frac{\sigma_Y^2}{\sigma_X^2} \cdot \hat{Q}_{XY}, \ \left(\hat{Q}_{XY} := \frac{\hat{s}_X^2}{\hat{s}_Y^2}\right)$$

яка має розподіл Фішера з $n_X - 1$ ступенями вільності чисельника та $n_Y - 1$ ступенями вільності знаменника. Будувати довірчий інтервал можна з аналогічних міркувань як у випадку дисперсії. Розглянемо h_- та h_+ – такі числа, що

$$\mathbf{P}\left(h_- \leq \hat{f}_{n_X, n_Y} \leq h_+\right) = p.$$

Тоді

$$\frac{\hat{Q}_{XY}}{h_+} \leq \frac{\sigma_X^2}{\sigma_Y^2} \leq \frac{\hat{Q}_{XY}}{h_-}$$

задає довірчий інтервал для відношення дисперсій (при невідомих середніх).

Ключова фігура у побудові довірчих інтервалів

Як ви могли побачити, то побудовані довірчі інтервали відштовхуються від таких функцій від вибірки, розподіл яких не залежить від невідомих параметрів статистичної моделі (принаймні асимптотично). Такі величини називають *опорними*.

Далі підуть задачі.

2 Задачі

2.1 Задача 1

Середній I.Q. (рівень інтелекту) 100 студентів університету Меріленду дорівнює 110 зі стандартним відхиленням 5.

1. Знайти 95%-ий довірчий інтервал для теоретичного середнього I.Q. генеральної сукупності всіх студентів університету.
2. Знайти 95%-ий довірчий інтервал для теоретичного розкиду I.Q. генеральної сукупності всіх студентів університету.

Розв'язання.

Припускається, що показник I.Q. мають гауссів розподіл $N(\mu, \sigma^2)$, де μ та σ^2 є невідомими. Значить, будемо використовувати довірчі інтервали, використовуючи цей факт.

Також відомо, що $\overline{X}_n = 110$ та $\sqrt{\hat{s}_n^2} = 5$ та $\sqrt{n} = \sqrt{100} = 10$.

1. Побудуємо довірчий інтервал для μ з рівнем довіри $p = 0.95$. У випадку невідомої σ^2, використовуємо інтервал вигляду:
$$\overline{X}_n - t_p\sqrt{\frac{\hat{s}_n^2}{n}} \leq \mu \leq \overline{X}_n + t_p\sqrt{\frac{\hat{s}_n^2}{n}},$$
де $t_p = Q^{T(n-1)}((1+p)/2) = Q^{T(99)}(0.975) \approx 1.9842$. Підрахуємо кінці інтервалу за вибіркою, підставивши відомі значення величин:
$$\overline{X}_n - t_p\sqrt{\frac{\hat{s}_n^2}{n}} \approx 110 - 1.9842 \cdot \frac{5}{10} = 109.0079$$
$$\overline{X}_n + t_p\sqrt{\frac{\hat{s}_n^2}{n}} \approx 110 + 1.9842 \cdot \frac{5}{10} = 110.9921$$
Отже, за отриманими даними довірчий інтервал для μ має кінці (109.0079, 110.9921).
2. Побудуємо довірчий інтервал для σ^2 з рівнем довіри $p = 0.95$. У випадку невідомого μ, використовуємо інтервал вигляду:
$$\frac{(n-1)\hat{s}_n^2}{h_+} \leq \sigma^2 \leq \frac{(n-1)\hat{s}_n^2}{h_-},$$
де $h_- = Q^{\chi^2_{n-1}}((1-p)/2) = Q^{\chi^2_{99}}(0.025) \approx 73.3611$ та $h_+ = Q^{\chi^2_{n-1}}(1-(1-p)/2) = Q^{\chi^2_{99}}(0.975) \approx 128.422$. Підрахуємо кінці інтервалу:
$$\frac{(n-1)\hat{s}_n^2}{h_+} \approx \frac{99 \cdot 25}{128.422} \approx 19.2724$$
$$\frac{(n-1)\hat{s}_n^2}{h_-} \approx \frac{99 \cdot 25}{73.3611} \approx 33.7372$$
Отже, за отриманими даними довірчий інтервал для σ^2 має кінці (19.2724, 33.7372).

2.2 Задача 2

Побудувати 99%-ий довірчий інтервал за $X = (X_1, \ldots, X_{100})$ з $X_1 \sim N(\mu, \sigma^2)$ ...

1. ...для середнього, якщо відомо, що $\overline{X}_{100} = 10$ та $\sigma^2 = 25$,
2. ...для середнього, якщо відомо, що $\overline{X}_{100} = 10$ та $\hat{s}^2_{100} = 25$,
3. ...для дисперсії, якщо відомо, що $\hat{\sigma}^2_{100} = 9$ (μ відоме),
4. ...для дисперсії, якщо відомо, що $\hat{s}^2_{100} = 9$.

Розв'язання.

Тут $p = 0.99$, $n = 100$.

1. Використаємо
$$\overline{X}_n - x_p \frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X}_n + x_p \frac{\sigma}{\sqrt{n}},$$
де $x_p = Q^{N(0,1)}((1+p)/2) \approx 2.5758$. Маємо такі кінці інтервалу:
$$(8.7121,\ 11.2879)$$

2. Використаємо
$$\overline{X}_n - t_p \sqrt{\frac{\hat{s}^2_n}{n}} \le \mu \le \overline{X}_n + t_p \sqrt{\frac{\hat{s}^2_n}{n}},$$
де $t_p = Q^{T(n-1)}((1+p)/2) \approx 2.6264$. Маємо такі кінці інтервалу:
$$(8.6868,\ 11.3132)$$

3. Використаємо
$$\frac{n\hat{\sigma}^2_n}{h_+} \le \sigma^2 \le \frac{n\hat{\sigma}^2_n}{h_-},$$
де $h_- = Q^{\chi^2_n}((1-p)/2) \approx 67.3276$ та $h_+ = Q^{\chi^2_n}(1-(1-p)/2) \approx 140.1695$. Маємо такі кінці інтервалу:
$$(6.4208,\ 13.3675)$$

4. Використаємо
$$\frac{(n-1)\hat{s}^2_n}{h_+} \le \sigma^2 \le \frac{(n-1)\hat{s}^2_n}{h_-},$$
де $h_- = Q^{\chi^2_{n-1}}((1-p)/2) \approx 66.5101$ та $h_+ = Q^{\chi^2_{n-1}}(1-(1-p)/2) \approx 138.9868$. Маємо такі кінці інтервалу:
$$(6.4754,\ 13.5318)$$

2.3 Задача 3

Дві лабораторії визначали концентрацію сірки в дизельному паливі за стандартним зразком. Вісім вимірювань в першій лабораторії дали результати:

$$0.869, 0.874, 0.867, 0.875, 0.870, 0.869, 0.864, 0.872.$$

Після десяти вимірювань другої лабораторії було отримано:

$$0.865, 0.870, 0.866, 0.871, 0.868, 0.870, 0.871, 0.870, 0.869, 0.874.$$

Істинне значення концентрації сірки в стандартному зразку дорівнює 0.870. Вважаючи систематичні похибки нульовими, знайти довірчий інтервал для відношення двох невідомих дисперсій. Випадкові похибки вважаємо однаково розподіленими нормальними величинами. Надійність інтервалу становить 0.95.

Розв'язання.

Тут $p = 0.95$.

Модель експерименту описується двома незалежними кратними вибірками $X = (X_1, \ldots, X_{n_X})$ та $Y = (Y_1, \ldots, Y_{n_Y})$, де $X_j = \mu + \varepsilon_j^X$, $\varepsilon_j^X \sim N(0, \sigma_X^2)$ та $Y_j = \mu + \varepsilon_j^Y$, $\varepsilon_j^Y \sim N(0, \sigma_Y^2)$, $\mu = 0.870$. Тобто

$$X_i \sim N(\mu, \sigma_X^2),\ Y_j \sim N(\mu, \sigma_Y^2).$$

Побудуємо довірчий інтервал для відношення σ_X^2/σ_Y^2 на основі величини

$$\hat{F} := \frac{\sigma_Y^2}{\sigma_X^2} \cdot \frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2},\ \left(\hat{\sigma}_X^2 = \frac{1}{n} \sum_{j=1}^{n_X} (X_j - \mu)^2,\ \hat{\sigma}_Y^2 = \frac{1}{n} \sum_{j=1}^{n_Y} (Y_j - \mu)^2\right)$$

яка має розподіл Фішера з кількістю ступенів вільності $n_X = 8$ чисельника та $n_Y = 10$ знаменника відповідно. Використовуючи міркування, схожі з тими що були у попередніх викладках, маємо довірчий інтервал з кінцями:

$$\frac{1}{h_+} \cdot \frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2} \leq \frac{\sigma_X^2}{\sigma_Y^2} \leq \frac{1}{h_-} \cdot \frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2},$$

де $h_- = Q^{F(n_X, n_Y)}((1-p)/2) \approx 0.2328$ та $h_+ = Q^{F(n_X, n_Y)}(1 - (1-p)/2) \approx 3.8549$. За наявними даними, $\hat{\sigma}_X^2/\hat{\sigma}_Y^2 \approx 1.7969$. Тепер підрахуємо кінці інтервалу:

$$\frac{1}{h_+} \cdot \frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2} \approx 0.4661,$$
$$\frac{1}{h_-} \cdot \frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2} \approx 7.7178.$$

Отже, за отриманими даними довірчий інтервал для σ_X^2/σ_Y^2 має кінці $(0.4661,\ 7.7178)$.

Додатково. Як рахувати ваші дивні квантилі?

Два підходи. Один може допомогти, коли у вас в руках є лише ручка, папір та калькулятор, а інший – за наявності комп'ютера.

1. **Табульовані значення.** У збірнику задач [1] починаючи з ст. 359 ви знайдете статистичні таблиці – табульовані значення квантилів спеціальних розподілів.

   Якщо раптом для x_* значення $f(x_*)$ відсутнє в таблиці, можна спробувати отримати наближення на основі лінійної інтерполяції:

   $$f(x_*) \approx f(x') \cdot \frac{x'' - x_*}{x'' - x'} + f(x'') \cdot \frac{x_* - x'}{x'' - x'},\ x' < x_* < x''.$$

   Або ж використати кращий підхід до наближення.

2. **Програмні методи R, Python.** Якщо ви маєте доступ до комп'ютера (якщо не маєте, купіть), то можна обчислити квантилі імовірнісного розподілу доклавши трохи зусиль у програмуванні:

   - Якщо ви дружите з Python, то вам допоможе бібліотека SciPy. Наведемо приклад обчислення квантиля стандартного нормального розподілу рівня 0.975:

     ```
     > import scipy
     > p = 0.975
     > norm_dist = scipy.stats.norm(loc=0, scale=1)
     > print(norm_dist.ppf(p))
     1.959963984540054
     ```

   - Якщо ви маєте навички в R, то обчислення квантилів розподілу реалізовано в вбудованих функцій.

Назва розподілу	Параметри розподілу	Функція в R
Нормальний $N(\mu, \sigma^2)$	$\mu \in \mathbb{R}$, $\sigma^2 > 0$	qnorm(p, mean=μ, sd=σ)
Хі-квадрат χ^2_k	$k = 1, 2, 3, \ldots$	qchisq(p, df=k)
Стьюдента $T(n)$	$n = 1, 2, 3, \ldots$	qt(p, df=n)
Фішера $F(n, d)$	$n, d = 1, 2, 3, \ldots$	qf(p, df1=n, df2=d)

     Табл. 1: Деякі вбудовані функції R для обчислення квантилів розподілу.

     Приклад обчислення квантиля стандартного нормального розподілу рівня 0.975:

     ```
     > p <- 0.975
     > qnorm(p)
     [1] 1.959964
     ```

Література

[1] В.В. Голомозий, М.В. Карташов , О.Г. Кукуш, С.В. Кушніренко, Р.Є. Майборода, Ю.С. Мішура, К.В. Ральченко, Г.М. Шевченко. *Збірник задач з теорії ймовірностей та математичної статистики* (клікабельне посилання)