**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**
**НАВЧАЛЬНО-НАУКОВИЙ КОМПЛЕКС**
**"ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ"**
**НАЦІОНАЛЬНОГО ТЕХНІЧНОГО УНІВЕРСИТЕТУ УКРАЇНИ**
**"КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ"**

РОЗРАХУНКОВА ГРАФІЧНА РОБОТА
З КУРСУ "МАТЕМАТИЧА СТАТИСТИКА"
Варіант: 104

Виконав:
Вітковськй Д.О.

Прийняв:
Ільєнко А.Б.

Київ - 2021

# Зміст

**Завдання.**

1) Провести первинний аналіз вибірки. Це включає статистичний ряд, емпіричну функцію розподілу, її графік, полігон частот, гістограму, box-and-whiskers plot.

2) Знайти вибіркове середнє, вибіркову дисперсію, виправлену вибіркову дисперсію, вибіркову медіану, вибіркову моду, вибіркові коефіцієнти асиметрії та ексцесу.

3) Обґрунтувати та висунути (нову) гіпотезу про розподіл генеральної сукупності.

4) Методом моментів та методом максимальної вірогідності знайти оцінки параметрів розподілу.

5) Для кожного з параметрів кращу з цих двох оцінок перевірити на (асимптотичну незміщеність), консистентність та ефективність.

6) Побудувати довірчі інтервали надійністю 0.95 для параметрів розподілу.

7) Перевірити висунуту гіпотезу про розподіл генеральної сукупності за допомогою критерію $\chi^2$. Якщо гіпотеза суперечить вибірковим даним, перейти до п. 3.

8) Зробити висновки.

Вихідна реалізація вибірки:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| -1.66 | -0.10 | 2.84 | 5.04 | 1.60 | -2.80 | -0.72 | 5.47 | -4.20 | 0.67 |
| -2.49 | -2.40 | 5.99 | -3.26 | 2.16 | 0.51 | -3.49 | 5.10 | -1.24 | 2.58 |
| -1.06 | -0.84 | 4.09 | 6.54 | -2.96 | 6.77 | 7.29 | -2.87 | 2.29 | 1.37 |
| 2.51 | 1.67 | -4.48 | -1.46 | 0.84 | -2.01 | -1.07 | 0.01 | -4.33 | 3.41 |
| 0.49 | -3.41 | 9.65 | -3.84 | 6.15 | 8.17 | -1.55 | -3.60 | -2.73 | **18.49** |
| 1.38 | 6.09 | -2.15 | 9.68 | -0.47 | 7.67 | -1.47 | 3.30 | 4.58 | 0.43 |
| 2.19 | 5.19 | 1.95 | -4.44 | -0.22 | **-4.49** | -3.06 | 1.09 | -1.03 | -1.18 |
| 3.52 | 2.15 | -3.48 | 3.64 | -3.21 | -0.82 | -3.29 | -2.09 | 2.81 | -0.92 |
| 0.24 | 1.39 | -4.22 | 1.20 | -2.68 | -1.93 | 1.49 | -3.64 | 1.58 | 1.59 |
| 0.66 | -2.04 | 3.41 | 2.69 | 3.73 | -2.11 | 4.02 | -0.66 | 1.25 | -3.07 |

Відсортована реалізація вибірки:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| -4.49 | -4.48 | -4.44 | -4.33 | -4.22 | -4.20 | -3.84 | -3.64 | -3.60 | -3.49 |
| -3.48 | -3.41 | -3.29 | -3.26 | -3.21 | -3.07 | -3.06 | -2.96 | -2.87 | -2.80 |
| -2.73 | -2.68 | -2.49 | -2.40 | -2.15 | -2.11 | -2.09 | -2.04 | -2.01 | -1.93 |
| -1.66 | -1.55 | -1.47 | -1.46 | -1.24 | -1.18 | -1.07 | -1.06 | -1.03 | -0.92 |
| -0.84 | -0.82 | -0.72 | -0.66 | -0.47 | -0.22 | -0.1 | 0.01 | 0.24 | 0.43 |
| 0.49 | 0.51 | 0.66 | 0.67 | 0.84 | 1.09 | 1.20 | 1.25 | 1.37 | 1.38 |
| 1.39 | 1.49 | 1.58 | 1.59 | 1.60 | 1.67 | 1.95 | 2.15 | 2.16 | 2.19 |
| 2.29 | 2.51 | 2.58 | 2.69 | 2.81 | 2.84 | 3.30 | 3.41 | 3.41 | 3.52 |
| 3.64 | 3.73 | 4.02 | 4.09 | 4.58 | 5.04 | 5.10 | 5.19 | 5.47 | 5.99 |
| 6.09 | 6.15 | 6.54 | 6.77 | 7.29 | 7.67 | 8.17 | 9.65 | 9.68 | 18.49 |

# 1. Первинний аналіз вибірки.

Оскільки з даних видно, що розподіл неперервний, то для побудови статистичного ряду дані з вибірки треба розбити на рівновеликі інтервали. Емпіричною формулою кількості таких інтервалів є формула Стерджеса:

$$m = 1 + \log_2 n = 1 + \log_2 100 = 7.644 \approx 8. \tag{1}$$

Але мені така кількість видається замалою, тому використаю значення

$$m = 12. \tag{2}$$

Отже
$x_{min} = \min\{x\} = -4.49$;
$x_{max} = \max\{x\} = 18.49$.
Розмах вибірки
$R = x_{max} - x_{min} = 22.98$.
Таким чином довжина кроку дорівнює 1.915.

| Інтервал | $[-4.49, -2.575)$ | $[-2.575, -0.66)$ | $[-0.66, 1.255)$ | $[1.255, 3.17)$ |
|---|---|---|---|---|
| Середина інтервалу | $-3.525$ | $-1.6175$ | 0.2975 | 4.425 |
| Частота | 22 | 21 | 15 | 18 |
| Кумулятивна частота | 22 | 43 | 58 | 76 |
| Відносна частота | 0.22 | 0.21 | 0.15 | 0.18 |
| Кумулятивна відносна частота | 0.22 | 0.43 | 0.58 | 0.76 |

Таблиця 1. Статистичний ряд.

| Інтервал | $[3.17, 5.085)$ | $[5.085, 7)$ | $[7, 8.915)$ | $[8.915, 10.83)$ |
|---|---|---|---|---|
| Середина інтервалу | 4.1275 | 6.0425 | 7.9575 | 9.8725 |
| Частота | 10 | 8 | 3 | 2 |
| Кумулятивна частота | 86 | 94 | 97 | 99 |
| Відносна частота | 0.1 | 0.08 | 0.03 | 0.02 |
| Кумулятивна відносна частота | 0.86 | 0.91 | 0.97 | 0.99 |

| Інтервал | $[10.83, 12.745)$ | $[12.745, 14.66)$ | $[14.66, 16.575)$ | $[16.575, 18.49]$ |
|---|---|---|---|---|
| Середина інтервалу | 11.7875 | 13.7025 | 15.6175 | 17.5325 |
| Частота | 0 | 0 | 0 | 1 |
| Кумулятивна частота | 99 | 99 | 99 | 100 |
| Відносна частота | 0 | 0 | 0 | 0.01 |
| Кумулятивна відносна частота | 0.99 | 0.99 | 0.99 | 1 |

Таблица 2. Статистичний ряд (продовження).

Емпірична функція розподілу (інтервальна):

$$F_n^*(x) = \begin{cases} 0, & x < x_0^*; \\ \omega_k^{\text{кумул}}, & x_k^* \leq x < x_{k+1}^*. \end{cases} \tag{3}$$

, де $x_k^*$ – початок k-того інтервалу, кінець (k-1)-го, а $\omega_k^{\text{кумул}}$ – кумулятивна відносна частота для k-того інтервалу.

Відповідна функція для даного інтервального розбиття:

$$F_n^*(x) = \begin{cases} 0, & x < -4.49; \\ 0.22, & x \in [-4.49, -2.575); \\ 0.43, & x \in [-2.575, -0.66); \\ 0.58, & x \in [-0.66, 1.255); \\ 0.76, & x \in [1.255, 3.17); \\ 0.86, & x \in [3.17, 5.085); \\ 0.91, & x \in [5.085, 7); \\ 0.97, & x \in [7, 8.915); \\ 0.99, & x \in [8.915, 16.575); \\ 1, & x \geq 16.575. \end{cases} \tag{4}$$

Її графік:

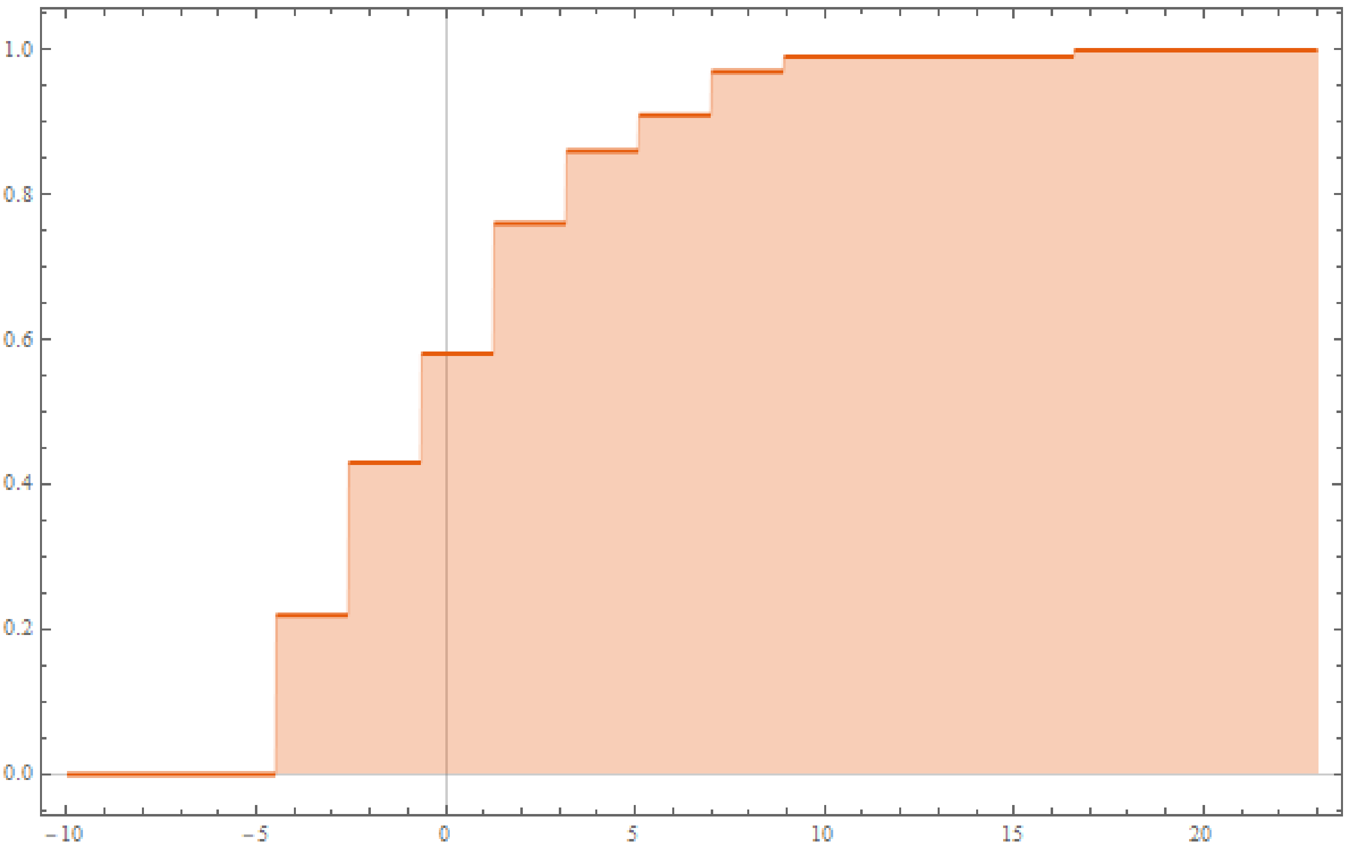


Рис. 1. Інтервальна емпірична функція розподілу.

Гістограма за заданими інтервалами:

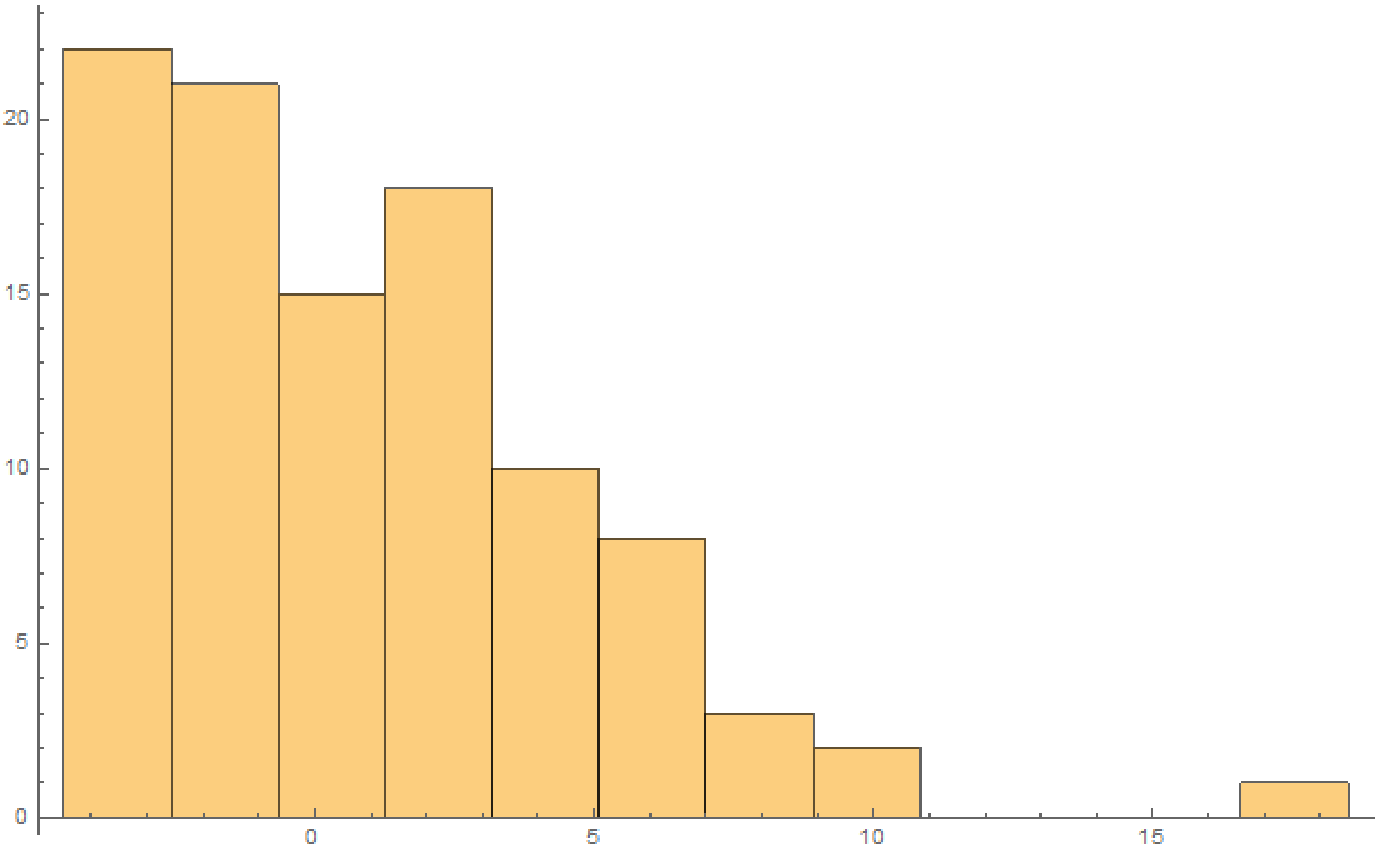


Рис. 2. Гістограма розподілу.

Для побудови графіку "ящик з вусами"використано значення, обчислені у наступному розділі, а саме: вибіркове середнє, медіану, перший та третій квартилі і мінімум та максимум як значення кінців вусів.

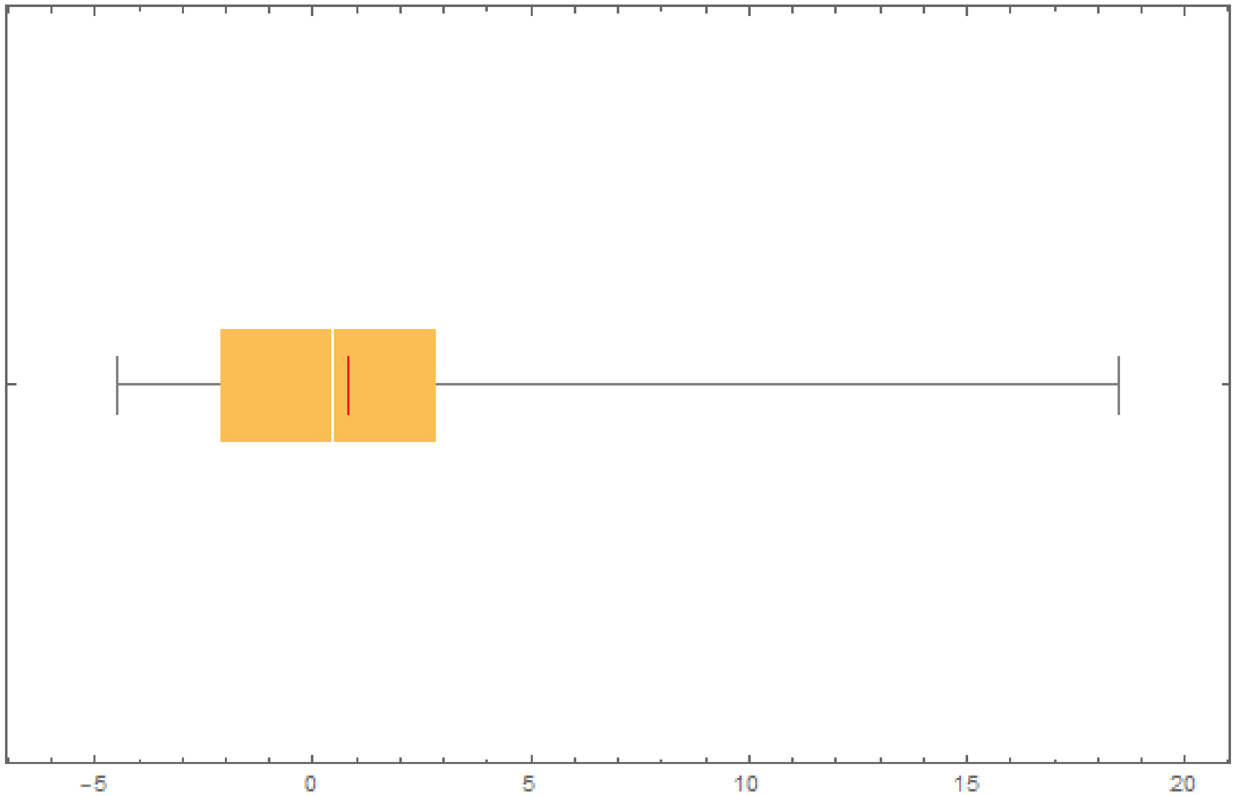


Рис. 3. Ящик з вусами.

# 2. Знаходження описових статистик.

Вибіркове середнє:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{100}\sum_{i=1}^{100} x_i = 0.7938 \tag{5}$$

Вибіркова дисперсія:

$$s^2 = \mathbb{D}_{\xi}^{**} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{100}\sum_{i=1}^{100}(x_i - 0.7938)^2 = 15.0129 \tag{6}$$

Виправлена вибіркова дисперсія:

$$s_0^2 = \mathbb{D}_{\xi}^{***} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{99}\sum_{i=1}^{100}(x_i - 0.7938)^2 = 15.1645 \tag{7}$$

Вибіркова медіана:

Оскільки маємо відсортований набір усіх $x$, можна доволі легко обчислити медіану за наступною формулою:

$$\mathcal{M}_e^* = \begin{cases} x_{(k+1)}, & n = 2k+1 \\ \frac{x_{(k)} + x_{(k+1)}}{2}, & n = 2k \end{cases}$$

, де $k \in \mathbb{N}$
Тоді у нашому випаддку медіана обчислюється наступним чином:

$$\mathcal{M}_e^* = \frac{x_{(k)} + x_{(k+1)}}{2} = \frac{0.43 + 0.49}{2} = 0.46 \tag{8}$$

Вибіркова мода:

$$mo = 1 \text{— номер модального класу} \tag{9}$$

$$\mathcal{M}_o^* = y_{mo-1} + (y_{mo} - y_{mo-1}) \cdot \frac{n_{mo} - n_{mo-1}}{(n_{mo} - n_{mo-1}) + (n_{mo} - n_{mo+1})} \tag{10}$$

, де $y_{mo-1}$, $y_{mo}$ – відповідно нижня та верхня межі модального класу; $n_{mo}$, $n_{mo-1}$, $n_{mo+1}$ – частоти відповідно модального, передмодального та післямодального інтервалів.

$$\boxed{=} -4.49 + (-2.575 - (-4.49)) \cdot \frac{22-0}{(22-0)+(22-21)} = -2.65826$$

Вибіркові коефіцієнти асиметрії та екцесу:

Для знаходження цих коефіцієнтів потрібно обчислювати вибіркові центральні моменти за наступною формулою:

$$\overline{\mu_k} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^k \tag{11}$$

Вибірковий коефіцієнт асиметрії:

$$As = \frac{\overline{\mu_3}}{s_0^3} = \frac{\frac{1}{100}\sum\limits_{i=1}^{100}(x_i - \overline{x})^3}{\left(\mathbb{D}_\xi^{***}\right)^{\frac{3}{2}}} = 1.22459 \tag{12}$$

Вибірковий коефіцієнт ексцесу:

$$Ek = \frac{\overline{\mu_4}}{s_0^4} - 3 = \frac{\frac{1}{100}\sum\limits_{i=1}^{100}(x_i - \overline{x})^4}{\left(\mathbb{D}_\xi^{***}\right)^{2}} - 3 = 2.822 \tag{13}$$

Перший квартиль:

$$\begin{gathered}\alpha = 0.25 \Rightarrow K = n \cdot \alpha = 100 * 0.25 = 25 \in \mathbb{Z} \Rightarrow \\ \Rightarrow Q_1 = q_\alpha = \frac{x_{(K)} + x_{(K+1)}}{2} = \frac{-2.15 - 2.11}{2} = -2.13\end{gathered} \tag{14}$$

Третій квартиль:

$$\begin{gathered}\alpha = 0.75 \Rightarrow K = n \cdot \alpha = 100 * 0.75 = 75 \in \mathbb{Z} \Rightarrow \\ \Rightarrow Q_1 = q_\alpha = \frac{x_{(K)} + x_{(K+1)}}{2} = \frac{2.81 - 2.84}{2} = 2.825\end{gathered} \tag{15}$$

## 3. Гіпотеза про розподіл

Дані вибірки майже не мають повторень та не є цілими числами, тому розподіл є неперервним. Вибірка найбільше схожа на зміщену експоненційно розподілену, оскільки має найбільшу частоту у лівій своїй частині яка швидко зменшується при русі у додатному напрямку. Її гістограма та емпірична функція розподілу схожі на відповідні функції для зміщеного на 4.49 у від'ємному напрямку експоненціного розподілу.

Таким чином висунемо гіпотезу що ГС розподілена за експоненційним законом зі зміщенням $a$ та відповідним параметром $\frac{1}{\lambda}$, де $\lambda = \mathbb{E}_\xi - a$:

$$H_0 = \left\{ \xi \sim Exp(\lambda, a) = Exp(\frac{1}{\lambda}) + a \right\}$$

$$f_{Exp(\lambda,a)}(x) = \begin{cases} \frac{1}{\lambda} e^{-\frac{x-a}{\lambda}}, & x \geq a \\ 0, & x < a \end{cases}$$

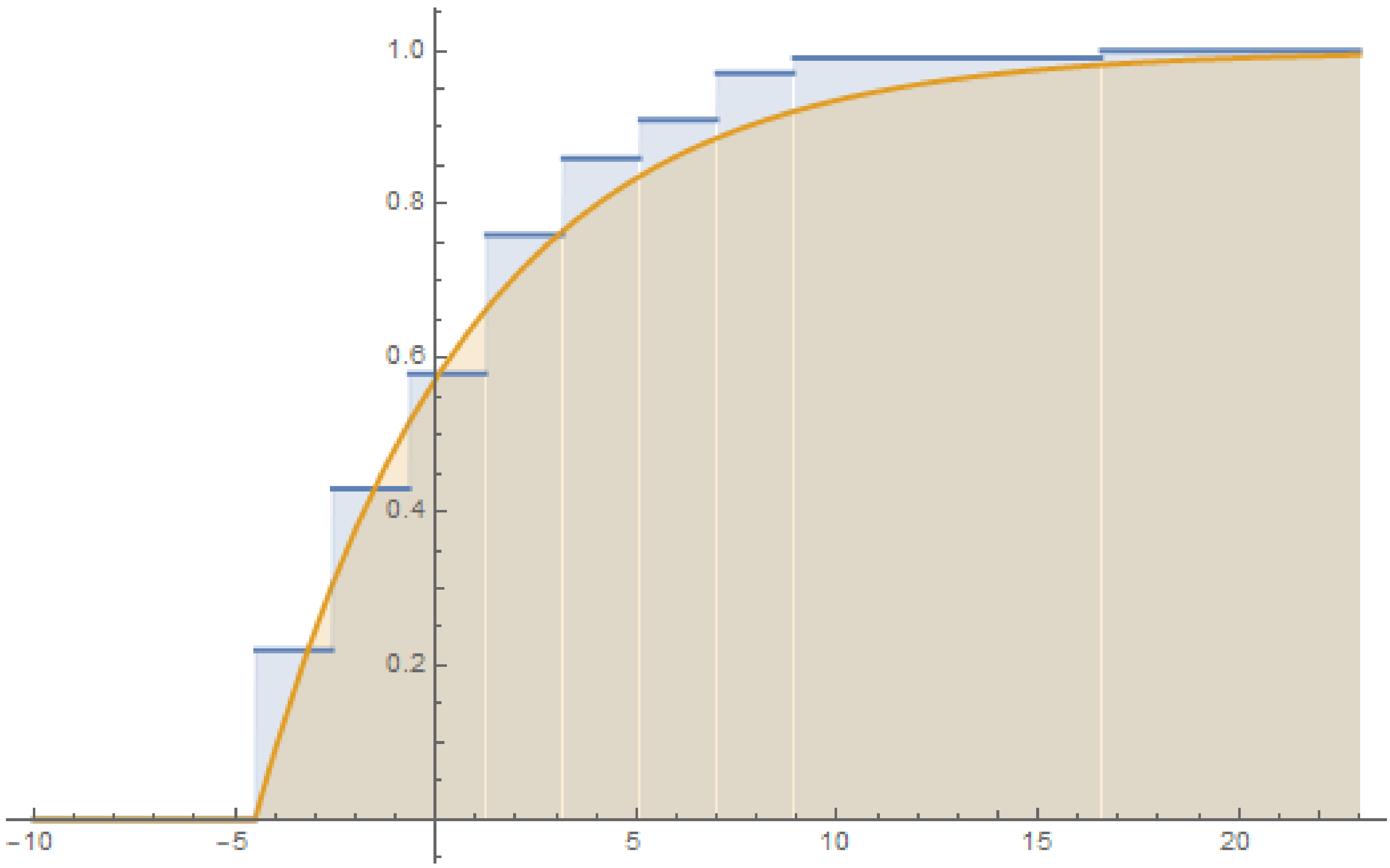


Рис. 4. Емпірична функція розподілу у порівнянні з відповідною функцією розподілу.

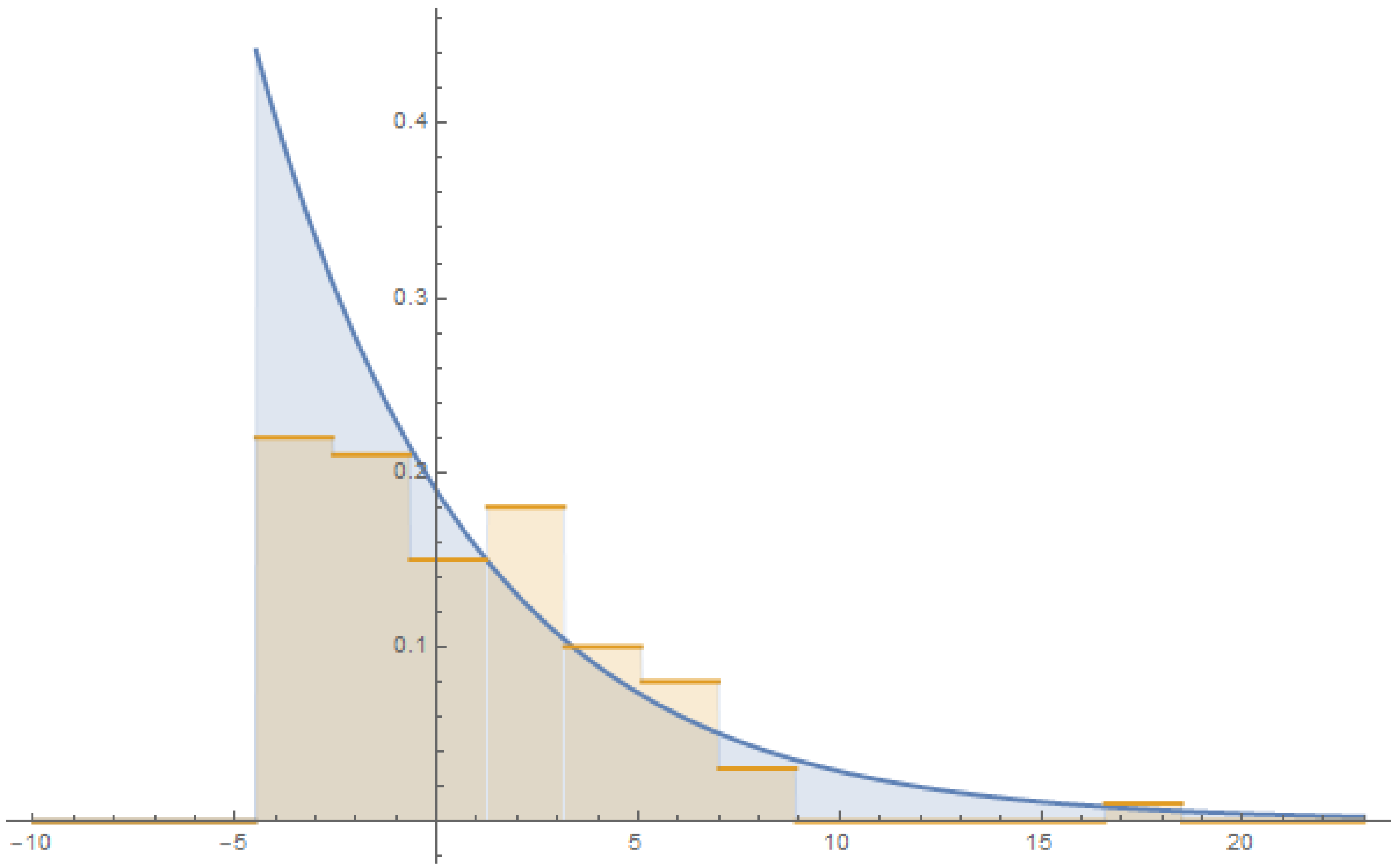

Рис. 5. Гістограма у порівнянні із відповідною функцією щільності.

Тоді альтернативною гіпотезою буде $H_1 = \{\xi \sim Exp(\lambda, a)\}$.

# 4. Оцінка параметрів.

Для зміщеного експоненційного розподівлу маємо параметри $\lambda$ та $a$. Треба їх оцінити.

## Метод моментів.

Для знаходження оцінки методом моментів, треба розв'язати систему рівнянь, що включає s моментів розподілу, де s – кількість параметрів розподілу, а також їх емпіричні еквіваленти.

$$\mathbb{E}\xi = \int\limits_{-\infty}^{+\infty} x f_\xi(x)\,dx = \int\limits_{a}^{+\infty} x \cdot \frac{1}{\lambda} e^{-\frac{x-a}{\lambda}}\,dx = \frac{1}{\lambda}\int\limits_{a}^{+\infty} x e^{-\frac{x-a}{\lambda}}\,dx =$$

$$= \left| \begin{array}{cc} u = x; & dv = e^{-\frac{x-a}{\lambda}}\,dx \\ du = dx; & v = -\lambda e^{-\frac{x-a}{\lambda}} \end{array} \right| = \frac{1}{\lambda}\left( -\lambda x e^{-\frac{x-a}{\lambda}} \Big|_a^{+\infty} + \lambda \int\limits_a^{+\infty} e^{-\frac{x-a}{\lambda}}\,dx \right) =$$

$$= \frac{1}{\lambda}\left(0 - (-\lambda a) - \lambda^2 e^{-\frac{x-a}{\lambda}}\Big|_a^{+\infty}\right) = \frac{1}{\lambda}\big(\lambda a - \lambda^2(0-1)\big) = \boxed{a+\lambda}; \tag{16}$$

$$\mathbb{E}\xi^2 = \int\limits_{-\infty}^{+\infty} x^2 \cdot f_\xi(x)\,dx = \frac{1}{\lambda}\int\limits_a^{+\infty} x^2 e^{-\frac{x-a}{\lambda}}\,dx = \left| \begin{array}{cc} u = x^2; & dv = e^{-\frac{x-a}{\lambda}}\,dx \\ du = 2x\,dx; & v = -\lambda e^{-\frac{x-a}{\lambda}} \end{array} \right| =$$

$$\frac{1}{\lambda}\left( -\lambda x^2 e^{-\frac{x-a}{\lambda}}\Big|_a^{+\infty} + 2\lambda \int\limits_a^{+\infty} x e^{-\frac{x-a}{\lambda}}\,dx \right) = \frac{1}{\lambda}\left(\lambda a^2 + 2\lambda\left(\lambda a + \lambda^2\right)\right) =$$

$$= a^2 + 2\lambda a + 2\lambda^2 = \boxed{(a+\lambda)^2 + \lambda^2} \tag{17}$$

Отримуємо наступну систему рівнянь:

$$\begin{cases} a^* + \lambda^* = \overline{\xi} \\ (a^* + \lambda^*)^2 + (\lambda^*)^2 = \overline{\xi^2} \end{cases}$$

$$\begin{cases} \lambda^* = \sqrt{\overline{\xi^2} - \overline{\xi}^2} \approx 3.0174 \\ a^* = \overline{\xi} - \sqrt{\overline{\xi^2} - \overline{\xi}^2} \approx -2.2236 \end{cases} \tag{18}$$

## Метод максимальної вірогідності.

Для знаходження оцінки методом максимальної вірогідності треба максимізувати функцію вірогідності за параметрами $\lambda$ та $a$. Можемо спробувати знайти її глобальний максимум.

$$\mathcal{L}(x_1,\ldots,x_n,\lambda,a) = f_{\xi_1,\ldots,\xi_n}(x_1,\ldots,x_n) = \prod_{i=1}^{n} f_{\xi_i}(x_i) = \prod_{i=1}^{n} f_{\xi}(x_i) =$$

$$= |x \geq a| = \prod_{i=1}^{n} \frac{1}{\lambda} e^{-\frac{x_i - a}{\lambda}} = \frac{1}{\lambda^n} e^{-\frac{1}{\lambda}(x_1+\cdots+x_n-na)}$$

Для легшого обчислення похідних варто використати функцію логарифма, оскільки вона монотонно зростаюча, а логарифм добутку родівнює сумі логарифмів множників, що спрощує диференціювання.

$$\ln \mathcal{L} = -n \ln \lambda - \frac{x_1+\cdots+x_n-na}{\lambda}$$

Знайдемо часткові похідні першого та другого порядків за параметрами:

$$\frac{\partial \ln \mathcal{L}}{\partial \lambda} = -\frac{n}{\lambda} + \frac{x_1+\cdots+x_n-na}{\lambda^2} = \frac{x_1+\cdots+x_n-na-n\lambda}{\lambda^2};$$

$$\frac{\partial \ln \mathcal{L}}{\partial a} = \frac{n}{\lambda};$$

$$\frac{\partial^2 \ln \mathcal{L}}{\partial \lambda^2} = \frac{n}{\lambda^2} - 2\frac{x_1+\cdots+x_n-na}{\lambda^3};$$

$$\frac{\partial^2 \ln \mathcal{L}}{\partial a^2} = 0;$$

$$\frac{\partial^2 \ln \mathcal{L}}{\partial \lambda \partial a} = -\frac{n}{\lambda^2};$$

$$\det \begin{pmatrix} \frac{\partial^2 \ln \mathcal{L}}{\partial \lambda^2} & \frac{\partial^2 \ln \mathcal{L}}{\partial \lambda \partial a} \\ \frac{\partial^2 \ln \mathcal{L}}{\partial \lambda \partial a} & \frac{\partial^2 \ln \mathcal{L}}{\partial a^2} \end{pmatrix} = 0 - \left(-\frac{n}{\lambda^2}\right)^2 < 0$$

Отже функція не має екстремумів, тому принаймні спробуємо максимізувати її значення.

$$\begin{cases} \frac{\partial \ln \mathcal{L}}{\partial \lambda^*} = 0; \\ \frac{\partial \ln \mathcal{L}}{\partial a^*} = 0. \end{cases} \Rightarrow \begin{cases} a^* + \lambda^* = \overline{x}; \\ \lambda^* \neq 0; \\ \lambda^* \to \infty. \end{cases} \Rightarrow \begin{cases} a^* = \min\{x_i\} = -4.49; \\ \lambda^* = \overline{x} - a^* = 5.2838. \end{cases} \quad (19)$$

# 5. Дослідження оцінки.

Кращою з двох отриманих оцінок є оцінка методом максимальної вірогідності. Проведемо її дослідження.

**<u>Незміщеність.</u>**

$$\mathbb{E}a^* = \mathbb{E}\min\{\xi_i\}\boxed{=}$$

Щільність розподілу мінімума:

$$f_\xi(x) = \frac{1}{\lambda}e^{-\frac{x-a}{\lambda}} \cdot \mathbb{I}\{x \geq a\}; \quad F_\xi(x) = (1 - e^{-\frac{x-a}{\lambda}}) \cdot \mathbb{I}\{x \geq a\}$$

$$f_{\min\{\xi\}}(x) = n(1 - F_\xi(x))^{n-1} f_\xi(x) = \frac{n}{\lambda}e^{-n\frac{x-a}{\lambda}} \cdot \mathbb{I}\{x \geq a\}$$

$$\boxed{=} \int\limits_a^{+\infty} x \cdot \frac{n}{\lambda}e^{-n\frac{x-a}{\lambda}}\,dx = \left| \begin{array}{cc} u = x; & dv = e^{-n\frac{x-a}{\lambda}}\,dx \\ du = dx; & v = -\frac{\lambda}{n}e^{-n\frac{x-a}{\lambda}} \end{array} \right| =$$

$$\frac{n}{\lambda}\left( -\frac{\lambda}{n}xe^{-n\frac{x-a}{\lambda}}\Big|_a^{+\infty} + \frac{\lambda}{n}\int\limits_a^{+\infty} e^{-n\frac{x-a}{\lambda}}\,dx \right) = a - \frac{\lambda}{n}e^{-n\frac{(x-a)}{\lambda}}\Big|_a^{+\infty} = \frac{na+\lambda}{n} \neq a \Rightarrow$$

$$\Rightarrow a^* \text{ – не є незміщеною.} \tag{20}$$

$$\lim_{n\to\infty}\mathbb{E}a^* = \lim_{n\to\infty}\frac{na+\lambda}{n} = a \Rightarrow$$

$$\Rightarrow a^* \text{ – є асимптотично незміщеною.} \tag{21}$$

$$\mathbb{E}\lambda^* = \mathbb{E}\left(\overline{\xi} - a^*\right) = \mathbb{E}\frac{\xi_1 + \cdots + \xi_n}{n} - \frac{na+\lambda}{n} = \frac{1}{n}n\mathbb{E}\xi_i - \frac{na+\lambda}{n} =$$

$$a + \lambda - \frac{na+\lambda}{n} = \frac{n\lambda - \lambda}{n} \neq \lambda \Rightarrow$$

$$\Rightarrow \lambda^* \text{ – не є незміщеною.} \tag{22}$$

$$\lim_{n\to\infty}\mathbb{E}\lambda^* = \lim_{n\to\infty}\frac{n\lambda - \lambda}{n} = \lambda \Rightarrow$$

$$\Rightarrow \lambda^* \text{ – є асимптотично незміщеною.} \tag{23}$$

**Консистентність.**

$$\mathbb{P}\lim_{n\to\infty} a^* \stackrel{?}{=} a$$

$$\mathbb{P}\lim_{n\to\infty} \lambda^* \stackrel{?}{=} \lambda$$

Перевіримо за лемою про консистентність. Для цього потрібно обчислити дисперсії оцінок.

$$\mathbb{D}a^* = \mathbb{E}(a^*)^2 - (\mathbb{E}a^*)^2 \ ;$$

$$\mathbb{E}(a^*)^2 = \frac{n}{\lambda}\int\limits_a^{+\infty} x^2 \cdot e^{-n\frac{x-a}{\lambda}}\,dx = \begin{vmatrix} u = x^2; & dv = e^{-n\frac{x-a}{\lambda}}\,dx \\ du = 2x\,dx; & v = -\frac{\lambda}{n}e^{-n\frac{x-a}{\lambda}} \end{vmatrix} =$$

$$= -x^2 e^{-n\frac{x-a}{\lambda}}\Big|_a^{+\infty} + 2\int\limits_a^{+\infty} x e^{-n\frac{x-a}{\lambda}}\,dx = a^2 + 2\left(-\frac{\lambda}{n} x e^{-n\frac{x-a}{\lambda}}\Big|_a^{+\infty} + \frac{\lambda}{n}\int\limits_a^{+\infty} e^{-n\frac{x-a}{\lambda}}\,dx\right) =$$

$$= a^2 + 2\cdot\frac{\lambda}{n}\cdot\frac{na+\lambda}{n} = a^2 + \frac{2na\lambda + 2\lambda^2}{n^2}\,;$$

$$\mathbb{D}a^* = a^2 + \frac{2na\lambda + 2\lambda^2}{n^2} - \frac{(na+\lambda)^2}{n^2} = \frac{\lambda^2}{n^2}\,; \qquad (24)$$

$$\begin{cases} \lim\limits_{n\to\infty} \mathbb{D}a^* = a^2 + 0 - a^2 = 0\,; \\ a^* \text{ – асимптотично незміщена.} \end{cases} \Rightarrow$$

$$\Rightarrow a^* \text{ – консистентна} \qquad (25)$$

$$\mathbb{D}\lambda^* = \mathbb{D}(\overline{\xi} - a^*) = \mathbb{D}\overline{\xi} + \mathbb{D}a^* = \mathbb{D}\frac{\xi_1 + \cdots + \xi_n}{n} + \mathbb{D}a^* = \frac{1}{n^2}\mathbb{D}(\xi_1 + \cdots + \xi_n) + \mathbb{D}a^* =$$

$$= \frac{1}{n}\mathbb{D}\xi_i + \mathbb{D}a^* = \frac{1}{n}\left(\mathbb{E}\xi^2 - (\mathbb{E}\xi)^2\right) + \mathbb{D}a^* = \frac{1}{n}\left(\cancel{(a+\lambda)^2} + \lambda^2 - \cancel{(a+\lambda)^2}\right) + \frac{\lambda^2}{n^2} =$$

$$= \frac{(n+1)\lambda^2}{n^2} \qquad (26)$$

$$\begin{cases} \lim\limits_{n\to\infty} \mathbb{D}\lambda^* = \lim\limits_{n\to\infty} \frac{(n+1)\lambda^2}{n^2} = 0\,; \\ \lambda^* \text{ – асимптотично незміщена.} \end{cases} \Rightarrow$$

$$\Rightarrow \lambda^* \text{ – консистентна.} \qquad (27)$$

**Ефективність.**

Оскільки оцінки виявилися зміщеними, перевіряти їх на ефективність немає сенсу.

# 6. Інтервальні оцінки параметрів.

Задача: побудувати інтервальні оцінки параметрів розподілу з довірчою імовірністю $\gamma = 0.95$.
Оскільки розподіл нестандартний, доволі важко буде знайти точний довірчий інтервал, але при великих n доволі непоганої точності можна досягти і з використанням асимптотичних довірчих інтервалів, тому знайдемо саме такий інтервал.
**Озн.** $(\theta^*_{1n}, \theta^*_{2n})$ називається асимптотичним довірчим інтервалом для параметру $\theta$, якщо $\lim\limits_{n\to\infty} \mathbb{P}\{\theta^*_{1n} < \theta < \theta^*_{2n}\} \overset{(\geq)}{=} \gamma$.

При великих $n$

$$\frac{\theta^* - \mathbb{E}\theta^*}{\sqrt{\mathbb{D}\theta^*}} \approx \mathcal{N}(0,1)$$

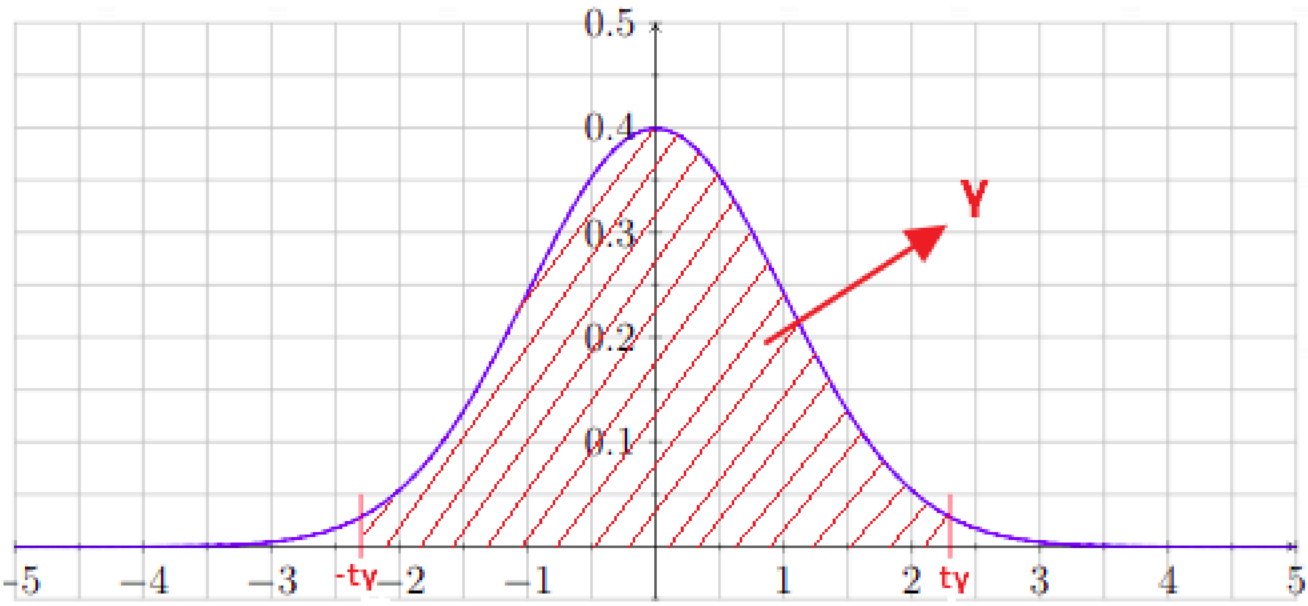


Рис. 6. Стандартний нормальний розподіл.

$$t_\gamma : \Phi(t_\gamma) = \frac{\gamma}{2} \tag{28}$$

$$\mathbb{P}\{-t_\gamma < \mathcal{N}(0,1) < t_\gamma\} = \gamma$$

$$\mathbb{P}\{-t_\gamma < \frac{\theta^* - \mathbb{E}\theta^*}{\sqrt{\mathbb{D}\theta^*}} < t_\gamma\} = \gamma \tag{29}$$

Для знаходження асимптотичної інтервальної оцінки параметра $\theta$ потрібно розв'язати нерівніть під оператором $\mathbb{P}$ відносно $\theta$.
Для $\gamma = 0.95$ маємо значення $t_\gamma = 1.95996$.

**Асимптотичний інтервал для параметра зміщеного середнього $\lambda$.**

Оскільки для даної оцінки рахувати дисперсію потрібно через коваріацію, бо її компоненти залежні, то можна спробувати трохи знизити точніть, але спростити обчислення – прийняти $a^*$ відомим, підставивши оцінене раніше обчислену оцінку ММВ.

$$\frac{\lambda^*-\mathbb{E}\lambda^*}{\sqrt{\mathbb{D}\lambda^*}} = \left| \begin{array}{c} \mathbb{D}\lambda^* = \mathbb{D}(\overline{\xi}-a) = \mathbb{D}\overline{\xi} = \frac{1}{n}\mathbb{D}\xi_i = \frac{1}{n}\mathbb{D}\xi = \\ = \frac{1}{n}\left(\mathbb{E}\xi^2-(\mathbb{E}\xi)^2\right) = \frac{1}{n}\left((a+\lambda)^2+\lambda^2-(a+\lambda)^2\right) = \frac{\lambda^2}{n} \end{array} \right| =$$

$$= \frac{\lambda^* - \frac{(n-1)\lambda}{n}}{\lambda}\sqrt{n} = \frac{n\lambda^*-(n-1)\lambda}{n\lambda}\sqrt{n} = \frac{n\lambda^*-(n-1)\lambda}{\sqrt{n}\lambda}\,;$$

$$-t_\gamma < \frac{n\lambda^*-(n-1)\lambda}{\sqrt{n}\lambda} < t_\gamma \quad \Rightarrow \quad \frac{(n\lambda^*-(n-1)\lambda)^2}{n\lambda^2} < t_\gamma^2 \Rightarrow$$

$$\Rightarrow (n-1)^2\lambda^2 - 2(n^2-n)\lambda^*\lambda + n^2(\lambda^*)^2 < nt_\gamma^2\lambda^2 \Rightarrow$$

$$\Rightarrow (n^2-2n-nt_\gamma^2+1)\lambda^2 - 2(n^2-n)\lambda^*\lambda + n^2(\lambda^*)^2 < 0 \; \boxed{\Rightarrow}$$

$$\mathcal{D} = b^2 - 4\hat{a}c = 4(n^2-n)^2(\lambda^*)^2 - 4(n^2-2n-nt_\gamma^2+1)n^2(\lambda^*)^2 =$$

$$= 4(\lambda^*)^2(\cancel{n^4-2n^3}+n\cancel{-n^4+2n^3}+n^3t_\gamma^2-n^2) = 4n(\lambda^*)^2(n^2t_\gamma^2-n+1)\,;$$

$$\boxed{\Rightarrow}\; \lambda_{1,2}^* = \frac{-b\pm\sqrt{\mathcal{D}}}{2\hat{a}} = \frac{2(n^2-n)\lambda^* \pm 2\lambda^*\sqrt{n(n^2t_\gamma^2-n+1)}}{2(n^2-n(2-t_\gamma^2)+1)} =$$

$$= \frac{\lambda^*\left((n^2-n)\pm\sqrt{n(n^2t_\gamma^2-n+1)}\right)}{n^2-n(2-t_\gamma^2)+1} \tag{30}$$

**Асимптотичний інтервал для параметра зміщення $a$.**

$$\frac{a^* - \mathbb{E}a^*}{\sqrt{\mathbb{D}a^*}} = \frac{\min\{\xi_i\} - \frac{na+\lambda}{n}}{\frac{\lambda}{n}} = \frac{n}{\lambda}\min\{\xi_1\} - \frac{na}{\lambda} - 1\,;$$

$$-t_\gamma < \frac{n}{\lambda}\min\{\xi_1\} - \frac{na}{\lambda} - 1 < t_\gamma \Rightarrow$$

$$\Rightarrow -t_\gamma - \frac{n}{\lambda}\min\{\xi_1\} + 1 < -\frac{na}{\lambda} < t_\gamma - \frac{n}{\lambda}\min\{\xi_1\} + 1 \Rightarrow$$

$$\Rightarrow \frac{\lambda t_\gamma}{n} + \min\{\xi_i\} - \frac{\lambda}{n} > a > -\frac{\lambda t_\gamma}{n} + \min\{\xi_i\} - \frac{\lambda}{n}\,;$$

$$a \in \left(-\frac{\lambda t_\gamma}{n} + \min\{\xi_i\} - \frac{\lambda}{n}, \quad \frac{\lambda t_\gamma}{n} + \min\{\xi_i\} - \frac{\lambda}{n}\right) \tag{31}$$

Оскільки ми не маємо точного значення параметра $\lambda$, також можемо взяти його оцінку, що знизить точність, але інтервал і так був асимптотичним, тому це не сильно вплине. Отже:

$$a \in \left(-\frac{\lambda^* t_\gamma}{n} + \min\{x_i\} - \frac{\lambda}{n}, \quad \frac{\lambda^* t_\gamma}{n} + \min\{x_i\} - \frac{\lambda}{n}\right) =$$

$$= (-4.6464, \quad -4.4393).$$

Як видно, отримана методом максимальної вірогідності оцінка потрапляє у заданий проміжок, однак оцінка методом моментів – ні.