Міністерство освіти і науки України
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Інститут прикладного системного аналізу

Розрахункова робота
З дисципліни «Математична статистика»

Виконав студент 2 курсу групи КА-01
Вагін Олександр Вікторович
Перевірила
Каніовська Ірина Юріївна

Київ-2022

Завдання розрахункової роботи

1. Побудувати варіаційний (дискретний або інтервальний) ряд наданої вибірки.
2. Зробити графічне зображення вибірки.
3. Побудувати емпіричну функцію розподілу.
4. Знайти незміщену оцінку математичного сподівання та дисперсії..
5. Обчислити значення вибіркової медіани, моди, асиметрії.
6. Висунути гіпотезу про розподіл, за яким отримано вибірку.
7. Знайти точкові оцінки параметрів гіпотетичного закону розподілу та перевірити їх властивості.
8. Перевірити за допомогою критерію χ^2 (Пірсона) гіпотезу про розподіл з рівнем значущості $\alpha = 0{,}05$.
9. Знайти довірчий інтервал для параметрів гіпотетичного закону розподілу, взяв рівень надійності $\gamma = 0{,}95$.
10. Висновки.

Задана конкретна реалізація вибірки:

6.71	5.66	3.06	4.25	3.4	6.45	3.46	2.57	6.55	8.5
3.63	7.22	7.61	2.72	3.91	2.98	7.05	11.79	7.51	4.44
13.59	6.61	4.53	12.28	9.5	4.26	4.52	5.55	11.34	4.21
4.02	4.34	2.67	2.95	8.86	4.87	5.76	7.68	2.75	2.62
2.86	5.16	2.88	6.62	5.03	3.3	5.47	7.8	7.8	5.43
3.05	3.99	4.1	2.85	3.28	4.7	6.67	2.85	4.71	6.94
7.17	3.02	2.97	10.96	2.75	3.03	3.61	15.5	6.24	2.89
5.57	5.75	4.86	10.36	14.49	15.38	2.8	3.86	4.67	4.58
8.64	3.65	3.65	4.53	3.63	11.68	3.14	6.07	6.57	3.22
6.12	3.37	3.9	6.63	2.88	9.18	2.96	4.86	3.18	6.59

З початку відсортуємо задану реалізацію вибірки для зручності подальшої роботи:

2.57	2.62	2.67	2.72	2.75	2.75	2.8	2.85	2.85	2.86
2.88	2.88	2.89	2.95	2.96	2.97	2.98	3.02	3.03	3.05
3.06	3.14	3.18	3.22	3.28	3.3	3.37	3.4	3.46	3.61
3.63	3.63	3.65	3.65	3.86	3.9	3.91	3.99	4.02	4.1
4.21	4.25	4.26	4.34	4.44	4.52	4.53	4.53	4.58	4.67
4.7	4.71	4.86	4.86	4.87	5.03	5.16	5.43	5.47	5.55
5.57	5.66	5.75	5.76	6.07	6.12	6.24	6.45	6.55	6.57
6.59	6.61	6.62	6.63	6.67	6.71	6.94	7.05	7.17	7.22
7.51	7.61	7.68	7.8	7.8	8.5	8.64	8.86	9.18	9.5
10.36	10.96	11.34	11.68	11.79	12.28	13.59	14.49	15.38	15.5

Можемо побачити, що задана конкретна реалізація містить майже всі унікальні значення, а саме 92 унікальних значення. Тому доцільним є будувати саме інтервальний варіаційний ряд.

1. Побудувати інтервальний варіаційний ряд

Скористуємося правилом Стерджеса, щоб знайти на скільки інтервалів ділити відрізок $[x_{min}; x_{max}]$.
А саме $k = 1 + [3{,}332 \lg n] = 1 + +[3{,}332 \lg 100] = 1 + [6{,}664] = 7$.
Розмах вибірки $R = x_{max} - x_{min} = 15{,}5 - 2{,}57 = 12{,}93$.
Побудувавши програмним шляхом гістограми мені здається, що краще взяти трохи більше число інтервалів для більшої наочності, тож я візьму кількість інтервалів $k = 8$.
Тоді поділимо на 8 пів інтервалів довжиною $\frac{R}{k} = \frac{12{,}93}{8} \approx 1{,}62$
Маємо наступний інтервальний варіаційний ряд:

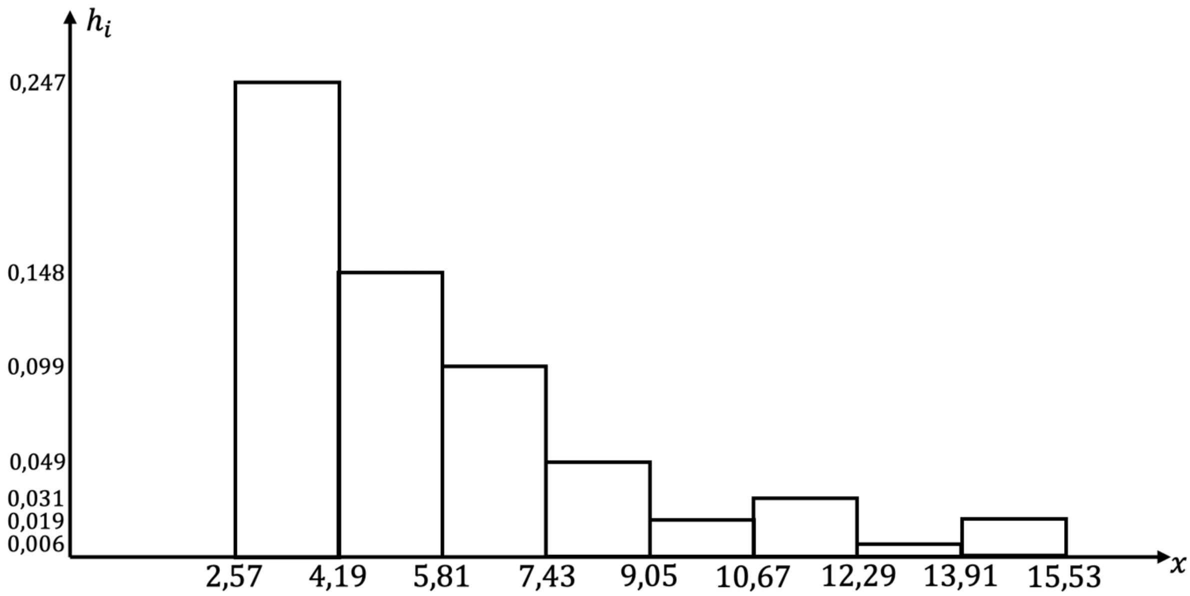
Інтервал Δ_i	$[2{,}57; 4{,}19)$	$[4{,}19; 5{,}81)$	$[5{,}81; 7{,}43)$	$[7{,}43; 9{,}05)$	$[9{,}05; 10{,}67)$	$[10{,}67; 12{,}29)$	$[12{,}29; 13{,}91)$	$[13{,}91; 15{,}53)$
Частоти n_i	40	24	16	8	3	5	1	3
Частості ω_i	0,4	0,24	0,16	0,08	0,03	0,05	0,01	0,03
Накопичені частості $\omega_i^{нак}$	0,4	0,64	0,8	0,88	0,91	0,96	0,97	1

2. Зробити графічне зображення вибірки

Графічним зображення конкретної реалізації вибірки є гістограма, що складається з прямокутників, які побудовані на Δ_i та мають висоти $h_i = \frac{\omega_i}{d_i}$, де d_i – довжина Δ_i. У нашому випадку d_i однакове для всіх інтервалів: $d_i = d = 1{,}62$. Для зручності побудови гістограми додаємо до ІВР ще один рядок – висоти.

Інтервал Δ_i	$[2{,}57; 4{,}19)$	$[4{,}19; 5{,}81)$	$[5{,}81; 7{,}43)$	$[7{,}43; 9{,}05)$	$[9{,}05; 10{,}67)$	$[10{,}67; 12{,}29)$	$[12{,}29; 13{,}91)$	$[13{,}91; 15{,}53)$
Частоти n_i	40	24	16	8	3	5	1	3
Частості ω_i	0,4	0,24	0,16	0,08	0,03	0,05	0,01	0,03
Накопичені частості $\omega_i^{\text{нак}}$	0,4	0,64	0,8	0,88	0,91	0,96	0,97	1
Висоти h_i	0,247	0,148	0,099	0,049	0,019	0,031	0,006	0,019

Побудуємо тепер гістограму, яка і є геометричною інтерпретацією даної реалізації:

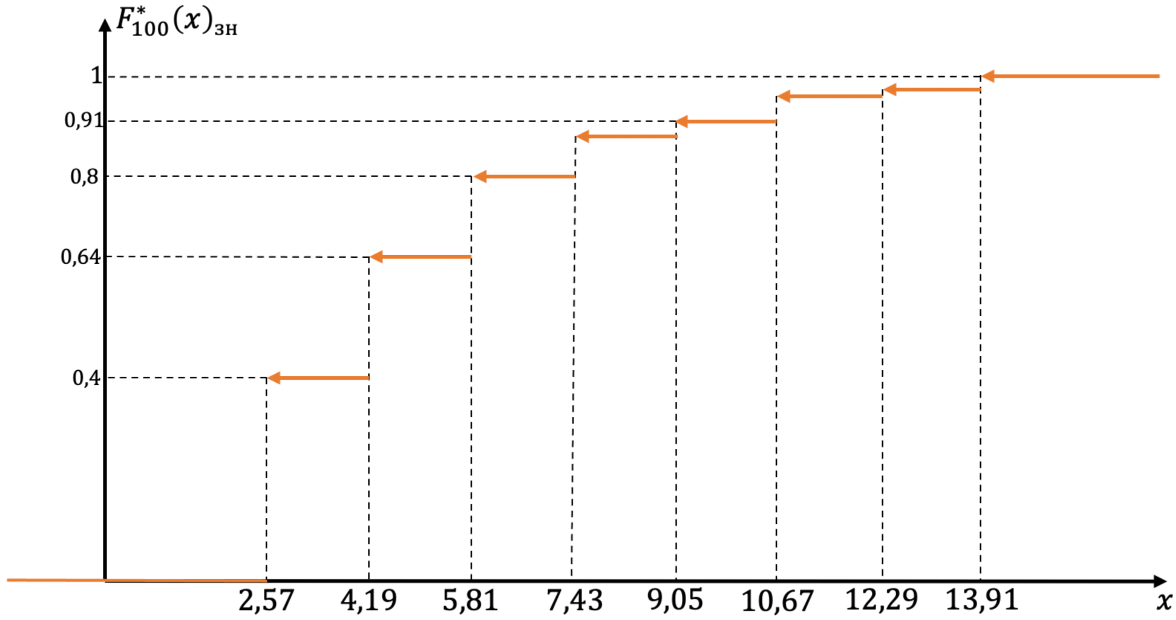


3. Побудувати емпіричну функцію розподілу

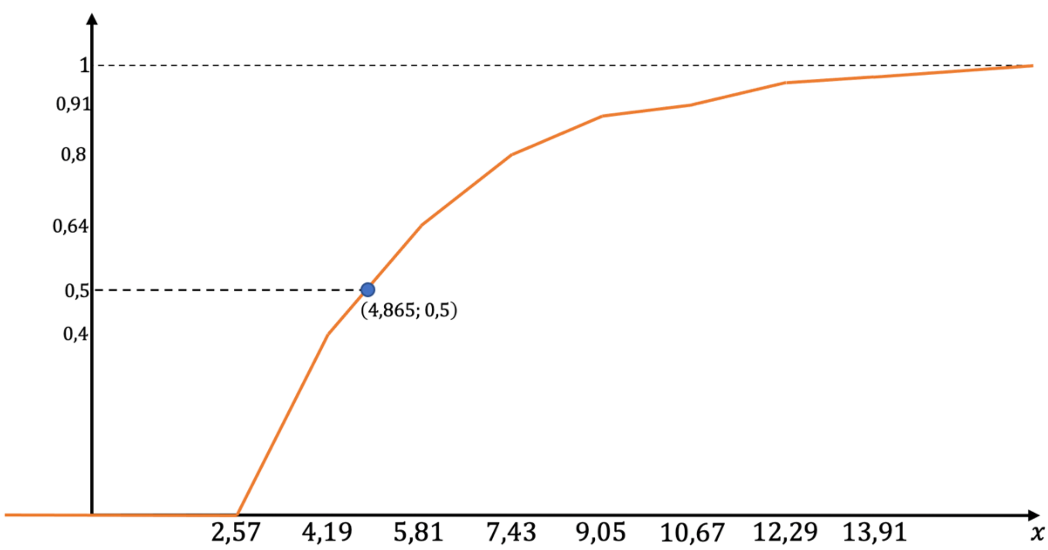
Побудуємо емпіричну функцію розподілу нашої конкретної реалізації вибірки:

$$\left(F_{100}^{*}(x)\right)_{зн}=\begin{cases}0, & x\le x_{min}\\ \omega_1^{нак}, & x_{min}<x\le t_1\\ \omega_2^{нак}, & t_1<x\le t_2\\ \dots & \\ 1, & x>t_r\end{cases}=\begin{cases}0, & x\le 2{,}57\\ 0{,}4, & 2{,}57<x\le 4{,}19\\ 0{,}64, & 4{,}19<x\le 5{,}81\\ 0{,}8, & 5{,}81<x\le 7{,}43\\ 0{,}88, & 7{,}43<x\le 9{,}05\\ 0{,}91, & 9{,}05<x\le 10{,}67\\ 0{,}96, & 10{,}67<x\le 12{,}29\\ 0{,}97, & 12{,}29<x\le 13{,}91\\ 1, & x>13{,}91\end{cases}$$

Побудуємо графік:



Оскільки майже всі значення нашої конкретної реалізації унікальні можемо зробити припущення, що ГС, що породила цю реалізацію є неперервною, тобто функція розподілу цієї ГС має бути неперервною. Тому доцільним буде ще і побудувати кумулятивну криву.



4. Знайти незміщену оцінку математичного сподівання та дисперсії

Точкова оцінка θ^* невідомого параметру θ називається незміщеною, якщо $\mathbb{E}\theta^* = \theta$

За оцінку математичного сподівання візьмемо вибіркове середнє:

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

Перевіримо незміщеність цієї оцінки:

$$\mathbb{E}\bar{\xi} = \mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k\right) = [\text{за властивістю мат. сподівання}] = \frac{1}{n}\mathbb{E}\sum_{k=1}^{n}\xi_k =$$

$$= [\text{за властивістю мат. сподівання}] = \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}\xi_k = [\text{Всі } \xi_k \text{ розподілені}$$

$$\text{однаково}] = \frac{1}{n}\cdot n\,\mathbb{E}\xi = \mathbb{E}\xi$$

Отже, за означення $\bar{\xi}$ - незміщена оцінка математичного сподівання

За оцінку дисперсії будемо брати виправлену вибіркову дисперсію, щоб мати незміщеність, а не асимптотичну незміщеність.

$$\mathbb{D}^{**}\xi = \frac{1}{n-1}\sum_{k=1}^{n}\left(\xi_k - \bar{\xi}\right)^2$$

Перевіримо, чи дійсно така точкова оцінка є незміщеною:

$$\mathbb{E}(\mathbb{D}^{**}\xi) = \mathbb{E}\left(\frac{1}{n-1}\sum_{k=1}^{n}\left(\xi_k - \bar{\xi}\right)^2\right) = [\text{за властивістю мат. сподівання}] =$$

$$= \frac{1}{n-1}\mathbb{E}\sum_{k=1}^{n}\left(\xi_k - \bar{\xi}\right)^2 = \frac{1}{n-1}\mathbb{E}\sum_{k=1}^{n}\left((\xi_k - \mathbb{E}\xi) - \left(\bar{\xi} - \mathbb{E}\xi\right)\right)^2 =$$

$$= \frac{1}{n-1}\mathbb{E}\sum_{k=1}^{n}\left((\xi_k - \mathbb{E}\xi)^2 - 2(\xi_k - \mathbb{E}\xi)\left(\bar{\xi} - \mathbb{E}\xi\right) + \left(\bar{\xi} - \mathbb{E}\xi\right)^2\right) =$$

$$= \frac{1}{n-1}\mathbb{E}\left(\sum_{k=1}^{n}(\xi_k - \mathbb{E}\xi)^2 - 2\left(\bar{\xi} - \mathbb{E}\xi\right)\sum_{k=1}^{n}(\xi_k - \mathbb{E}\xi) + n\left(\bar{\xi} - \mathbb{E}\xi\right)^2\right) =$$

$$= \left[\sum_{k=1}^{n}(\xi_k - \mathbb{E}\xi) = \sum_{k=1}^{n}\xi_k - \sum_{k=1}^{n}\mathbb{E}\xi = n\cdot\frac{1}{n}\sum_{k=1}^{n}\xi_k - n\mathbb{E}\xi = n\left(\bar{\xi} - \mathbb{E}\xi\right)\right] =$$

$$= \frac{1}{n-1}\mathbb{E}\left(\sum_{k=1}^{n}(\xi_k - \mathbb{E}\xi)^2 - 2n\left(\bar{\xi} - \mathbb{E}\xi\right)^2 + n\left(\bar{\xi} - \mathbb{E}\xi\right)^2\right) =$$

$$= \frac{1}{n-1}\mathbb{E}\left(\sum_{k=1}^{n}(\xi_k - \mathbb{E}\xi)^2 - n\left(\bar{\xi} - \mathbb{E}\xi\right)^2\right) = [\text{за незміщеністю } \bar{\xi}] =$$

$$= \frac{1}{n-1}\mathbb{E}\left(\sum_{k=1}^{n}(\xi_k - \mathbb{E}\xi)^2 - n\big(\bar{\xi} - \mathbb{E}\bar{\xi}\big)^2\right) = [\text{Всі } \xi_k \text{ розподілені однаково}]$$

$$= \frac{1}{n-1}\big(n\mathbb{D}\xi - n\mathbb{D}\bar{\xi}\big) = \frac{n}{n-1}\left(\mathbb{D}\xi - \mathbb{D}\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k\right)\right) = [\text{ за вл. дисперсії}]$$

$$= \frac{n}{n-1}\left(\mathbb{D}\xi - \frac{1}{n^2}\mathbb{D}\sum_{k=1}^{n}\xi_k\right) = [\xi_k - \text{незалежні у сукупності}] =$$

$$= \frac{n}{n-1}\left(\mathbb{D}\xi - \frac{1}{n^2}\sum_{k=1}^{n}\mathbb{D}\xi_k\right) = [\text{Всі } \xi_k \text{ розподілені однаково}] =$$

$$= \frac{n}{n-1}\left(\mathbb{D}\xi - \frac{1}{n}\mathbb{D}\xi\right) = \mathbb{D}\xi\frac{n}{n-1}\frac{n-1}{n} = \mathbb{D}\xi$$

Отже, за означенням $\mathbb{D}^{**}\xi$ – незміщена оцінка дисперсії

Обчислимо значення вибіркового середнього на нашій конкретній реалізації:

$$\big(\bar{\xi}\big)_{\text{зн}} = \bar{x} = \frac{1}{n}\sum_{k=1}^{n}x_k = \frac{1}{100}\sum_{k=1}^{100}x_k = 5{,}6278$$

Обчислимо значення виправленої вибіркової дисперсії:

$$(\mathbb{D}^{**}\xi)_{\text{зн}} = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})^2 = \frac{1}{99}\sum_{k=1}^{100}(x_k - 5{,}6278)^2 \approx 9{,}0019$$

5. Обчислити значення вибіркової медіани, моди, асиметрії

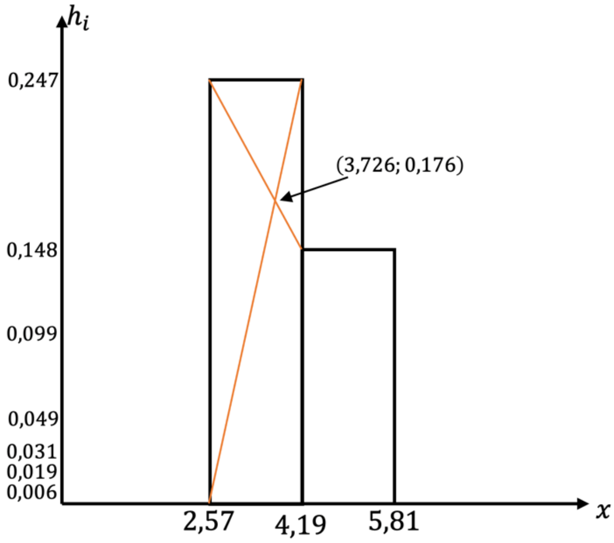
Вибіркова мода

Оскільки у нас ІВР, то:

$$(Mo^*\xi)_{зн} = y_1 + h_1 \frac{n_1}{(n_1) + (n_1 - n_2)} = 2{,}57 + 1{,}62 \frac{40}{40 + (40 - 24)}$$

$\approx 3{,}727$

Також обчислимо геометричним шляхом:



Таким шляхом отримали:

$(Mo^*\xi)_{зн} \approx 3{,}726$

Вибіркова медіана

Оскільки у нас ІВР, то:

$$(Me^*\xi)_{зн} = y_2 + \frac{h_2}{n_2}\left(\frac{n}{2} - \sum_{k=1}^{2} n_k\right) = 4{,}19 + \frac{1{,}62}{24}\left(\frac{100}{2} - 40\right) = 4{,}865$$

Також знайдемо значення вибіркової медіани на нашій реалізації за допомогою кумулятивної кривої. Це буде абсциса точки, в якій кумулята набуває значення 0,5. Оскільки кумулята є аналогом функції розподілу, а медіана це абсциса точки в якій функція розподілу набуває значення 0,5.

Таким чином маємо: $(Me^*\xi)_{зн} = 4{,}865$

Також можемо обрахувати значення медіани як серединне значення нашої відсортованої конкретної реалізації:

$(Me^*\xi)_{зн} = (4{,}67 + 4{,}7)/2 \ = 4{,}685$

Вибіркова асиметрія

$$As^*\xi = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(\xi_i - \bar{\xi}\right)^3}{(\mathbb{D}^{**}\xi)^{\frac{3}{2}}}$$

$$(As^*\xi)_{зн} = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(x_i^0 - \bar{x}\right)^3}{(\mathbb{D}^{**}\xi)_{зн}^{3/2}} = \frac{\frac{1}{100}\sum_{i=1}^{n}\left(x_i^0 - 5{,}6278\right)^3}{9{,}0019^{3/2}} \approx 1{,}4436$$

6. Висунити гіпотезу про розподіл, за яким отримано вибірку

Для висування гіпотези звернемо свою увагу на наступні речі:

Обвідна гістограми розподілу є своєрідним аналогом графіку розподілу щільності ГС. З гістограми можемо припустити, що щільність є спадною. А характер спаду нагадує графік щільності експоненційного закону. До того ж найменшим значенням у реалізації є 2,57, то це більш схоже на експоненційний закон із зсувом.

Графік кумуляти на нашій реалізації нагадує графік функції розподілу зсунутого експоненційного розподілу.

Оскільки на нашій конкретній реалізації майже всі значення унікальні, то припустимо, що ГС – неперервна.

Асиметрія не близька до нуля, то це скоріш за все не рівномірний і не нормальний закони

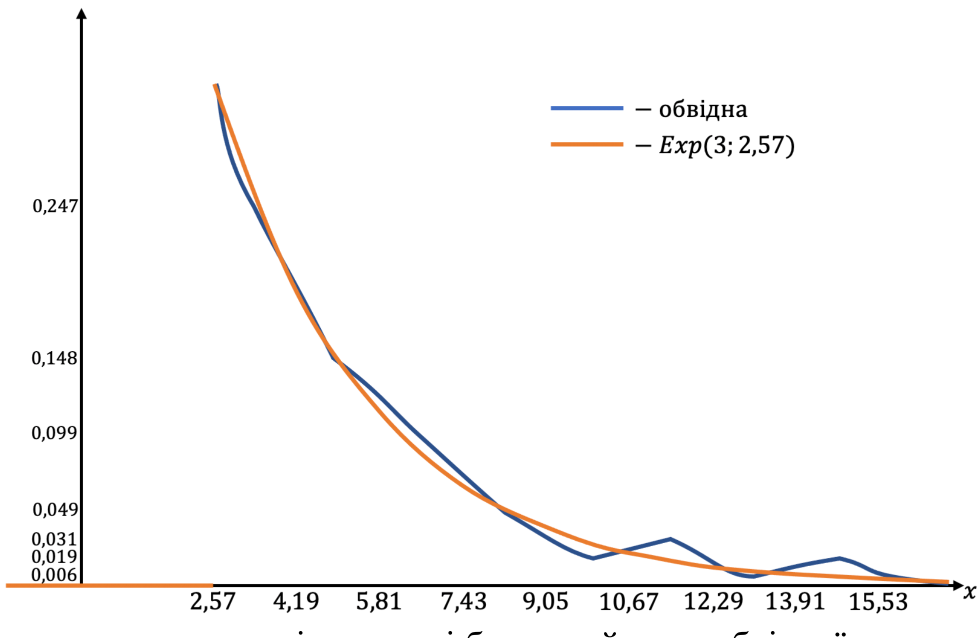
Щоб підкріпити думку про зсунутий експоненційний розподіл спробуємо розглянути $\eta \sim Exp(\lambda, b)$. $\mathbb{D}\eta = \lambda^2$, а ми маємо $(\mathbb{D}^{**}\xi)_{\text{зн}} = 9{,}0019$.

Припустимо, що $\lambda \approx \sqrt{(\mathbb{D}^{**}\xi)_{\text{зн}}} \approx 3$

Зсув у експоненційному законі має значення найменшого значення, яке може приймати ВВ. Тож припустимо, що $b \approx \min\limits_{1 \le k \le n} x_k = 2{,}57$

$\mathbb{E}\eta = \lambda + b \approx 5{,}57$, що видодить достатньо близько до $\bar{x} = 5{,}6278$

Тепер накладемо графік щільності η на обвідну гістограму частот:



Бачимо, що крива розподілу доволі близько йде до обвідної.

Тож, виходячи з вищезазначеного, вважаю доцільним висунути гіпотезу: $H_0: \xi \sim Exp(\lambda, b)$, тобто випадкова вибірка, з якої узята наша конкретна реалізація, породжена ГС розподіленою за зсунутим експоненційним законом.

7. Знайти точкові оцінки параметрів гіпотетичного закону розподілу та перевірити їх властивості

Для знаходження точкових оцінок невідомих параметрів закону розподілу генеральної сукупності випадкової вибірки будемо використовувати метод моментів та максимальної правдоподібності.

Метод моментів

$$\begin{cases} \mathbb{E}\xi = \lambda + b \\ \mathbb{D}\xi = \lambda^2 \end{cases} \Rightarrow \begin{cases} b = -\lambda + \mathbb{E}\xi \\ \lambda = \sqrt{\mathbb{D}\xi} \end{cases} \Rightarrow \begin{cases} b = -\sqrt{\mathbb{D}\xi} + \mathbb{E}\xi \\ \lambda = \sqrt{\mathbb{D}\xi} \end{cases} \Rightarrow \begin{cases} b^*_{\text{ММ}} = \bar{\xi} - \sqrt{\mathbb{D}^{**}\xi} \\ \lambda^*_{\text{ММ}} = \sqrt{\mathbb{D}^{**}\xi} \end{cases}$$

Метод максимальної правдоподібності

Нам необхідно максимізувати значення функції правдоподібності. Беручи до уваги, що натуральний логарифм монотонна зростаюча функція, будемо досліджувати логарифм функції правдоподібності.

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

Для $x_k \geq b$:

$$\mathcal{L}_{Exp(\lambda,b)}(\vec{x}, \lambda, b) = \prod_{k=1}^{n} f_\xi(x_k) = \prod_{k=1}^{n} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(x_k-b)} = \frac{1}{\lambda^n} \exp\left\{-\frac{1}{\lambda}\sum_{k=1}^{n}(x_k - b)\right\}$$

$$\ln \mathcal{L}_{Exp(\lambda,b)}(\vec{x}, \lambda, b) = -n \ln \lambda - \frac{1}{\lambda}\sum_{k=1}^{n}(x_k - b) = -n \ln \lambda - \frac{1}{\lambda}\sum_{k=1}^{n} x_k + \frac{nb}{\lambda}$$

Застосуємо необхідну умову екстремуму:

$$\begin{cases} \frac{\partial \ln \mathcal{L}_{Exp(\lambda,b)}(\vec{x}, \lambda, b)}{\partial \lambda} = 0 \\ \frac{\partial \ln \mathcal{L}_{Exp(\lambda,b)}(\vec{x}, \lambda, b)}{\partial b} = 0 \end{cases} \Rightarrow \begin{cases} -\frac{n}{\lambda} + \frac{1}{\lambda^2}\sum_{k=1}^{n} x_k - \frac{nb}{\lambda^2} = 0 \\ \frac{n}{\lambda} \neq 0 \end{cases}$$

Як бачимо екстремуму немає, але водночас ми бачимо, що частинна похідна по b більше за нуль, а отже зі зростанням b і логарифм функції правдоподібності буде зростати, але одночасно $b \leq \min\limits_{1 \leq k \leq n} x_k$, оскільки зсув не може бути більше за найменше значення реалізації. А отже $\ln \mathcal{L}_{Exp(\lambda,b)}(\vec{x}, \lambda, b)$ буде приймати найбільшого значення, коли $b = \min\limits_{1 \leq k \leq n} x_k$

Тепер знайдемо при якому λ, має найбільше значення:

$$\frac{\partial \ln \mathcal{L}_{Exp(\lambda,b)}(\vec{x}, \lambda, b)}{\partial \lambda} = -\frac{n}{\lambda} + \frac{1}{\lambda^2}\sum_{k=1}^{n} x_k - \frac{nb}{\lambda^2} = \frac{n}{\lambda^2}(-\lambda + \bar{x} - b) = 0$$

Оскільки $\frac{n}{\lambda^2} > 0$, то $\lambda_{\text{кр}} = \bar{x} - b$ – точка в якій може досягатися максимум.
Тепер перевіримо знак другої похідної у цій точці:
$$\frac{\partial^2 \ln \mathcal{L}_{Exp(\lambda,b)}(\vec{x}, \lambda, b)}{\partial \lambda^2}|_{\lambda=\bar{x}-b} = \left(\frac{n}{\lambda^2} - \frac{2n(\bar{x}-b)}{\lambda^3}\right)|_{\lambda=\bar{x}-b} = \frac{n-2n}{(\bar{x}-b)^2} < 0$$
А отже $argmax_{\lambda,b} \ln \mathcal{L}_{Exp(\lambda,b)}(\vec{x}, \lambda, b) = \ln \mathcal{L}_{Exp(\lambda,b)}\left(\vec{x}, \bar{x} - b\,, \min\limits_{1\le k\le n} x_k\right)$
Отже методом максимальної правдоподібності отримали:
$$\begin{cases} \lambda^*_{\text{ММП}} = \bar{\xi} - b \\ b^*_{\text{ММП}} = \min\limits_{1\le k\le n} \xi_k \end{cases}$$

Будемо перевіряти властивості оцінок отриманих методом максимальної правдоподібності:

1) $\lambda^*_{\text{ММП}} = \bar{\xi} - b$

Незміщеність:
$$\mathbb{E}\lambda^*_{\text{ММП}} = \mathbb{E}(\bar{\xi} - b) = \mathbb{E}\bar{\xi} - \mathbb{E}b = \mathbb{E}\xi - b = \lambda + b - b = \lambda$$
Отже оцінка незміщена

Консистентність:
$$\mathbb{D}\lambda^*_{\text{ММП}} = \mathbb{D}(\bar{\xi} - b) = \mathbb{D}\bar{\xi} = \mathbb{D}\left(\frac{1}{n}\sum_{k=1}^{n} \xi_k\right) = \frac{1}{n^2}\sum_{k=1}^{n} \mathbb{D}\xi_k = \frac{1}{n}\mathbb{D}\xi = \frac{\lambda^2}{n} \to 0, n \to \infty$$
Отже за достатньою умовою консистентності незміщеною оцінки, маємо що $\lambda^*_{\text{ММП}}$ - консистентна

Ефективність:
$$\frac{\partial \ln \mathcal{L}_{Exp(\lambda,b)}(\vec{\xi}, \lambda, b)}{\partial \lambda} = -\frac{n}{\lambda} + \frac{1}{\lambda^2}\sum_{k=1}^{n} \xi_k - \frac{nb}{\lambda^2} = \frac{n}{\lambda^2}(-\lambda + \bar{\xi} - b) =$$
$$= \frac{n}{\lambda^2}(\lambda^*_{\text{ММП}} - \lambda)$$
За наслідком з нерівності Рао-Крамера, бачимо що $\lambda^*_{\text{ММП}}$ - ефективна

Асимптотична нормальність:
$$\frac{\lambda^*_{\text{ММП}} - \mathbb{E}\lambda^*_{\text{ММП}}}{\sqrt{\mathbb{D}\lambda^*_{\text{ММП}}}} = \frac{\lambda^*_{\text{ММП}} - \lambda}{\sqrt{\frac{\lambda^2}{n}}} = \frac{\sqrt{n}(\bar{\xi} - b - \lambda)}{\lambda} = \frac{\sqrt{n}(\bar{\xi} - \mathbb{E}\xi)}{\sqrt{\mathbb{D}\xi}} =$$
$$= \frac{\sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n} \xi_k - \mathbb{E}\xi\right)}{\sqrt{\mathbb{D}\xi}} = \frac{1}{\sqrt{n}}\frac{\sum_{k=1}^{n} \xi_k - \mathbb{E}\xi}{\sqrt{\mathbb{D}\xi}} = \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\frac{(\xi_k - \mathbb{E}\xi)}{\sqrt{\mathbb{D}\xi}} \xrightarrow{F}$$

$\xrightarrow{F} N(0,1), n \to \infty$, бо оскільки всі ξ_k – однаково розподілені, тому виконується умова Ляпунова і за ЦГТ маємо збіжність за розподілом. Отже отримали, що $\lambda^*_{\text{ММП}}$ - асимптотична нормальна

2) $b^*_{\text{ММП}} = \min\limits_{1 \le k \le n} \xi_k$

Знайдемо закон розподілу мінімуму ξ_k:

$$f_{b^*_{\text{ММП}}} = n\left(1 - F_\xi(x)\right)^{n-1} f_\xi(x) = n\left(1 - 1 + e^{-\frac{1}{\lambda}(x-b)}\right)^{n-1} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(x-b)} =$$
$$= \frac{n}{\lambda} e^{-\frac{n}{\lambda}(x-b)}, x \ge b$$

Маємо, що $\min\limits_{1 \le k \le n} \xi_k \sim Exp\left(\frac{\lambda}{n}, b\right)$

Незміщеність

$$\mathbb{E}b^*_{\text{ММП}} = \mathbb{E} \min_{1 \le k \le n} \xi_k = \frac{\lambda}{n} + b$$

Така оцінка є асимптотична незміщеною. Розглянемо $b^{**}_{\text{ММП}} = \min\limits_{1 \le k \le n} \xi_k - \frac{\lambda}{n}$

$$\mathbb{E}b^{**}_{\text{ММП}} = \mathbb{E}\left(\min_{1 \le k \le n} \xi_k - \frac{\lambda}{n}\right) = \mathbb{E} \min_{1 \le k \le n} \xi_k - \mathbb{E}\frac{\lambda}{n} = \frac{\lambda}{n} + b - \frac{\lambda}{n} = b$$

Отже оцінка $b^{**}_{\text{ММП}} = \min\limits_{1 \le k \le n} \xi_k - \frac{\lambda}{n}$ – є незміщеною, тож будемо надалі використовувати її.

Консистентність

$$\mathbb{D}b^{**}_{\text{ММП}} = \mathbb{D}\left(\min_{1 \le k \le n} \xi_k - \frac{\lambda}{n}\right) = \mathbb{D} \min_{1 \le k \le n} \xi_k = \left(\frac{\lambda}{n}\right)^2 \to 0, n \to \infty$$

Отже за достатньою умовою консистентності незміщеною оцінки, маємо що $b^{**}_{\text{ММП}}$ - консистентна

Ефективність

$$\mathcal{I}(b) = \mathbb{E}\left(\frac{\partial \ln \mathcal{L}_{Exp(\lambda,b)}(\vec{x}, \lambda, b)}{\partial b}\right)^2 = \mathbb{E}\left(\frac{n}{\lambda}\right)^2 = \left(\frac{n}{\lambda}\right)^2 = \frac{1}{\mathbb{D}b^{**}_{\text{ММП}}}$$

За нерівністю Рао-Крамера маємо, що $b^{**}_{\text{ММП}}$ - ефективна

Асимптотична нормальність

$$\frac{b^{**}_{\text{ММП}} - \mathbb{E}b^{**}_{\text{ММП}}}{\sqrt{\mathbb{D}b^{**}_{\text{ММП}}}} = \frac{\min\limits_{1 \le k \le n} \xi_k - \frac{\lambda}{n} - b}{\sqrt{\left(\frac{\lambda}{n}\right)^2}} = \frac{n}{\lambda}\left(\min_{1 \le k \le n} \xi_k - \frac{\lambda}{n} - b\right)$$

Для подальшого просування треба розглянути декілька фактів:
$\xi \sim Exp(\lambda)$. Розглянемо $\eta = \xi + b$. Знайдемо розподіл η.
$\varphi(x) = x + b$ – монотонна неперервна функція
$\varphi^{-1}(y) = y - b$ – диференційована
Тоді маємо:

$$f_\eta(y) = f_\xi(\varphi^{-1}(y)) \cdot \left|(\varphi^{-1}(y))'\right| = \begin{cases} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(y-b)}, & y - b \geq 0 \\ 0, & y - b < 0 \end{cases}$$

Отже, маємо що $\eta \sim Exp(\lambda, b)$
Тоді повертаючись до нашої рівності, оскільки $\min\limits_{1\leq k\leq n} \xi_k \sim Exp\left(\frac{\lambda}{n}, b\right)$ маємо:

$$\frac{n}{\lambda}\left(\min_{1\leq k\leq n} \xi_k - \frac{\lambda}{n} - b\right) = \left[\eta \sim Exp\left(\frac{\lambda}{n}\right)\right] = \frac{n}{\lambda}\left(\eta + b - \frac{\lambda}{n} - b\right) = \frac{n}{\lambda}\eta - 1$$

Тепер розглянемо розподіл $\frac{n}{\lambda}\eta$. Для цього скористаємося характеристичними функціями:

$$\chi_{\frac{n}{\lambda}\eta}(t) = \chi_\eta\left(\frac{n}{\lambda}t\right) = \frac{\frac{n}{\lambda}}{\frac{n}{\lambda} - i\frac{n}{\lambda}t} = \frac{1}{1 - it} = \chi_\zeta(t), \zeta \sim Exp(1)$$

А отже знову повертаючись до рівності:

$$\frac{n}{\lambda}\eta - 1 = \zeta - 1$$

Тепер знайдемо характеристичну функцію $\zeta - 1$:

$$\chi_{\zeta-1}(t) = e^{it}\chi_\zeta(t) = e^{it}\frac{1}{1-it} \neq e^{-\frac{t^2}{2}}$$

А тепер за теоремою Леві маємо, що якщо була б збіжність за розподілом, то мала б бути і по точкова збіжність характеристичних функцій. А це не так, отже $b^{**}_{\text{ММП}}$ - не є асимптотично нормальною.

Отже остаточно найкращими оцінками виявилися:

$$\begin{cases} \lambda^*_{\text{ММП}} = \bar{\xi} - b \\ b^{**}_{\text{ММП}} = \min\limits_{1\leq k\leq n} \xi_k - \frac{\lambda}{n} \end{cases}$$

Оскільки $\lambda^*_{\text{ММП}}, b^{**}_{\text{ММП}}$ - це статистики, а точні значення b та λ нам невідомі, ми не зможемо знайти точні значення оцінок. Але ми можемо знайти наближені значення оцінок заміняючи b та λ на значення їх найкращих оцінок, що і зробимо:

$$(\lambda^*_{\text{ММП}})_{\text{зн}} = \bar{x} - b \approx \bar{x} - \min_{1\leq k\leq n} x_k + \frac{\lambda}{n} \approx \bar{x} - \min_{1\leq k\leq n} x_k + \frac{(\lambda^*_{\text{ММП}})_{\text{зн}}}{n}$$

$$(\lambda^*_{\text{ММП}})_{\text{зн}} \approx \frac{n}{n-1}\left(\bar{x} - \min_{1\leq k\leq n} x_k\right)$$

$$(b^{**}_{\text{ММП}})_{\text{зн}} = \min_{1\le k\le n} x_k - \frac{\lambda}{n} \approx \min_{1\le k\le n} x_k - \frac{(\lambda^{*}_{\text{ММП}})_{\text{зн}}}{n} \approx$$
$$\approx \min_{1\le k\le n} x_k - \frac{1}{n-1}\left(\bar{x} - \min_{1\le k\le n} x_k\right) \approx \frac{n}{n-1}\min_{1\le k\le n} x_k - \frac{\bar{x}}{n-1} \approx$$
$$\approx \frac{n}{n-1}\left(\min_{1\le k\le n} x_k - \frac{1}{n}\bar{x}\right)$$

Підрахуємо ці значення:

$$(\lambda^{*}_{\text{ММП}})_{\text{зн}} \approx \frac{100}{99}(5{,}6278 - 2{,}57) \approx 3{,}089$$
$$(b^{**}_{\text{ММП}})_{\text{зн}} \approx \frac{100}{99}\left(2{,}57 - \frac{5{,}6278}{100}\right) \approx 2{,}539$$

Можемо висунити гіпотезу: $H_0 \colon \xi \sim Exp(3{,}089; 2{,}539)$

8. Перевірити за допомогою критерія Пірсона гіпотезу про розподіл з рівнем значущості $\boldsymbol{\alpha = 0,05}$

Будемо перевіряти гіпотезу $H_0: \xi \sim Exp(3{,}089; 2{,}539)$. За гіпотезою множина можливих значень ξ $X = [2{,}539; +\infty)$. Розіб'ємо множину на такі попарно не перетинаючі підмножини: $[2{,}539; 2{,}57)$, $[2{,}57; 4{,}19)$, $[4{,}19; 5{,}81)$, $[5{,}81; 7{,}43)$, $[7{,}43; 9{,}05)$, $[9{,}05; 10{,}67)$, $[10{,}67; 12{,}29)$, $[12{,}29; 13{,}91)$, $[13{,}91; 15{,}53)$, $[15{,}53; +\infty)$

X_i	p_i	np_i
$[2{,}539; 2{,}57)$	0,01	1
$[2{,}57; 4{,}19)$	0,404	40,4
$[4{,}19; 5{,}81)$	0,239	23,9
$[5{,}81; 7{,}43)$	0,142	14,2
$[7{,}43; 9{,}05)$	0,084	8,4
$[9{,}05; 10{,}67)$	0,05	5
$[10{,}67; 12{,}29)$	0,029	2,9
$[12{,}29; 13{,}91)$	0,017	1,7
$[13{,}91; 15{,}53)$	0,01	1
$[15{,}53; +\infty)$	0,015	1,5

Для виконання умови $np_i > 10$ об'єднаємо такі інтервали:
X_1 з X_2 та X_5 з $X_6, X_7, X_8, X_9, X_{10}$

X_i	p_i	np_i	n_i
$[2{,}539; 4{,}19)$	0,414	41,4	40
$[4{,}19; 5{,}81)$	0,239	23,9	24
$[5{,}81; 7{,}43)$	0,142	14,2	16
$[7{,}43; +\infty)$	0,205	20,5	20

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

$$(\eta)_{\text{зн}} = \frac{(40 - 41{,}4)^2}{41{,}4} + \frac{(24 - 23{,}9)^2}{23{,}9} + \frac{(16 - 14{,}2)^2}{14{,}2} + \frac{(20 - 20{,}5)^2}{20{,}5} = $$
$$= 0{,}288$$

Оскільки за теоремою Пірсона якщо справджується H_0, то

$$\eta \xrightarrow{F} \chi^2_{4-2-1}, n \to \infty$$

На рівні значущості $\alpha = 0{,}05$ з таблиці хі-квадрат критичне значення $t_{\text{кр}} = 3{,}84$. А ми отримали, що $(\eta)_{\text{зн}} < t_{\text{кр}}$, тому дослідні дані на заданому рівні значущості не суперечать висунутій гіпотезі

9. Знайти довірчий інтервал для параметрів гіпотетичного закону розподілу з рівнем надійності $\boldsymbol{\gamma = 0,95}$

Довірчий інтервал для параметру λ:

Як було доведено $\lambda^*_{\text{ММП}} = \bar{\xi} - b$ - асимптотична нормальна, а отже:

$$\frac{\lambda^*_{\text{ММП}} - \mathbb{E}\lambda^*_{\text{ММП}}}{\sqrt{\mathbb{D}\lambda^*_{\text{ММП}}}} = \frac{\sqrt{n}(\lambda^*_{\text{ММП}} - \lambda)}{\lambda} \approx N(0,1)$$

Тому:

$$0{,}95 = \mathbb{P}\{|\lambda^*_{\text{ММП}} - \lambda| < \varepsilon\} = \mathbb{P}\left\{\frac{\sqrt{n}|\lambda^*_{\text{ММП}} - \lambda|}{\lambda} < \frac{\sqrt{n}\varepsilon}{\lambda}\right\} =$$

$$= \mathbb{P}\left\{\frac{\sqrt{n}|\lambda^*_{\text{ММП}} - \lambda|}{\lambda} < t\right\} \approx 2\Phi(t)$$

З таблиці Лапласа маємо, що $t = 1{,}96$. Тоді:

$$\frac{\sqrt{n}|\lambda^*_{\text{ММП}} - \lambda|}{\lambda} < 1{,}96 \Rightarrow |\lambda^*_{\text{ММП}} - \lambda| < \frac{1{,}96\lambda}{\sqrt{n}} \Rightarrow -\frac{1{,}96\lambda}{\sqrt{n}} < \lambda^*_{\text{ММП}} - \lambda < \frac{1{,}96\lambda}{\sqrt{n}} \Rightarrow$$

$$\Rightarrow \begin{cases} \lambda^*_{\text{ММП}} - \lambda > -\dfrac{1{,}96\lambda}{\sqrt{n}} \\ \lambda^*_{\text{ММП}} - \lambda < \dfrac{1{,}96\lambda}{\sqrt{n}} \end{cases} \Rightarrow \begin{cases} \dfrac{\lambda^*_{\text{ММП}}}{1 - \dfrac{1{,}96}{\sqrt{n}}} > \lambda \\ \dfrac{\lambda^*_{\text{ММП}}}{1 + \dfrac{1{,}96}{\sqrt{n}}} < \lambda \end{cases}$$

Отримали довірчий інтервал для $\gamma = 0{,}95$: $\left(\frac{\lambda^*_{\text{ММП}}}{1+\frac{1{,}96}{\sqrt{n}}}; \frac{\lambda^*_{\text{ММП}}}{1-\frac{1{,}96}{\sqrt{n}}}\right)$

Отже маємо, що з імовірністю $0{,}95$ $\lambda \in \left(\frac{(\lambda^*_{\text{ММП}})_{\text{зн}}}{1+\frac{1{,}96}{\sqrt{100}}}; \frac{(\lambda^*_{\text{ММП}})_{\text{зн}}}{1-\frac{1{,}96}{\sqrt{100}}}\right) = (2{,}583; 3{,}842)$

Довірчий інтервал для параметру b:

Оскільки $f_\xi(x) = 0$, при $x < b$, то $\mathbb{P}\left\{b \le \min\limits_{1\le k\le n} \xi_k\right\} = 1$

Отже будемо шукати довірчий інтервал такого виду:

$$0{,}95 = \mathbb{P}\left\{b^{**}_{\text{ММП}} - \varepsilon < b \le \min_{1\le k\le n} \xi_k\right\} = \mathbb{P}\{b^{**}_{\text{ММП}} - b < \varepsilon\}$$

Коли ми досліджували оцінку $b^{**}_{\text{ММП}} = \min\limits_{1\le k\le n} \xi_k - \frac{\lambda}{n}$ отримали що:

$$\frac{b^{**}_{\text{ММП}} - \mathbb{E}b^{**}_{\text{ММП}}}{\sqrt{\mathbb{D}b^{**}_{\text{ММП}}}} = \zeta - 1, \zeta \sim Exp(1)$$

Як було вище отримано $\zeta - 1 \sim Exp(1, -1)$

Тоді маємо:

$$\mathbb{P}\left\{b^{**}_{\text{ММП}} - \varepsilon < b \le \min_{1\le k\le n} \xi_k\right\} = \mathbb{P}\{b^{**}_{\text{ММП}} - b < \varepsilon\} = \mathbb{P}\left\{n\frac{b^{**}_{\text{ММП}} - b}{\lambda} < \frac{n\varepsilon}{\lambda}\right\}$$

$$0{,}95 = F_{Exp(1,-1)}\left(\frac{n\varepsilon}{\lambda}\right) \approx F_{Exp(1,-1)}\left(\frac{n\varepsilon}{(\lambda^*_{\text{ММП}})_{\text{зн}}}\right) =$$

$$= 1 - \exp\left\{-\left(\frac{n\varepsilon}{(\lambda^*_{\text{ММП}})_{\text{зн}}} + 1\right)\right\}$$

$$0{,}05 \approx \exp\left\{-\left(\frac{n\varepsilon}{(\lambda^*_{\text{ММП}})_{\text{зн}}} + 1\right)\right\} \Longrightarrow -\left(\frac{n\varepsilon}{(\lambda^*_{\text{ММП}})_{\text{зн}}} + 1\right) \approx -\ln 20$$

$$\varepsilon \approx \frac{(\lambda^*_{\text{ММП}})_{\text{зн}}}{n}(\ln 20 - 1)$$

Отримали довірчий інтервал для $\gamma = 0{,}95$:

$$\left(b^{**}_{\text{ММП}} - \frac{(\lambda^*_{\text{ММП}})_{\text{зн}}}{n}(\ln 20 - 1);\ \min_{1\le k\le n} \xi_k\right]$$

Отже маємо, що з імовірністю 0,95

$$b \in \left((b^{**}_{\text{ММП}})_{\text{зн}} - \frac{(\lambda^*_{\text{ММП}})_{\text{зн}}}{100}(\ln 20 - 1);\ \min_{1\le k\le n} x_k\right] = (2{,}477; 2{,}57]$$

10. Висновки

Проаналізувавши вхідні дані більшою мірою за допомогою графічних інтерпретацій та особливостей деяких розподілів було зроблено припущення щодо закону розподілу ГС сукупності випадкової вибірки, що породила задану реалізацію.
На основі припущень були знайдені найкращі очкові оцінки невідомих параметрів генеральної сукупності. Під час знаходження були доведені їх властивості. Підрахувавши значення цих статистик на нашій реалізації була висунута гіпотеза, що ГС сукупність розподілена за експоненційним законом зі зсувом з параметрами $\lambda = 3{,}089$ та $b = 2{,}539$. За критеріям Пірсона виявилося що дослідні дані не суперечать цій гіпотезі на рівні значущості $\alpha = 0{,}05$. Також були знайдені 95% довірчі інтервали для обох параметрів: $\lambda \in (2{,}583; 3{,}842)$ та $b \in (2{,}477; 2{,}57]$