

**Министерство науки и высшего образования Российской
Федерации Калужский филиал федерального государственного
бюджетного образовательного учреждения высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»**

Факультет «Машиностроительный»

Кафедра МК10 «Высшая математика и физика»

**Домашнее задание № 2
по курсу «Теория вероятностей и математическая статистика»
на тему
«Элементы математической статистики»**

Вариант 17

Выполнил студент гр. ИУК5-42Б Ли Р. В.

Проверил Супельняк М.И.

Домашнее задание сдано

дата

подпись студента

Домашнее задание проверено

дата

подпись студента

Домашнее задание принято

дата

подпись студента

Домашнее задание защищено

дата

подпись студента

Количество рейтинговых баллов _____

Калуга, 2023 г.

Задача 1

На промышленном предприятии установлен газоанализатор, который фиксирует превышение предельно допустимой концентрации вредных веществ в окружающем воздухе. В конце каждого месяца подсчитывается количество дней, в которые наблюдалось превышение предельно допустимой концентрации вредных веществ. В результате годичных наблюдений получена выборка объёма $n = 12$ из генеральной совокупности X , где случайная величина X – количество дней месяца, в которые наблюдалось превышение предельно допустимой концентрации вредных веществ в воздухе. Для полученной выборки необходимо:

- ~ построить статистический ряд;
- ~ построить полигон относительных частот;
- ~ определить эмпирическую функцию распределения $F(x)$ и построить её график;
- ~ вычислить числовые характеристики выборки: среднее \bar{x} , дисперсию $\hat{\sigma}^2$, стандартное отклонение $\hat{\sigma}$, исправленную дисперсию s^2 , исправленное стандартное отклонение s .

Задача 2

В результате сбоя на сортировочной линии цеха в большую партию болтов попало небольшое количество гаек, после чего они были расфасованы по упаковкам. В ходе исследования содержимого $n = 300$ упаковок был построен статистический ряд для сделанной выборки из генеральной совокупности X , где случайная величина X – количество гаек в упаковке. Полагая, что генеральная совокупность имеет распределение Пуассона с вероятностями

$$p(l; \lambda) = P(X = l) = \frac{\lambda^l e^{-\lambda}}{l!}, l = 0, 1, 2, \dots,$$

необходимо с помощью полученного статистического ряда выборки:

- ~ получить значения точечных оценок $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4$ частоты появления гаек в упаковке λ методом моментов, используя начальные моменты выборки первого и второго порядка и центральный момент выборки второго порядка, и методом наибольшего правдоподобия соответственно;
- ~ построить полигоны вероятностей для найденных значений точечных оценок и сравнить их с полигоном относительных частот.

Задача 4

В результате испытаний n двигателей построен интервальный статистический ряд для выборки объёма $n = 50$ из генеральной совокупности X , где случайная величина X – длительность безотказной работы двигателя, ч. Полагая, что длительность безотказной работы имеет показательное распределение с плотностью

$$0, x < 0;$$

$$f(x; \lambda) = \{\lambda e^{-\lambda x}, x \geq 0,$$

необходимо:

- ~ получить значения точечных оценок $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4$ частоты отказов λ методом моментов, используя начальные моменты выборки первого и второго порядка и центральный момент выборки второго порядка, и методом наибольшего правдоподобия соответственно;
- ~ для найденных значений точечных оценок построить графики плотности распределения вероятностей $f(x; \hat{\lambda})$ и сравнить их с графиком эмпирической плотности распределения $\hat{f}(x)$;
- ~ построить доверительный интервал для λ с коэффициентом доверия γ .

$n = 12$;

| | | | | | | | | | | | | |
|-------|---|---|---|---|---|---|---|----|---|----|----|----|
| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| x_i | 3 | 7 | 5 | 9 | 3 | 9 | 7 | 10 | 8 | 1 | 6 | 4 |

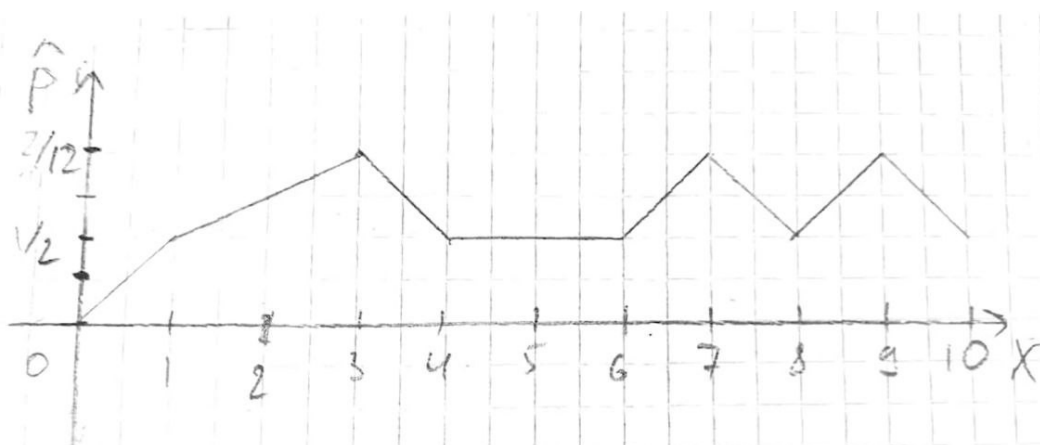
Найти: \hat{p}_i , $F(x)$, \bar{x} , $\hat{\sigma}^2$, $\hat{\sigma}$, s^2 , s .

Решение.

Выборка содержит $k = 6$ различных элементов, из которых можно составить возрастающую последовательность чисел \hat{x}_i , $i = 1, k$. Составляем статистический ряд выборки, в котором указываем частоту m_i каждого элемента \hat{x}_i и относительную частоту $\hat{p}_i = m_i/n$:

| | | | | | | | | | |
|-------------|------|------|------|------|------|------|------|------|------|
| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| \hat{x}_i | 1 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| m_i | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 2 | 1 |
| \hat{p}_i | 1/12 | 2/12 | 1/12 | 1/12 | 1/12 | 2/12 | 1/12 | 2/12 | 1/12 |

С помощью статистического ряда строим полигон относительных частот.

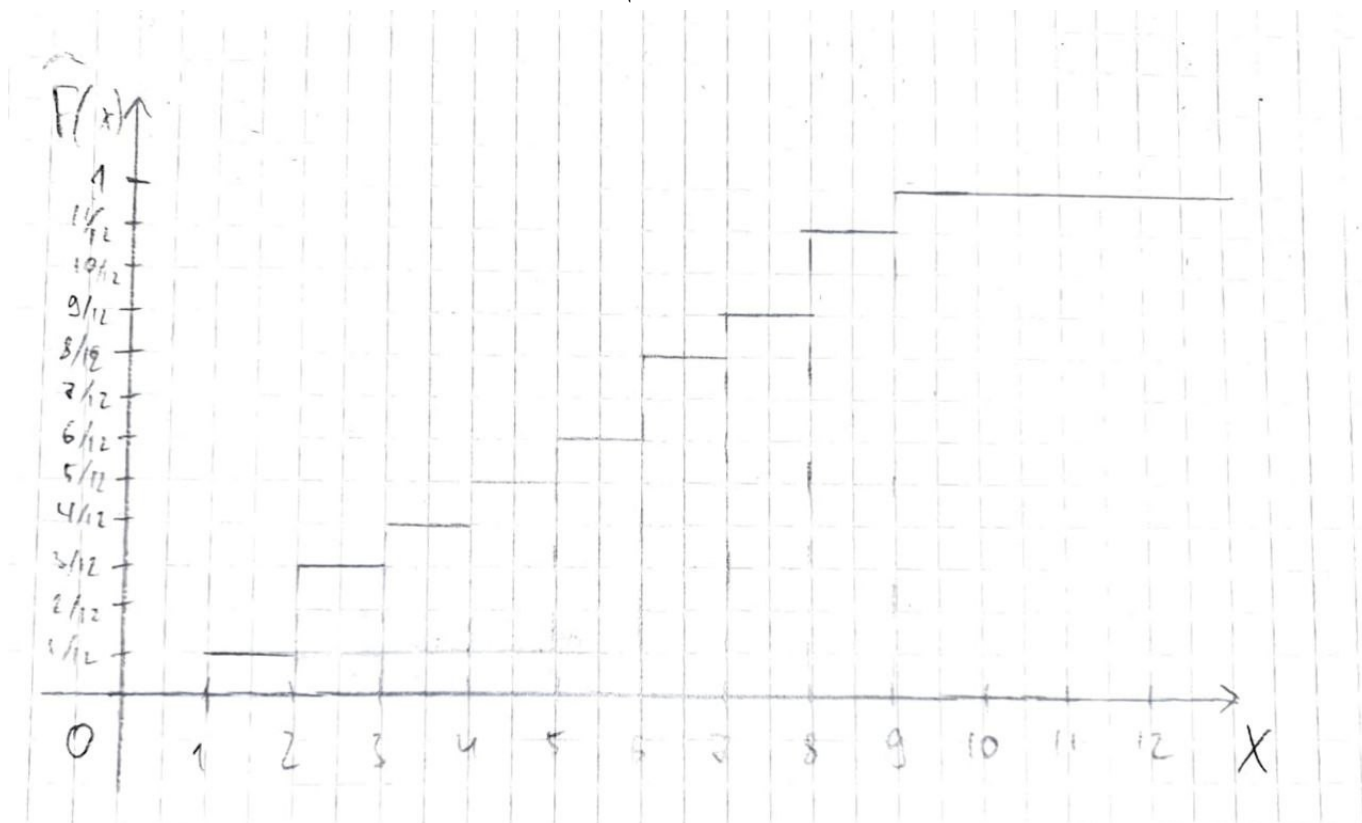


Определяем эмпирическую функцию распределения

$$\hat{F}(x) = \begin{cases} 0, & x \leq \hat{x}_1 \\ \sum_{j=1}^{i-1} \hat{p}_j, & \hat{x}_{i-1} < x \leq \hat{x}_i, i = \overline{2, k} \\ 1, & x > \hat{x}_k \end{cases}$$

для статистического ряда выборки:

$$\hat{F}(x) = \begin{cases} 0, & x \leq 0; \\ 1/12, & 1 < x \leq 2; \\ 3/12, & 2 < x \leq 3; \\ 4/12, & 3 < x \leq 4; \\ 5/12, & 4 < x \leq 5; \\ 6/12, & 5 < x \leq 6; \\ 8/12, & 6 < x \leq 7; \\ 9/12, & 7 < x \leq 8; \\ 11/12, & 8 < x \leq 9; \\ 1, & x > 9 \end{cases}$$



Вычисляем числовые характеристики выборки:

- среднее

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k m_i \hat{x}_i = (1 \cdot 1 + 3 \cdot 2 + 4 \cdot 1 + 5 \cdot 1 + 6 \cdot 1 + 7 \cdot 2 + 8 \cdot 1 + 9 \cdot 2 + 10 \cdot 1) / 12 = 6.278$$

- средний квадрат

$$\overline{x^2} = \frac{1}{n} \sum_{i=1}^k m_i \hat{x}_i^2 = (1 \cdot 1 + 9 \cdot 2 + 16 \cdot 1 + 25 \cdot 1 + 36 \cdot 1 + 49 \cdot 2 + 64 \cdot 1 + 81 \cdot 2 + 100 \cdot 1) / 12 = 43.333$$

- дисперсия

$$\hat{\sigma}^2 = \overline{x^2} - \hat{x}^2 = 43.333 - 39.415 = 3.918$$

- стандартное отклонение

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \sqrt{3.918} = 1.979$$

- исправленная дисперсия

$$s^2 = n\hat{\sigma}^2 / (n-1) = 12 * 1,979 / 11 = 2.159$$

- исправленное стандартное отклонение

$$s = \sqrt{s^2} = \sqrt{2,08} = 1,469$$

Ответ: $\bar{x} = 6.278$; $\hat{\sigma}^2 = 3.918$; $\hat{\sigma} = 1.979$; $s^2 = 2.159$; $s = 1.469$

2

Дано:

$$p(l; \lambda) = P(X=l) = \frac{\lambda^l}{l!} e^{-\lambda}, l=0, 1, 2, \dots;$$

$n = 300$;

| | | | | | | | | |
|-------------|----|----|----|----|----|----|---|---|
| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| \hat{x}_i | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 8 |
| m_i | 29 | 96 | 72 | 63 | 24 | 10 | 4 | 2 |

Найти: $\hat{\lambda}$, \hat{p}_i , $p(\hat{x}_i; \hat{\lambda})$.

Решение.

Выборка содержит $k = 8$ различных элементов. Находим числовые характеристики выборки:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k m_i \hat{x}_i = (29 * 0 + 96 * 1 + 2 * 72 + 3 * 63 + 4 * 24 + 5 * 10 + 4 * 6 + 8 * 2) / 300 = 2,05$$

$$\overline{x^2} = \frac{1}{n} \sum_{i=1}^k m_i \hat{x}_i^2 = (29 * 0 + 72 * 1 + 4 * 96 + 9 * 72 + 16 * 63 + 25 * 10 + 36 * 4 + 64 * 2) / 300 = 8,78$$

$$\hat{\sigma}^2 = \overline{x^2} - \hat{x}^2, 8,78 - 4,20 = 4,578$$

Чтобы получить значения точечных оценок $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\lambda}_3$ параметра λ методом моментов, составляем уравнения

$$v_1 = \hat{v}_1;$$

$$v_2 = \hat{v}_2;$$

$$\mu_2 = \hat{\mu}_2,$$

где теоретические моменты

$$v_1 = M(X) = \lambda, v_2 = M(X^2) = \lambda^2 + \lambda, \mu_2 = D(X) = \lambda;$$

моменты выборки

$$\hat{v}_1 = \bar{x}, \hat{v}_2 = \overline{x^2}, \hat{\mu}_2 = \hat{\sigma}^2.$$

Решая относительно λ уравнения

$$\lambda = \bar{x};$$

$$\lambda^2 + \lambda = \bar{x}^2; \lambda = \hat{\sigma}^2,$$

находим значения точечных оценок

$$\begin{aligned}\hat{\lambda}_1 &= \bar{x} = 2,05; \\ \hat{\lambda}_2 &= (-1 + \sqrt{1 + 4\bar{x}^2})/2 = (-1 + \sqrt{1 + 4 * 8,78})/2 = 2,005 \\ \hat{\lambda}_3 &= \hat{\sigma}^2 = 4,578.\end{aligned}$$

Для нахождения значения точечной оценки $\hat{\lambda}_4$ параметра λ методом наибольшего правдоподобия построим функцию правдоподобия

$$\begin{aligned}L = L(\vec{x}; \lambda) &= \prod_{i=1}^n p(x_i; \lambda) = \prod_{i=1}^{\kappa} p^{m_i}(\hat{x}_i; \lambda) = \prod_{i=1}^{\kappa} \frac{\lambda^{m_i \hat{x}_i}}{(\hat{x}_i!)^{m_i}} e^{-m_i \lambda} = \\ &= \underbrace{[(\hat{x}_1!)^{m_1} (\hat{x}_2!)^{m_2} \dots (\hat{x}_k!)^{m_k}]^{-1}}_{\beta} \lambda^{m_1 \hat{x}_1 + m_2 \hat{x}_2 + \dots + m_k \hat{x}_k} e^{-(m_1 + m_2 + \dots + m_k) \lambda} = \beta \lambda^{n\bar{x}} e^{-n\lambda}.\end{aligned}$$

Отсюда $\ln L = \ln \beta + n\bar{x} \ln \lambda - n\lambda$.

Находим точку экстремума $\lambda = \hat{\lambda}_4$ функции $\ln L$ из уравнения

$$\frac{d \ln L}{d \lambda} = \frac{n\bar{x}}{\lambda} - n = 0$$

которое имеет решение

$$\hat{\lambda}_4 = \bar{x} = 2,05.$$

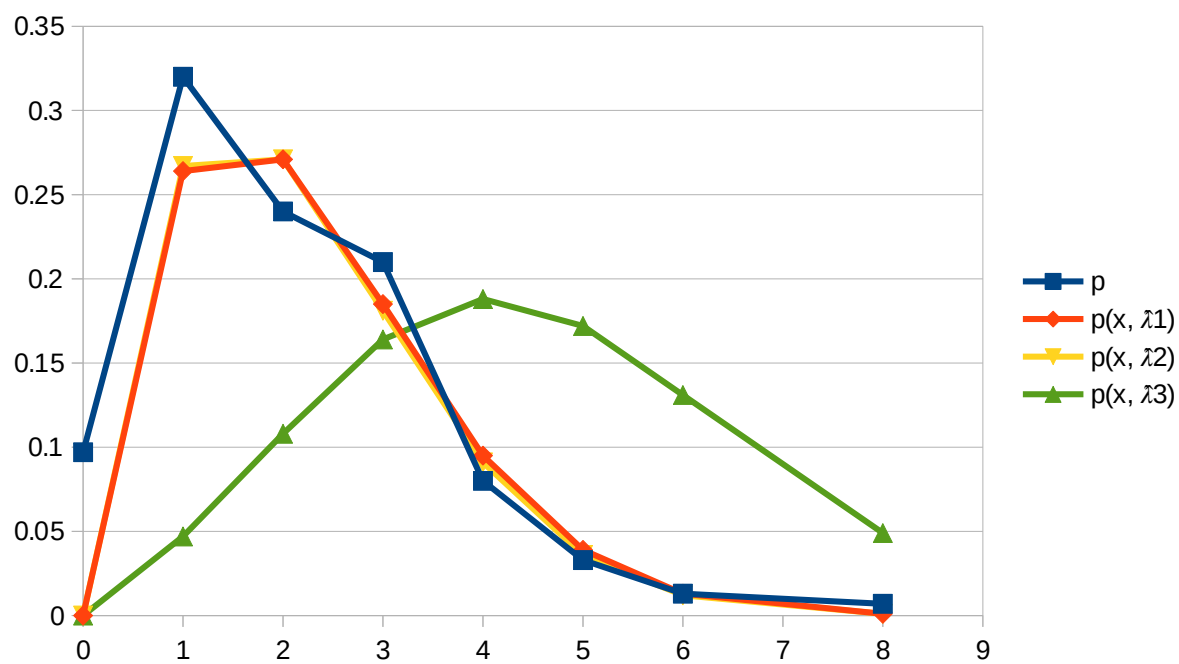
Поскольку

$$\frac{d^2 \ln L}{d \lambda^2} \big|_{\lambda = \hat{\lambda}_4} = \frac{-n\bar{x}}{\hat{\lambda}_4^2} < 0$$

то $\lambda = \hat{\lambda}_4$ является точкой максимума функции $\ln L$ и, соответственно, функции L , а значит, согласно методу наибольшего правдоподобия, $\hat{\lambda}_4$ является значением точечной оценки параметра λ . Можно заметить, что $\hat{\lambda}_4 = \hat{\lambda}_1$.

Для построения полигона относительных частот перепишем статистический ряд для относительных частот $\hat{p}_i = \hat{m}_i/n$:

| | | | | | | | | |
|-------------|-------|------|------|------|------|-------|-------|-------|
| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| \hat{x}_i | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 8 |
| \hat{p}_i | 0,097 | 0,32 | 0,24 | 0,21 | 0,08 | 0,033 | 0,013 | 0,007 |



Ответ: $\lambda_1 = 2,05$; $\lambda_2 = 2,005$; $\lambda_3 = 4,578$; $\lambda_4 = 2,05$.

Дано:

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

$n = 50$;

| | | | | | | | |
|---------------|----|----|----|----|----|----|-----|
| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| \tilde{x}_i | 9 | 27 | 45 | 63 | 81 | 99 | 117 |
| \tilde{m}_i | 31 | 13 | 3 | 1 | 0 | 1 | 1 |

$\gamma = 0,95$.

Найти: $\hat{\lambda}, f(x), f(x; \hat{\lambda}), (\hat{\lambda}_H, \hat{\lambda}_B)$.

Решение.

Данные выборки разбиты на $l = 7$ интервалов. Находим числовые характеристики выборки:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^l m_i \hat{x}_i = (9 * 31 + 27 * 13 + 45 * 3 + 63 * 1 + 81 * 0 + 99 * 1 + 117 * 1) / 50 = 2.88$$

$$\overline{x^2} = \frac{1}{n} \sum_{i=1}^l m_i \hat{x}_i^2 = (31 * 9^2 + 13 * 27^2 + 3 * 45^2 + 1 * 63^2 + 0 * 81^2 + 1 * 99^2 + 1 * 117^2) / 50 = 910.44$$

$$\hat{\sigma}^2 = \overline{x^2} - \bar{x}^2 = 910.44 - 2.88^2 \approx 902.146.$$

Чтобы получить значения точечных оценок $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3$ параметра λ методом моментов, составляем уравнения

$$v_1 = \hat{v}_1;$$

$$v_2 = \hat{v}_2;$$

$$\mu_2 = \hat{\mu}_2,$$

где теоретические моменты

$$v_1 = M(X) = \frac{1}{\lambda}, v_2 = M(X^2) = \frac{2}{\lambda^2}, \mu_2 = D(X) = \frac{1}{\lambda^2},$$

моменты выборки

$$\hat{v}_1 = \bar{x}, \hat{v}_2 = \overline{x^2}, \hat{\mu}_2 = \hat{\sigma}^2.$$

Решая относительно λ уравнения

$$\frac{1}{\lambda} = \bar{x}$$

$$\frac{2}{\lambda^2} = \overline{x^2}$$

$$\frac{1}{\lambda^2} = \hat{\sigma}^2$$

находим значения точечных оценок

$$\begin{aligned}\hat{\lambda}_1 &= 1/\bar{x} = 1/2,88 \approx 0,347; \\ \hat{\lambda}_2 &= \sqrt{2/\bar{x}^2} = \sqrt{2/910,44} \approx 0,047 \\ \hat{\lambda}_3 &= \sqrt{1/\hat{\sigma}^2} = \sqrt{1/902,146} \approx 0,047\end{aligned}$$

Для нахождения значения точечной оценки $\hat{\lambda}_4$ параметра λ методом наибольшего правдоподобия построим функцию правдоподобия

$$\begin{aligned}L = L(\vec{x}; \lambda) &= \prod_{i=1}^n f(x_i; \lambda) \approx \prod_{i=1}^l f^{\tilde{m}_i}(\tilde{x}_i; \lambda) = \prod_{i=1}^l \lambda^{\tilde{m}_i} e^{-\tilde{m}_i \tilde{x}_i \lambda} = \\ &= \lambda^{\tilde{m}_1 + \tilde{m}_2 + \dots + \tilde{m}_l} e^{-(\tilde{m}_1 \tilde{x}_1 + \tilde{m}_2 \tilde{x}_2 + \dots + \tilde{m}_l \tilde{x}_l) \lambda} = \lambda^n e^{-n \bar{x} \lambda}.\end{aligned}$$

Отсюда

$$\ln L = n \ln \lambda - n \bar{x} \lambda.$$

Находим точку экстремума $\lambda = \hat{\lambda}_4$ функции $\ln L$ из уравнения

$$\frac{d \ln L}{d \lambda} = \frac{n}{\lambda} - n \bar{x} = 0$$

которое имеет решение

$$\hat{\lambda}_4 = \frac{1}{\bar{x}} = 0,347$$

Поскольку

$$\frac{d^2 \ln L}{d \lambda^2} \big|_{\lambda = \hat{\lambda}_4} = -\frac{n}{\hat{\lambda}_4^2} < 0$$

то $\lambda = \hat{\lambda}_4$ является точкой максимума функции $\ln L$ и, соответственно, функции L , а значит, согласно методу наибольшего правдоподобия, $\hat{\lambda}_4$ является значением точечной оценки параметра λ . Можно заметить, что $\hat{\lambda}_4 = \hat{\lambda}_1$.

Чтобы построить доверительный интервал для параметра λ , учтём, что статистика

$$V = 2 \lambda n \bar{X}$$

имеет распределение χ^2 с $2n$ степенями свободы. Принимаем $\gamma = 1 - \delta - \varepsilon$, где $\delta = \varepsilon = (1 - \gamma)/2 = (1 - 0,9)/2 = 0,05$.

Поскольку

$$\frac{dV}{d\lambda} = 2n \bar{X} > 0$$

то $V(\vec{X}, \lambda)$ является возрастающей функцией параметра λ . Поэтому нижняя и верхняя границы доверительного интервала определяются из уравнений

$$V(\vec{x}, \hat{\lambda}_H) = 2 \hat{\lambda}_H n \bar{x} = v_\delta = \chi_\delta^2(2n) = \chi_{(1-\gamma)/2}^2(2n),$$

$$V(\vec{x}, \hat{\lambda}_B) = 2 \hat{\lambda}_B n \bar{x} = v_{1-\varepsilon} = \chi_{1-\varepsilon}^2(2n) = \chi_{(1+\gamma)/2}^2(2n),$$

где $\chi_{(1-\gamma)/2}^2(2n) = \chi_{0,05}^2(100) = 77,929$ – квантиль уровня $(1 - \gamma)/2$ распределения χ^2 с $2n$ степенями свободы; $\chi_{(1+\gamma)/2}^2(2n) = \chi_{0,95}^2(100) = 124,342$ – квантиль уровня

$(1 + \gamma)/2$ распределения χ^2 с $2n$ степенями свободы. Отсюда находим

$$\hat{\lambda}_n = \frac{\chi_{(1-\gamma)/2}^2(2n)}{2n\bar{x}} = \frac{77,929}{2 \cdot 50 \cdot 11,424} = 0,068$$

$$\hat{\lambda}_o = \frac{\chi_{(1+\gamma)/2}^2(2n)}{2n\bar{x}} = \frac{124,342}{2 \cdot 50 \cdot 11,424} = 0,109$$

Таким образом, $0,068 < \lambda < 0,109$ с коэффициентом доверия $\gamma = 0,9$.

Для построения графика функции $\tilde{f}(x)$ учтём, что

$$\tilde{f}(x) = \begin{cases} \tilde{p}_i / \Delta x, & x \in J_i, i = \overline{1, l}; \\ 0, & x \notin J, \end{cases}$$

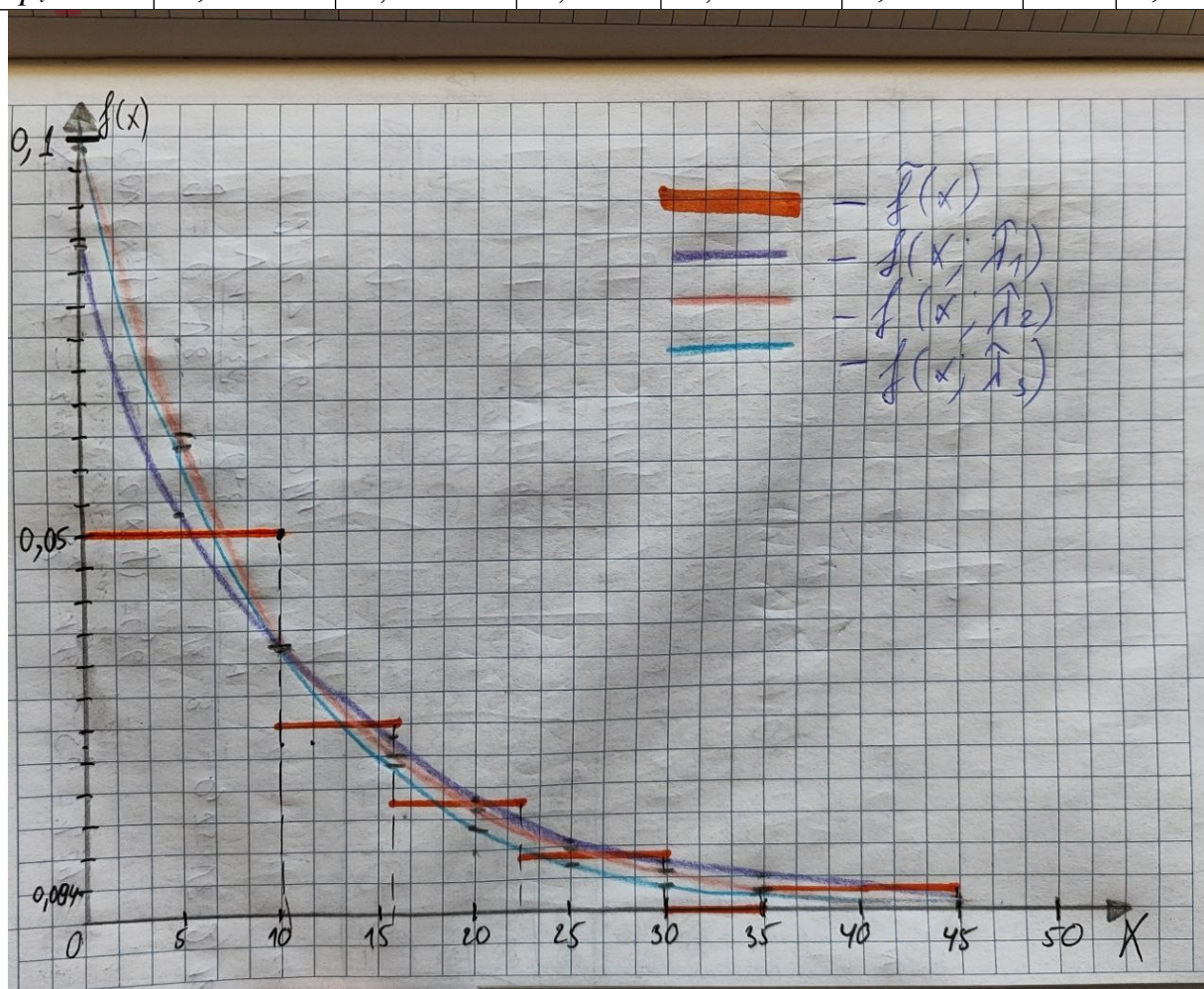
где $\tilde{p}_i = \tilde{m}_i / n$ – относительная частота попадания элементов выборки в интервал J_i ;

$$\Delta x = \hat{x}_{i+1} - \hat{x}_i = 6,72, i = \overline{1, l-1};$$

$$J_i = \begin{cases} [\tilde{x}_i - \Delta x/2, \tilde{x}_i + \Delta x/2), & i = \overline{1, l-1}; \\ [\tilde{x}_l - \Delta x/2, \tilde{x}_l + \Delta x/2], & i = l; \\ J = [\tilde{x}_1 - \Delta x/2, \tilde{x}_l + \Delta x/2], \end{cases}$$

и занесём необходимые числовые значения в таблицу:

| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|--------------------------|--------|--------|--------|--------|--------|---|--------|
| \tilde{p}_i | 0,34 | 0,34 | 0,2 | 0,06 | 0,04 | 0 | 0,02 |
| $\tilde{p}_i / \Delta x$ | 0,0506 | 0,0506 | 0,0298 | 0,0089 | 0,0059 | 0 | 0,0029 |



Ответ: $\hat{\lambda}_1 = 0,087$; $\hat{\lambda}_2 = 0,099$; $\hat{\lambda}_3 = 0,118$; $\hat{\lambda}_4 = 0,087$; $0,068 < \lambda < 0,109$ с коэффициентом доверия $\gamma = 0,9$.