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)»
(КФ МГТУ им. Н.Э. Баумана)***

ФАКУЛЬТЕТ ИУК «Информатика и управление»

КАФЕДРА ИУК4 «Программное обеспечение ЭВМ, информационные технологии»

Практическое занятие №3

«Точечное оценивание»

ДИСЦИПЛИНА: «Методы обработки информации»

Выполнил: студент гр. ИУК4-72Б ________________ (Сафронов Н.С.)
(подпись) (Ф.И.О.)

Проверил: ________________ (Никитенко У.В.)
(подпись) (Ф.И.О.)

Дата сдачи (защиты):

Результаты сдачи (защиты):

- Балльная оценка:

- Оценка:

Калуга, 2023

Постановка задачи

Сгенерировать выборку из 100 элементов, имеющих указанное в вашем варианте распределение. Считая один из параметров распределения неизвестным, найти его точечную оценку:

а) методом моментов (с помощью указанных в задании моментов);

б) методом максимального правдоподобия.

Построить график функции правдоподобия и убедиться, что найденная с помощью метода максимального правдоподобия оценка действительно является точкой максимума функции правдоподобия. Сравнить полученные точечные оценки с истинным значением параметра распределения.

Вариант 14

X - выборка из распределения χ_k^2 , где k = 3. Найти оценку параметра k, считая его неизвестным. Метод моментов реализовать с помощью моментов 1-го и 2-го порядков.

Ход выполнения практического задания

Выпишем формулы для нахождения математического ожидания и дисперсии для распределения χ^2:

$$E\chi^2 = k$$

$$D\chi^2 = 2k$$

Получаем следующие точечные оценки для k:

Для момента 1-го порядка:

$$k^* = \bar{X}$$

Для момента 2-го порядка:

$$k^* = \frac{s^2}{2}$$

Найдём выборочные характеристики распределения:

```
Выборочные показатели
Выборочное среднее: 3.263741116279519
Выборочная дисперсия: 6.751452867399329

Теоретические показатели
Математическое ожидание chi(k=3.00): 3.0
Дисперсия chi(k=3.00): 6.0
```

Рисунок 1 – Выборочные и теоретические показатели распределения

Воспользовавшись методом моментов, найдём точечную оценку параметра k^*:

```
Метод моментов

Точечная оценка по 2-му моменту:
k=3.38
Математическое ожидание chi(k=3.38): 3.3757264336996644
Дисперсия chi(k=3.38): 6.751452867399329

Точечная оценка по 1-му моменту
k=3.26
Математическое ожидание chi(k=3.26): 3.263741116279519
Дисперсия chi(k=3.26): 6.527482232559038
```

Рисунок 2 – Точечные оценки параметра, полученные методом моментов

Построим графики, соответствующие полученным значениям параметра:

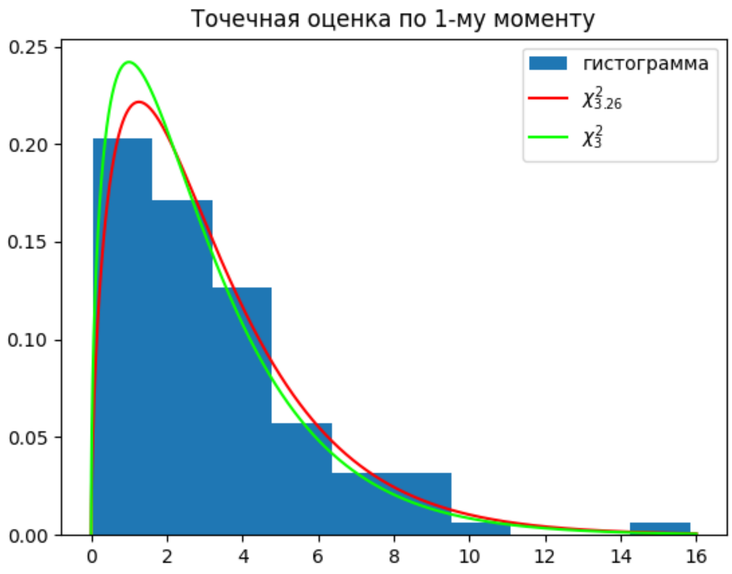

Рисунок 3 – График функции при точечной оценке, полученной по первому моменту

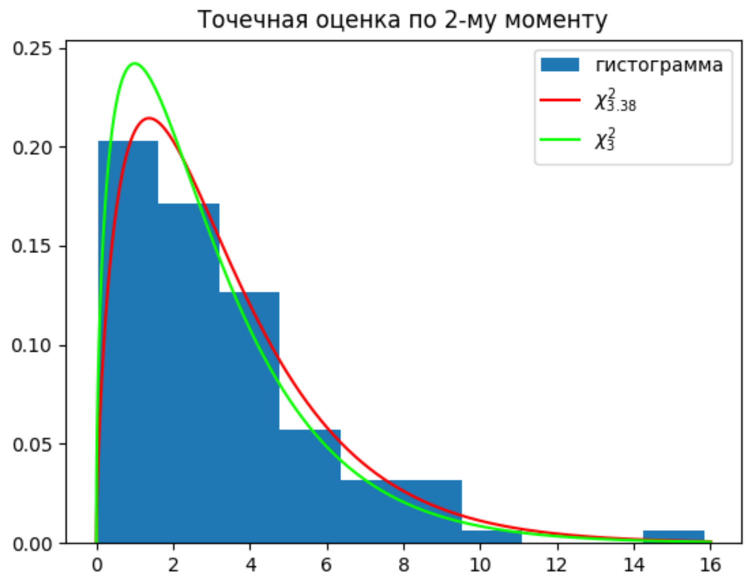

Рисунок 4 – График функции при точечной оценке, полученной по второму моменту

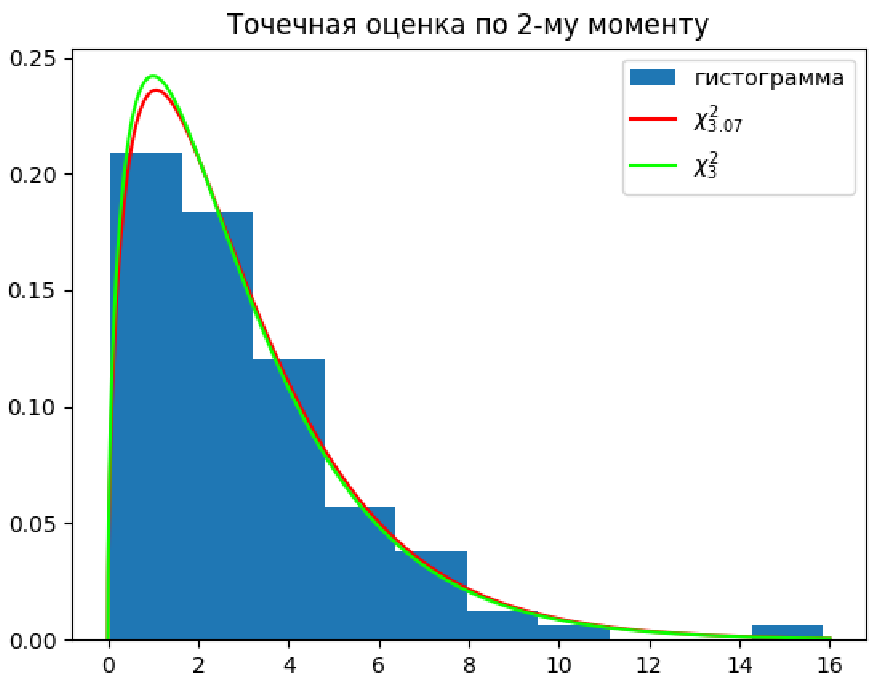

Рисунок 5 – Точечная оценка параметра k^*, вычисленная методом моментов 2-го порядка

Воспользуемся методом максимального правдоподобия.

Построим логарифмическую функцию правдоподобия для заданного распределения:

$$L(k) = \log\big(f(x;k)\big) = \log\left(\prod_{i=1}^{n} f(x_i;k)\right) = \sum_{i=1}^{n} \log\left(\frac{x_i^{\frac{k}{2}-1} e^{-\frac{x_i}{2}}}{\Gamma\left(\frac{k}{2}\right) 2^{\frac{k}{2}}}\right) =$$

$$= \left(\frac{k}{2} - 1\right) \sum_{i=1}^{n} \log x_i - \frac{1}{2} \sum_{i=1}^{n} x_i - n \log\left(\Gamma\left(\frac{k}{2}\right)\right) - \frac{nk}{2} \log 2$$

Построим график зависимости логарифмической функции правдоподобия на заданном промежутке значений k при заданных значениях выборки. Найдём максимальное значение функции и точку, соответствующую ему.

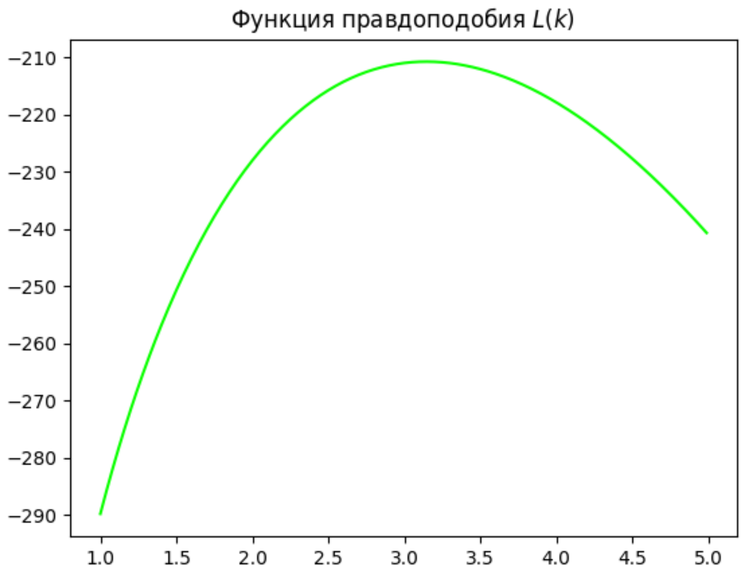

Рисунок 6 – Функция правдоподобия

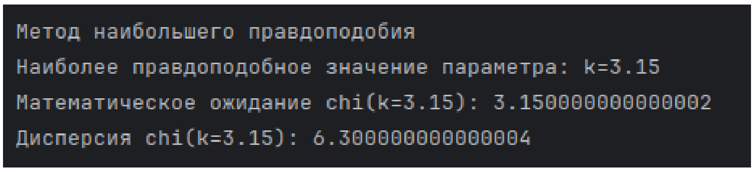
```
Метод наибольшего правдоподобия
Наиболее правдоподобное значение параметра: k=3.15
Математическое ожидание chi(k=3.15): 3.150000000000002
Дисперсия chi(k=3.15): 6.300000000000004
```

Рисунок 7 – Точечная оценка параметра, полученная методом максимального правдоподобия

Построим график, соответствующий полученному значению параметра:

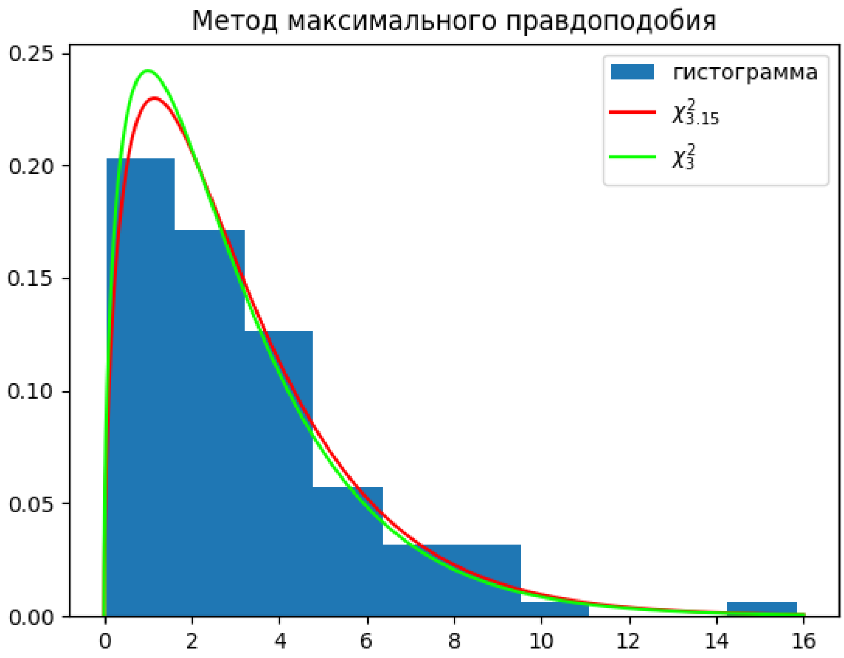

Рисунок 8 – График функции при точечной оценке, полученной методом максимального правдоподобия

Таким образом, получаем, что наиболее точной оказалась оценка, полученная методом моментов по второму моменту.

ПРИЛОЖЕНИЯ

Листинг программы

```
import matplotlib.pyplot as plt
import numpy as np
from scipy.special import gamma
from scipy.stats import chi2


def get_distribution_sample() -> list:
    k = 3
    n = 100
    xs = chi2(k)
    ys = xs.rvs(n)
    return ys


def get_distribution_mean(k: float) -> float:
    return chi2(k).mean()


def get_distribution_variance(k: float) -> float:
    return chi2(k).var()


def get_likelihood_value(xs: list, k: float) -> float:
    xs = np.delete(xs, 0)
    n = len(xs)
    result = (k / 2 - 1) * np.sum(np.log(xs)) - 1 / 2 * np.sum(xs) \
             - n * np.log(gamma(k / 2)) - n * k / 2 * np.log(2)
    return result


if __name__ == "__main__":
    ys = get_distribution_sample()

    sample_mean = np.mean(ys)
    sample_variance = np.var(ys)
    print("\nВыборочные показатели")

    print("Выборочное среднее:", sample_mean)
    print("Выборочная дисперсия:", sample_variance)

    k = 3
    mean = get_distribution_mean(k)
    variance = get_distribution_variance(k)

    print("\nТеоретические показатели")

    print(f"Математическое ожидание chi({k=:.2f}):", mean)
    print(f"Дисперсия chi({k=:.2f}):", variance)

    differences = []

    k = sample_variance / 2
    mean = get_distribution_mean(k)
    variance = get_distribution_variance(k)
    print("\nМетод моментов")
    print("\nТочечная оценка по 2-му моменту:")
    print(f"{k=:.2f}")
    print(f"Математическое ожидание chi({k=:.2f}):", mean)
    print(f"Дисперсия chi({k=:.2f}):", variance)
```

```
plt.hist(ys, density=True, label="гистограмма")
xs = np.arange(0, 16, 0.001)
plt.plot(
    xs, chi2.pdf(xs, df=k), label="$\chi^2_{" + f"{k:.2f}" + "}$",
    color="#fe0000"
)
plt.plot(
    xs, chi2.pdf(xs, df=3), label="$\chi^2_{3}$", color="#0bff01"
)
plt.title("Точечная оценка по 2-му моменту")
plt.legend()
plt.show()

k = sample_mean
mean = get_distribution_mean(k)
variance = get_distribution_variance(k)
print("\nТочечная оценка по 1-му моменту")
print(f"{k=:.2f}")
print(f"Математическое ожидание chi({k=:.2f}):", mean)
print(f"Дисперсия chi({k=:.2f}):", variance)

plt.hist(ys, density=True, label="гистограмма")
xs = np.arange(0, 16, 0.001)
plt.plot(
    xs, chi2.pdf(xs, df=k), label="$\chi^2_{" + f"{k:.2f}" + "}$",
    color="#fe0000"
)
plt.plot(
    xs, chi2.pdf(xs, df=3), label="$\chi^2_{3}$", color="#0bff01"
)
plt.title("Точечная оценка по 1-му моменту")
plt.legend()
plt.show()

ks = np.arange(1, 5, 0.01)
likelihood_values = np.array([get_likelihood_value(ys, k) for k in ks])
max_likelihood_indices = np.argmax(likelihood_values)
k = ks[max_likelihood_indices]

print("\nМетод наибольшего правдоподобия")
print(f"Наиболее правдоподобное значение параметра: {k=:.2f}")

mean = get_distribution_mean(k)
variance = get_distribution_variance(k)
print(f"Математическое ожидание chi({k=:.2f}):", mean)
print(f"Дисперсия chi({k=:.2f}):", variance)

plt.plot(
    ks, likelihood_values, label="Функция правдоподобия $L(k)$",
    color="#0bff01"
)
plt.title("Функция правдоподобия $L(k)$")
plt.show()

plt.hist(ys, density=True, label="гистограмма")
xs = np.arange(0, 16, 0.001)
plt.plot(
    xs, chi2.pdf(xs, df=k), label="$\chi^2_{" + f"{k:.2f}" + "}$",
    color="#fe0000"
)
plt.plot(
    xs, chi2.pdf(xs, df=3), label="$\chi^2_{3}$", color="#0bff01"
)
```

```
plt.title("Метод максимального правдоподобия")
plt.legend()
plt.show()
```