Санкт-Петербургский Национальный Исследовательский

Университет ИТМО

Расчетно-графическая работа №1 По дисциплине

Математическая статистика

Вариант 1

Выполнил:

Солодовников Данила Дмитриевич Группа: P3107

Альметов Тимур Айдарович: P3107

Преподаватель:

Иван Александрович Кононов

Санкт-Петербург 2025 г.

Задание №1

1. В файле iris.csv[1] (здесь и далее ссылки кликабельны) представлены данные о параметрах различных экземплярах цветка ириса. Какой вид в датасете представлен больше всего, какой – меньше? Рассчитайте выборочное среднее, выборочную дисперсию, выборочную медиану и выборочную квантиль порядка $2/5$ для суммарной площади (более точно – оценки площади) чашелистика и лепестка всей совокупности и отдельно для каждого вида. Построить график эмпирической функции распределения, гистограмму и box-plot суммарной площади чашелистика и лепестка для всей совокупности и каждого вида.

Код для задания 1

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from scipy.stats import geom, beta

table = pd.read_csv("iris.csv") # загружаем файл с таблицей
table_size = table.size
vue = table.head()
print(vue)
species_count = table['Species'].value_counts()
print("Количество экземпляров каждого ириса")
print(species_count)
max = species_count.idxmax()
min = species_count.idxmin()
if max == min:
    print("Вида с наибольшим количеством или наименьшим количеством цветков не существует!")
    print("Вид, который представлен чаще всего:", max)
    print("Вид, который представлен реже всего:", min)
else:
    print("Вид, который представлен чаще всего:", max)
    print("Вид, который представлен реже всего:", min)

table['Sepal_square'] = table['Sepal.Length'] * table["Sepal.Width"]
table['Petal_square'] = table['Petal.Length'] * table["Petal.Width"]
table['Square'] = table['Sepal_square'] + table['Petal_square']
print(table['Square'] ) # суммарная площадь

print("Вывод рассчетов для всей площади")
sample_average = table['Square'].mean()
sample_dispersion = table['Square'].var()
sample_median = table['Square'].median()
sample_quantile = table['Square'].quantile(0.4)
print("Выборочное среднее:", sample_average)
print("Выборочная Дисперсия:", sample_dispersion)
print("Выборочная Медиана:", sample_median)
print("Выборочный Квантиль 0.4:", sample_quantile)


grouped = table.groupby("Species")["Square"]
for species, values in grouped:
    print(f"\nВид: {species}")
    print("Выборочное среднее:", values.mean())
    print("Выборочная Дисперсия:", values.var())
    print("Выборочная Медиана:", values.median())
    print("Выборочный Квантиль 0.4:", values.quantile(0.4))
```

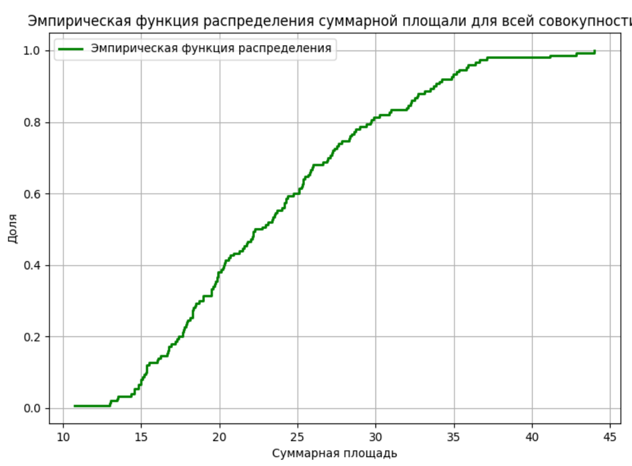
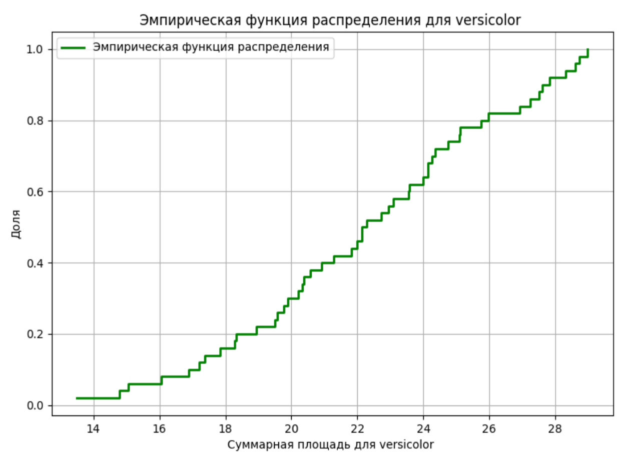
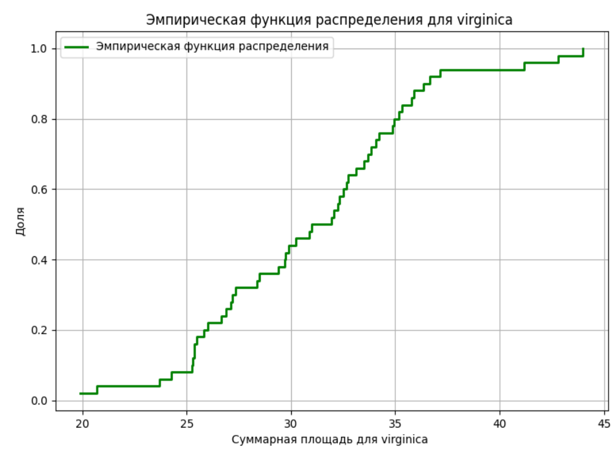
```python
def draw_efr(data, label='', xlabel='Значения', ylabel='Доля', title='Эмпирическая функция распределения'):
    sorted_table = np.sort(data)
    table_size = len(sorted_table)
    y = np.arange(1, table_size + 1) / table_size # доли значений
    plt.figure(figsize=(9, 6))
    plt.step(sorted_table, y, where="post", label=label if label else 'Доля', color='green', linewidth=2)
    plt.xlabel(xlabel)
    plt.ylabel(ylabel)
    plt.title(title)
    plt.grid(True) # сетка на графике
    plt.legend() # подпись
    plt.show()

draw_efr(
    data=table['Square'],  # данные для построения
    label='Эмпирическая функция распределения',
    xlabel='Суммарная площадь',
    title='Эмпирическая функция распределения суммарной площали для всей совокупности '
)

for species in table['Species'].unique():
    species_data = table[table['Species'] == species]['Square']
    draw_efr(
        data=species_data,
        label='Эмпирическая функция распределения',
        xlabel=f'Суммарная площадь для {species} ',
        title=f'Эмпирическая функция распределения для {species} '
    )

def draw_histogram(data, x_column='Square', hue_column=None, title="Гистограмма", xlabel="Значение", ylabel="Частота"):
    plt.figure(figsize=(10, 6))
    sns.histplot(data=data, x=x_column, hue=hue_column, bins=20, kde=False, palette='Set2')
    plt.title(title)
    plt.xlabel(xlabel)
    plt.ylabel(ylabel)
    plt.grid(True)
    plt.show()


draw_histogram(
    data=table,
    x_column='Square',
    hue_column='Species',
    title='Гистограмма суммарной площади по видам',
    xlabel='Суммарная площадь',
    ylabel='Частота'
)

plt.figure(figsize=(6, 5))
plt.boxplot(table['Square'])
plt.title('Boxplot суммарной площади для всей совокупности)')
plt.ylabel('Суммарная площадь')
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.xticks([1], ['Все ирисы'])
plt.show()


def draw_boxplot(dataframe, column='Square', group_column='Species', title='Boxplot по видам'):
    grouped_data = [group[column].values for _, group in dataframe.groupby(group_column)]
    labels = dataframe[group_column].unique()

    plt.figure(figsize=(8, 6))
    plt.boxplot(grouped_data, labels=labels, vert=True)
    plt.xlabel(group_column)
    plt.ylabel(column)
    plt.title(title)
    plt.grid(axis='y', linestyle='--', alpha=0.7)
    plt.show()

table.rename(columns={'Square': 'Суммарная площадь'}, inplace=True)
table.rename(columns={'Species': 'Виды ирисов'}, inplace=True)
draw_boxplot(
    dataframe=table,
    column='Суммарная площадь',
    group_column='Виды ирисов',
    title='Boxplot суммарной площади по видам ирисов'
)
```

Вывод для задания 1

```
   Sepal.Length  Sepal.Width  Petal.Length  Petal.Width Species
0           5.1          3.5           1.4          0.2  setosa
1           4.9          3.0           1.4          0.2  setosa
2           4.7          3.2           1.3          0.2  setosa
3           4.6          3.1           1.5          0.2  setosa
4           5.0          3.6           1.4          0.2  setosa
Количество экземпляров каждого ириса
Species
setosa        50
versicolor    50
virginica     50
Name: count, dtype: int64
Вида с наибольшим количеством или наименьшим количеством цветков не существует!
Вид, который представлен чаще всего: setosa
Вид, который представлен реже всего: setosa
0      18.13
1      14.98
2      15.30
3      14.56
4      18.28
       ...
145    32.06
146    25.25
147    29.90
148    33.50
149    26.88
Name: Square, Length: 150, dtype: float64
Вывод рассчетов для всей площади
Выборочное среднее: 23.616933333333332
Выборочная Дисперсия: 47.909493888143174
Выборочная Медиана: 22.5
Выборочный Квантиль 0.4: 20.316000000000003

Вид: setosa
Выборочное среднее: 17.6234
Выборочная Дисперсия: 8.940059632653064
Выборочная Медиана: 17.66
Выборочный Квантиль 0.4: 16.735999999999997

Вид: versicolor
Выборочное среднее: 22.246600000000004
Выборочная Дисперсия: 15.834108612244899
Выборочная Медиана: 22.21
Выборочный Квантиль 0.4: 21.142

Вид: virginica
Выборочное среднее: 30.9808
Выборочная Дисперсия: 27.00491771428571
Выборочная Медиана: 31.475
Выборочный Квантиль 0.4: 29.716
```

Графики для задания 1

График эмпирической функции распределения

Суммарной площади всей совокупности	Суммарной площади для вида setsoa
	
Суммарной площади для вида versicolor	**Суммарной площади для вида virgnica**
	

Гистограмма для суммарной площади для всей совокупности и для каждого вида

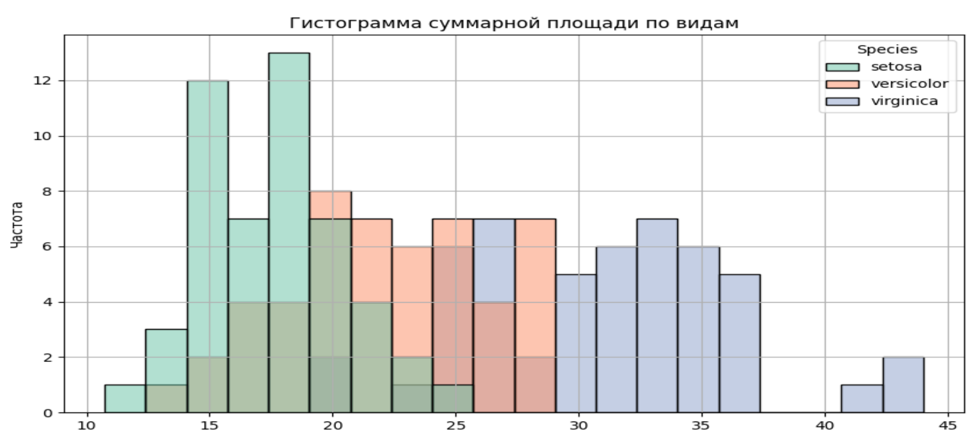
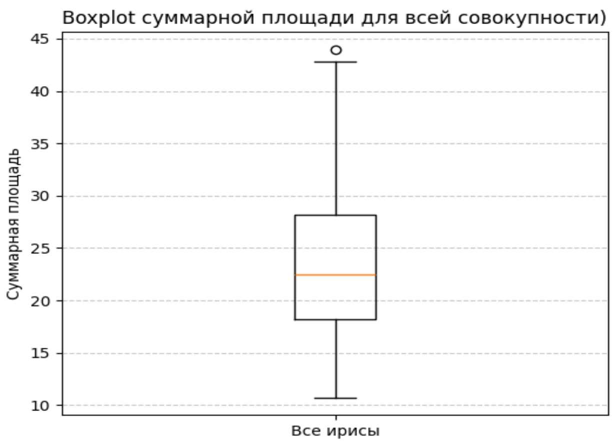
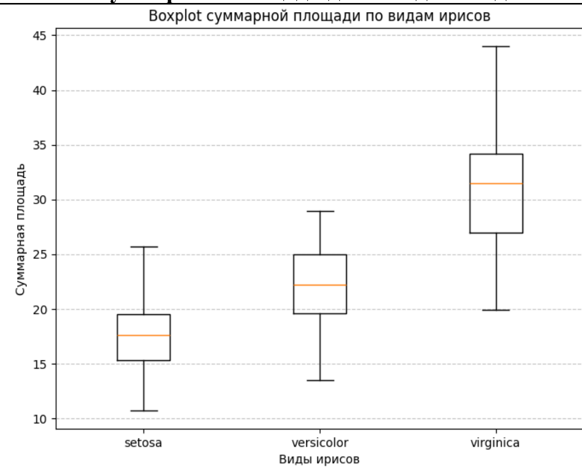


График Box-Plot для суммарной площади для всей совокупности и для каждого вида

<table>
<tr><th>Суммарной площади всей совокупности</th><th>Суммарной площади для каждого вида</th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Задание №2

Задание 2

Предположите, какому вероятностному закону соответствует распределение показателя, рассмотренного (расчет выборочных характеристик и визуализация) в задании №1. Оцените параметры данного распределения методом максимального правдоподобия или методом моментов (**математическое обоснование оценки строго обязательно**). Какими статистическими свойствами обладает найденная оценка (**обосновать**)? Найти **теоретические** смещение, дисперсию, MSE (или хотя бы написать теоретические формулы, по которым данные показатели вычиляются, если в итоге получается «очень сложный» интеграл/ряд), информацию Фишера (если определена для вашей модели).

Задание №2 частично было выполнено в LateX (подробнее в конце файла)

Вычисления:

```
n = table.shape[0] # колючество наблюдений в файле

# Логнормальное распределение
log_square = np.log(table['Square'])
estimated_average_lognormal_distribution_in_log = log_square.mean()
estimated_variance_lognormal_distribution_in_log = ((log_square -
estimated_average_lognormal_distribution_in_log) ** 2).mean()

# Переводим mu и Var в исходные единицы
estimated_average_lognormal_distribution = np.exp(estimated_average_lognormal_distribution_in_log +
estimated_variance_lognormal_distribution_in_log / 2)
estimated_variance_lognormal_distribution = (
    np.exp(estimated_variance_lognormal_distribution_in_log) *
    np.exp(2 * estimated_average_lognormal_distribution_in_log +
estimated_variance_lognormal_distribution_in_log)
    - np.exp(2 * estimated_average_lognormal_distribution_in_log +
estimated_variance_lognormal_distribution_in_log)
)

print("\n--- Логнормальное распределение ---")
print("Оценка среднего:", estimated_average_lognormal_distribution)
print("Оценка дисперсии:", estimated_variance_lognormal_distribution)


# MSE
def mse_formula(sigma_squared, n):
    return (sigma_squared ** 2 / n ** 2) + (2 * sigma_squared ** 2 / n)

mse_lognormal = mse_formula(estimated_variance_lognormal_distribution, n)

print("\n--- MSE дисперсий ---")
print("MSE (логнормальное):", mse_lognormal)

# Bias дисперсий
bias_lognormal_distribution = estimated_variance_lognormal_distribution ** 2 / n

print("\n--- Смещения дисперсий (bias) ---")
print("Bias (логнормальное):", bias_lognormal_distribution)
# Информация Фишера для оценки среднего
info_fisher_mu_lognormal = n / estimated_variance_lognormal_distribution

print("\n--- Информация Фишера для μ ---")
print("Информация Фишера (логнормальное):", info_fisher_mu_lognormal)

# Информация Фишера для оценки дисперсии
info_fisher_sigma2_lognormal = n / (2 * estimated_variance_lognormal_distribution ** 2)

print("\n--- Информация Фишера для σ² ---")
print("Информация Фишера (логнормальное):", info_fisher_sigma2_lognormal)
```

Вывод:

```
--- Логнормальное распределение ---
Оценка среднего: 23.619201777091575
Оценка дисперсии: 48.947513044656944

--- MSE дисперсий ---
MSE (логнормальное): 32.05126973379179

--- Смещения дисперсий (bias) ---
Bias (логнормальное): 15.972393555045743

--- Информация Фишера для μ ---
Информация Фишера (логнормальное): 3.0645070744074063

--- Информация Фишера для σ² ---
Информация Фишера (логнормальное):
0.031304012030310135
```

Задание №3

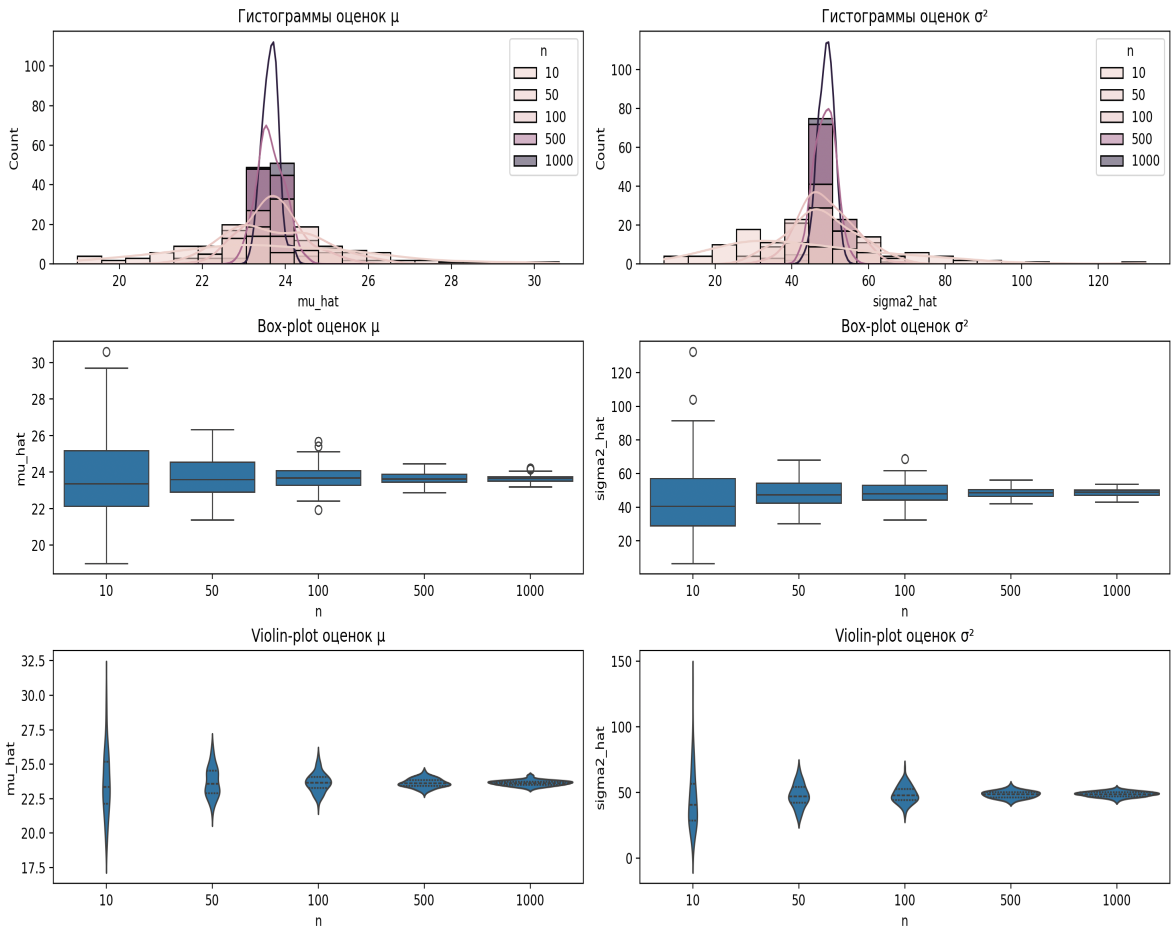
Задание 3

Пусть P_θ – выбранное в предыдущем задании распределение, параметризующееся вектором θ (*пример* – равномерное распределение на $[-2\theta; 4\theta]$, $\hat{\theta} = \overline{X}$), $\hat{\theta}$ – оценка параметра θ, полученная в предыдущем упражнении. Проведите численный эксперимент по следующей схеме:

- Зафиксируйте конкретное значение $\theta = \theta_0$
- Заведите массив $\{n_1, \ldots, n_k\}$ объемов выборки
- Сгенерируйте из распределения P_{θ_0} достаточно большое количество M выборок объёма n, где n принимает значения из массива $\{n_1, \ldots, n_m\}$. Для каждой сгенерированной выборки вычислите оценку $\hat{\theta}$
- Эмпирически рассмотреть поведение оценки $\hat{\theta}$ в зависимости от объема выборки (можно для каждого объема выборки n_i вывести описательные статистики для оценок, изобразить гисторгамму, box-plot, violin-plot).

Код

```
# Заданные параметры
mu_0 = estimated_average_lognormal_distribution
sigma2_0 = estimated_variance_lognormal_distribution
sigma_0 = np.sqrt(sigma2_0)
n_values = [10, 50, 100, 500, 1000]
M = 100

mu_estimates = {n: [] for n in n_values}
sigma2_estimates = {n: [] for n in n_values}

for n in n_values:
    for _ in range(M):
        sample = np.random.normal(mu_0, sigma_0, size=n)
        mu_hat = np.mean(sample)
        sigma2_hat = np.var(sample, ddof=0)  # Смещённая оценка по MLE
        mu_estimates[n].append(mu_hat)
        sigma2_estimates[n].append(sigma2_hat)

# Формируем DataFrame'ы
df_mu = pd.DataFrame([(n, mu) for n in n_values for mu in mu_estimates[n]], columns=["n", "mu_hat"])
df_sigma2 = pd.DataFrame([(n, sigma2) for n in n_values for sigma2 in sigma2_estimates[n]], columns=
["n", "sigma2_hat"])

# Визуализация
fig, axes = plt.subplots(3, 2, figsize=(12, 18), constrained_layout=True)
axes = axes.flatten()

sns.histplot(df_mu, x="mu_hat", hue="n", kde=True, bins=20, ax=axes[0])
axes[0].set_title("Гистограммы оценок μ")

sns.histplot(df_sigma2, x="sigma2_hat", hue="n", kde=True, bins=20, ax=axes[1])
axes[1].set_title("Гистограммы оценок σ²")

sns.boxplot(data=df_mu, x="n", y="mu_hat", ax=axes[2])
axes[2].set_title("Box-plot оценок μ")

sns.boxplot(data=df_sigma2, x="n", y="sigma2_hat", ax=axes[3])
axes[3].set_title("Box-plot оценок σ²")

sns.violinplot(data=df_mu, x="n", y="mu_hat", ax=axes[4], inner="quartile")
axes[4].set_title("Violin-plot оценок μ")

sns.violinplot(data=df_sigma2, x="n", y="sigma2_hat", ax=axes[5], inner="quartile")
axes[5].set_title("Violin-plot оценок σ²")

plt.show()
```


Гистограммы оценок μ
Гистограммы оценок σ²
Count
mu_hat
sigma2_hat
n
10
50
100
500
1000
Box-plot оценок μ
Box-plot оценок σ²
Violin-plot оценок μ
Violin-plot оценок σ²

1 Логнормальное распределение

1.1 Оценка методом максимального правдоподобия

Функция правдоподобия:

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{x_i\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ln x_i - \mu)^2}{2\sigma^2}\right)$$

Логарифм функции правдоподобия:

$$\ell(\mu, \sigma^2) = \sum_{i=1}^{n} \left[-\ln x_i - \frac{1}{2}\ln(2\pi\sigma^2) - \frac{(\ln x_i - \mu)^2}{2\sigma^2}\right]$$

Разделяем слагаемые:

$$\ell(\mu, \sigma^2) = -\sum_{i=1}^{n} \ln x_i - \frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i - \mu)^2$$

Берём производную по μ:

$$\frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(\ln x_i - \mu)$$

Приравниваем к нулю:

$$\sum_{i=1}^{n}(\ln x_i - \mu) = 0 \quad \Rightarrow \quad \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\ln x_i$$

Для оценки σ^2 берём производную:

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(\ln x_i - \mu)^2 = 0$$

Решение относительно σ^2 :

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(\ln x_i - \hat{\mu})^2$$

Mle-оценки параметров логнормального распределения:

Оценки параметров логнормального распределения методом максимального правдоподобия:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\ln x_i,$$

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(\ln x_i - \hat{\mu})^2.$$

2 Свойства:

2.1 Свойства μ:

Несмещенность

MLE-оценка для μ имеет вид:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} \ln x_i.$$

Найдём математическое ожидание:

$$\mathbb{E}[\hat{\mu}] = \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^{n} \ln X_i\right] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[\ln X_i].$$

Так как $X_i \sim \text{Lognormal}(\mu, \sigma^2)$, то $\mathbb{E}[\ln X_i] = \mu$. Тогда:

$$\mathbb{E}[\hat{\mu}] = \frac{1}{n} \sum_{i=1}^{n} \mu = \mu.$$

Таким образом, $\hat{\mu}$ является несмещённой оценкой параметра μ.

Состоятельность

Оценка максимального правдоподобия для μ в логнормальном распределении:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} \ln X_i$$

Дисперсия $\hat{\mu}$ вычисляется как:

$$\text{Var}[\hat{\mu}] = \text{Var}\left(\frac{1}{n} \sum_{i=1}^{n} \ln X_i\right)$$

Так как наблюдения X_i независимы:

$$\text{Var}[\hat{\mu}] = \frac{1}{n^2} \sum_{i=1}^{n} \text{Var}[\ln X_i] = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}$$

Поскольку:

- $\mathbb{E}[\hat{\mu}] = \mu$ (несмещенность)
- $\text{Var}[\hat{\mu}] = \frac{\sigma^2}{n} \to 0$ при $n \to \infty$

Согласно закону больших чисел:

$$\hat{\mu} \xrightarrow{P} \mu \quad \text{при} \quad n \to \infty$$

Таким образом, $\hat{\mu}$ является состоятельной оценкой параметра μ логнормального распределения.

Асимптотическая нормальность MLE-оценки $\hat{\mu}$

Доказательство

Пусть $X_1, X_2, \ldots, X_n$ – независимые одинаково распределённые случайные величины, где $X_i \sim \text{Lognormal}(\mu, \sigma^2)$. Тогда:

- $\ln X_i \sim N(\mu, \sigma^2)$
- Оценка максимального правдоподобия для μ:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \ln X_i$$

По центральной предельной теореме при $n \to \infty$:

$$\frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1)$$

Оценка $\hat{\mu}$ вычисляется как выборочное среднее логарифмов наблюдений:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \ln X_i.$$

Если $\ln X_i \sim N(\mu, \sigma^2)$, то оценка $\hat{\mu}$ имеет нормальное распределение:

$$\hat{\mu} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

Применяя стандартную нормировку, получаем:

$$\frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1).$$

Это означает, что для больших n оценка $\hat{\mu}$ асимптотически нормальна:

$$\hat{\mu} \approx N\left(\mu, \frac{\sigma^2}{n}\right), \quad \text{при } n \to \infty.$$

Таким образом, MLE-оценка $\hat{\mu}$ является асимптотически нормальной.

2.2 Свойства $\widehat{\sigma^2}$:

Несмещённость. Доказательство

MLE-оценка дисперсии логарифма:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (\ln X_i - \hat{\mu})^2.$$

Рассмотрим математическое ожидание:

$$E[\hat{\sigma}^2] = E\left[\frac{1}{n}\sum_{i=1}^{n}(\ln X_i - \hat{\mu})^2\right].$$

Известно, что несмещённая оценка дисперсии строится по формуле:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\ln X_i - \hat{\mu})^2.$$

Таким образом, MLE-оценка $\hat{\sigma}^2$ смещена, и её можно сделать несмещённой, умножив на $\frac{n}{n-1}$:

$$S^2 = \frac{n}{n-1}\hat{\sigma}^2.$$

Следовательно оценка дисперсии , MLE- смещена вниз, но можно получить несмещённую оценку домножив её на , $\frac{n}{n-1}$

Несмещённость. Доказательство

$$\mathrm{Var}[\hat{\sigma}^2] = \frac{1}{n}\left[E\left((\ln X_i - \mu)^4\right) - \sigma^4\right].$$

Поскольку $E\left((\ln X_i - \mu)^4\right)$ конечно (как центральный момент нормального распределения), получаем:

$$\mathrm{Var}[\hat{\sigma}^2] = O\left(\frac{1}{n}\right).$$

Следовательно, $\hat{\sigma}^2$ сходится по вероятности к σ^2:

$$\hat{\sigma}^2 \xrightarrow{p} \sigma^2 \quad \text{при} \quad n \to \infty.$$

Таким образом, $\hat{\sigma}^2$ является **состоятельной оценкой**.

Асимптотическая нормальность. Доказательство

MLE-оценка дисперсии для лог-нормального распределения имеет вид:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(\ln X_i - \hat{\mu})^2.$$

Напомним, что $\ln X_i \sim N(\mu, \sigma^2)$, а $\hat{\mu}$ - несмещённая оценка математического ожидания:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\ln X_i.$$

Для больших n оценка $\hat{\sigma}^2$ асимптотически нормальна. Известно, что:

$$\sqrt{n}\left(\frac{\hat{\sigma}^2}{\sigma^2}-1\right) \xrightarrow{d} N(0,2).$$

Это означает, что при $n \to \infty$:

$$\hat{\sigma}^2 \approx N\left(\sigma^2, \frac{2\sigma^4}{n}\right).$$

Таким образом, MLE-оценка $\hat{\sigma}^2$ асимптотически нормальна.

3 Оценки максимального правдоподобия

Для выборки $X_1, X_2, \dots, X_n \sim \text{Lognormal}(\mu, \sigma^2)$ оценки максимального правдоподобия параметров:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \ln X_i$$
$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n} (\ln X_i - \hat{\mu})^2$$

4 Теоретические свойства оценок

4.1 Смещение (Bias)

- Для $\hat{\mu}$:
$$\text{Bias}(\hat{\mu}) = \mathbb{E}[\hat{\mu}] - \mu = 0$$
- Для $\widehat{\sigma^2}$:
$$\text{Bias}(\widehat{\sigma^2}) = \mathbb{E}[\widehat{\sigma^2}] - \sigma^2 = -\frac{\sigma^2}{n}$$

4.2 Дисперсия оценок

- Дисперсия $\hat{\mu}$:
$$\text{Var}(\hat{\mu}) = \frac{\sigma^2}{n}$$
- Дисперсия $\widehat{\sigma^2}$:
$$\text{Var}(\widehat{\sigma^2}) = \frac{2(n-1)\sigma^4}{n^2} \approx \frac{2\sigma^4}{n} \text{ при } n \to \infty$$

4.3 Среднеквадратическая ошибка (MSE)

- Для $\hat{\mu}$:
$$\text{MSE}(\hat{\mu}) = \text{Var}(\hat{\mu}) + \text{Bias}^2(\hat{\mu}) = \frac{\sigma^2}{n}$$
- Для $\widehat{\sigma^2}$:
$$\text{MSE}(\widehat{\sigma^2}) = \frac{(2n-1)\sigma^4}{n^2} \approx \frac{2\sigma^4}{n} \text{ при } n \to \infty$$

Информация Фишера для логнормального распределения

Для выборки $X_1, X_2, \ldots, X_n \sim \mathrm{LN}(\mu, \sigma^2)$, где $\ln X_i \sim N(\mu, \sigma^2)$, информация Фишера вычисляется как:

$$I(\theta) = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\log L(\theta)\right]$$

Информация Фишера для оценки $\hat{\mu}$:

Вторая производная логарифма правдоподобия по μ:

$$\frac{\partial^2}{\partial\mu^2}\log L(\mu, \sigma^2) = -\frac{n}{\sigma^2}$$

$$I(\hat{\mu}) = \frac{n}{\sigma^2}$$

Информация Фишера для оценки $\widehat{\sigma^2}$:

Вторая производная логарифма правдоподобия по σ^2:

$$\frac{\partial^2}{\partial(\sigma^2)^2}\log L(\mu, \sigma^2) = \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{i=1}^{n}(\ln X_i - \mu)^2$$

$$I(\widehat{\sigma^2}) = \frac{n}{2\sigma^4}$$