

Федеральное государственное автономное образовательное учреждение высшего
образования «Национальный исследовательский университет ИТМО»

Лабораторная №1

По дисциплине «Математическая статистика»

Поленов Кирилл Александрович
Р3213

Пименова Екатерина Андреевна
Р3213

Санкт-Петербург
2025 год

Содержание

1. Задача №1 да	3
2. Задача №2	5
2.1. Предположение распределения	5
2.2. Оценка параметров методом максимального правдоподобия (ММП)	5
2.3. Статистические свойства	6
2.3.1. Смещение	6
2.3.1.1. $\hat{\mu}$:	6
2.3.1.2. $\hat{\sigma}^2$:	6
2.3.2. Дисперсия	6
2.3.2.1. $\hat{\mu}$	6
2.3.2.2. $\hat{\sigma}^2$	6
2.3.3. MSE	6
2.3.3.1. $\hat{\mu}$	6
2.3.3.2. $\hat{\sigma}^2$	6
2.3.4. Информация Фишера	6
3. Задача №3	7

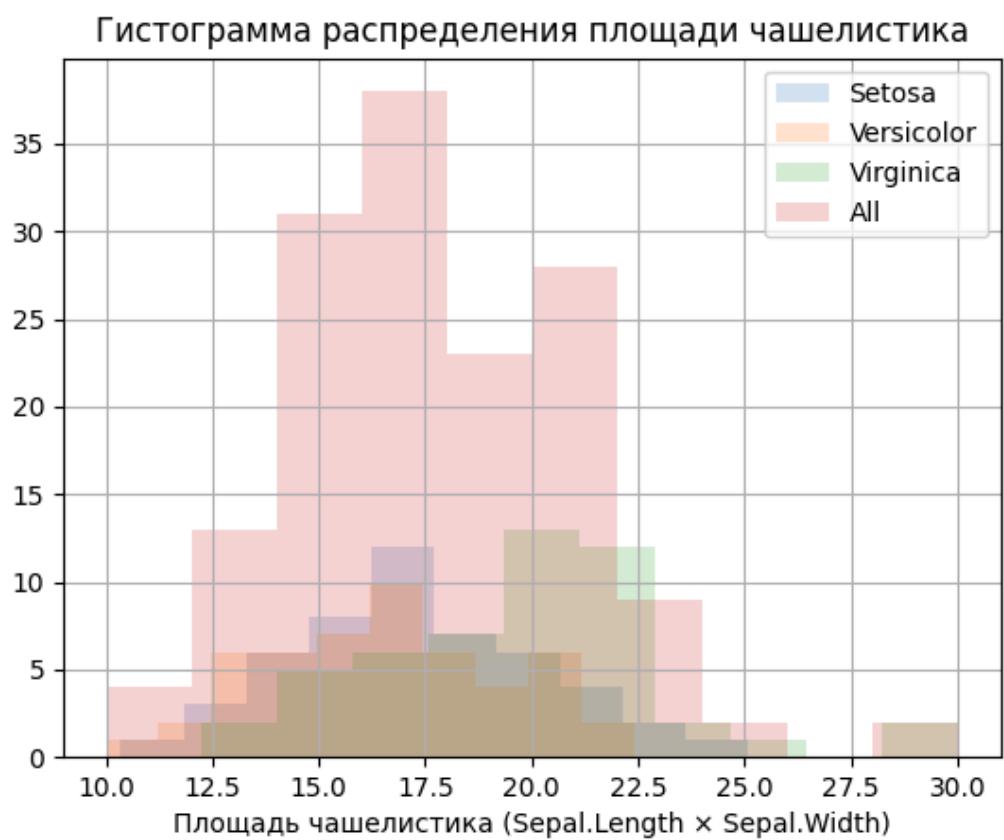
1. Задача №1 да

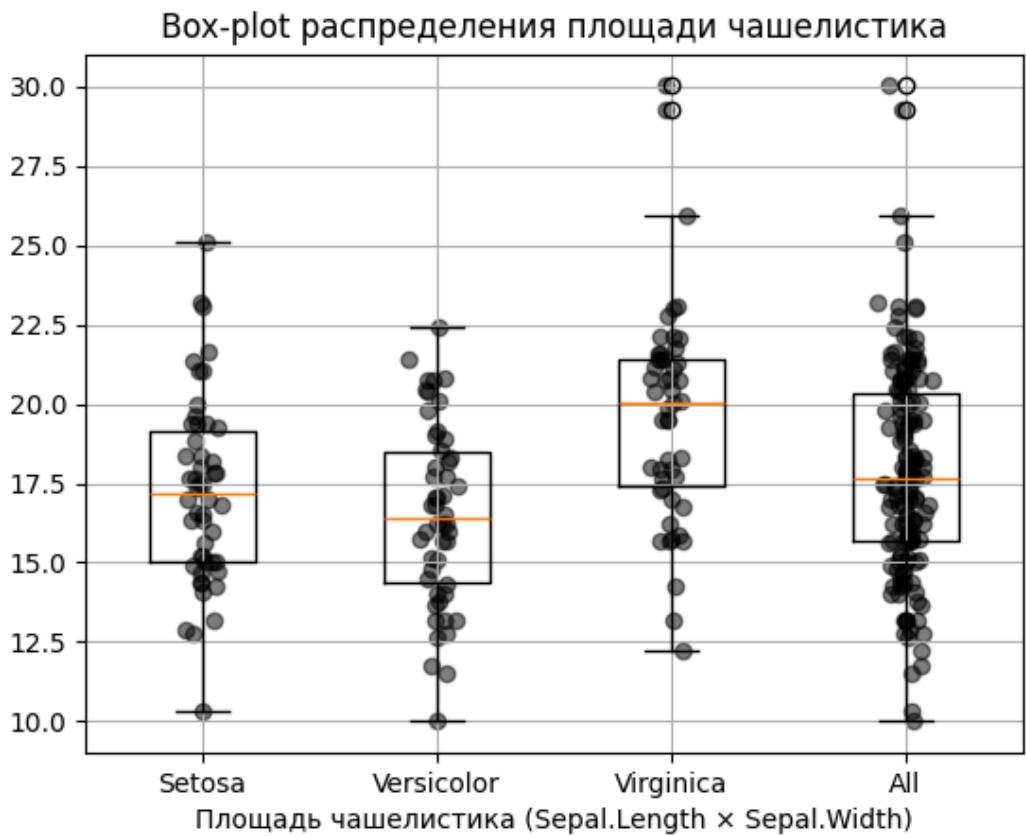
<i>setosa</i>	$E(x)$	$D(x)$	Q	M
Площадь чашелистика	17.258	8.435	16.428	15.0
Площадь лепестока	0.366	0.032	0.292	0.32
Площадь чашелистика и лепестка	8.812	75.57	0.486	10.35

<i>versicolor</i>	$E(x)$	$D(x)$	Q	M
Площадь чашелистика	16.526	8.055	15.876	19.8
Площадь лепестока	5.72	1.835	5.408	6.16
Площадь чашелистика и лепестка	11.123	34.136	6.75	10.0

<i>virginica</i>	$E(x)$	$D(x)$	Q	M
Площадь чашелистика	19.685	11.724	19.02	23.04
Площадь лепестока	11.296	4.561	10.32	10.8
Площадь чашелистика и лепестка	15.49	25.734	13.164	12.25

All flowers	$E(x)$	$D(x)$	Q	M
Площадь чашелистика	17.823	11.227	16.818	19.8
Площадь лепестока	5.794	22.059	4.524	6.16
Площадь чашелистика и лепестка	11.808	52.816	10.152	17.85





2. Задача №2

2.1. Предположение распределения

Предположим, что данное это гауссовское распределение

2.2. Оценка параметров методом максимального правдоподобия (ММП)

$$L(\mu, \sigma^2) = \prod_{i=1}^n f(x_i)$$

где $f(x) = \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

Дальше рассмотрим вот такое:

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

Оценим μ :

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

Оценим σ^2 :

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

2.3. Статистические свойства

2.3.1. Смещение

2.3.1.1. $\hat{\mu}$:

$$E[\hat{\mu}] = \mu \Rightarrow \text{несмешенная}$$

т.к. $\text{Bias}(\hat{\mu}) = E[\hat{\mu}] - \mu$

2.3.1.2. $\hat{\sigma}^2$:

$$E[\hat{\sigma}^2] = \sigma^2 \cdot \frac{n-1}{n} \Rightarrow \text{смещение: } \text{Bias}(\hat{\sigma}^2) = -\frac{\sigma^2}{n}$$

Т.к. $\frac{n}{\sigma^2} \hat{\sigma}^2 \sim \mathcal{X}^2(n-1)$, т.е. на 1 степень свободы меньше

Для устранения смещения используют:

$$\hat{\sigma}_{\text{несм}}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu})^2$$

2.3.2. Дисперсия

2.3.2.1. $\hat{\mu}$

Т.к. $\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$, где x_i независимы, $\text{Var}(x_i) = \sigma^2$

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

2.3.2.2. $\hat{\sigma}^2$

$$\text{Var}(\hat{\sigma}^2) = 2\sigma^4 \frac{n-1}{n^2}$$

Т.к. $\sum_{i=1}^n (x_i - \hat{\mu})^2 \sim \sigma^2 \mathcal{X}^2(n-1)$, потому что $\hat{\mu}$ теряет одну степень свободы

2.3.3. MSE

2.3.3.1. $\hat{\mu}$

$$\text{MSE}(\hat{\mu}) = \text{Var}(\hat{\mu}) + \text{Bias}(\hat{\mu})^2 = \frac{\sigma^2}{n}$$

2.3.3.2. $\hat{\sigma}^2$

$$\text{MSE}(\hat{\sigma}^2) = \text{Var}(\hat{\sigma}^2) + \text{Bias}(\hat{\sigma}^2)^2 = 2\sigma^4 \frac{n-1}{n^2} + \left(-\frac{\sigma^2}{n}\right)^2$$

2.3.4. Информация Фишера

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

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

3. Задача №3

$$M = 1000, \quad n = \{5, 10, 50, 100, 150\}, \quad \theta_0 = 17, 82$$

