



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 №5

«Проверка гипотез»

ДИСЦИПЛИНА: «Методы обработки информации»

Выполнил: студент гр. ИУК4-72Б _____ (_____)
(подпись) (Ф.И.О.)

Проверил: _____ (_____)
(подпись) (Ф.И.О.)

Дата сдачи (защиты):

Результаты сдачи (защиты):

- Балльная оценка:

- Оценка:

Калуга, 2023

Постановка задачи

Пусть проверяется простая гипотеза относительно параметра распределения $H_0: \theta = \theta_0$, с заданным уровнем значимости α .

Для нескольких альтернативных гипотез $H_1: \theta = \theta_{1i}$, при $\theta_{1i} = \theta_0 + i\Delta (i = 1, 2, 3, 4, 5)$. Построить графики мощности критерия значимости, если используется выборка (выборка из ПЗ-2):

1. объема $k_1 = 25$ (любые 25 значений из заданной выборки);
2. объема $k_2 = N$ (полный объем исходной выборки) Используя полученные результаты, построить таблицы "Ошибка II рода и мощность для нескольких альтернативных гипотез с объемом выборки k_i и α " и графики функций мощности критерия для случая 1 и 2.

Ход выполнения практического задания

За нулевую гипотезу будем брать среднее значение выборки.

Уровень значимости $\alpha = 0.1$.

Все значения выборки:

Значение параметра распределения	Мощность критерия	Ошибка II рода
0.1461757984121978	0.4004016061579658	0.5995983938420342
0.22031553737164683	0.7642757723355271	0.23572422766447287
0.2944552763310958	0.9571518535622745	0.0428481464377255
0.36859501529054484	0.9967206201316634	0.003279379868336596
0.44273475424999387	0.999899775068096	0.00010022493190398052

Рисунок 1 – Ошибки II рода и мощность критерия для полной выборки

Любые 25 из заданной выборки:

Значение параметра распределения	Мощность критерия	Ошибка II рода
0.1461757984121978	0.24078266609083318	0.7592173339091668
0.22031553737164683	0.35300632649265795	0.646993673507342
0.2944552763310958	0.508814786124697	0.491185213875303
0.36859501529054484	0.6712869326824206	0.32871306731757943
0.44273475424999387	0.8081769602965775	0.19182303970342252

Рисунок 3 – Ошибки II рода и мощность критерия для малой выборки

($N = 25$)

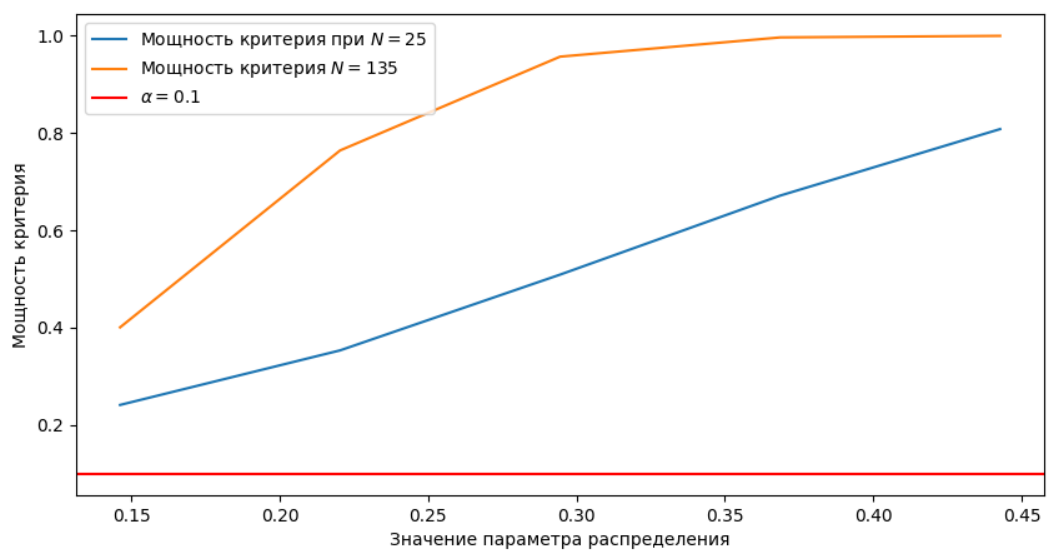


Рисунок 3 – График мощности критерия значимости для малой и полной выборок

ПРИЛОЖЕНИЯ

Листинг программы

```
import argparse
import csv

import matplotlib.pyplot as plt
import numpy as np
from prettytable import PrettyTable
from scipy.stats import norm

def get_test_power(
    data: list[float], theta_0: float, alpha: float, theta: float,
    sample_size: int
) -> float:
    critical_value = norm.ppf(1 - alpha)
    standard_error = np.std(data) / np.sqrt(sample_size)
    critical_region = (
        theta_0 - critical_value * standard_error,
        theta_0 + critical_value * standard_error
    )
    power = 1 - norm.cdf(critical_region[1], theta, standard_error) +
    norm.cdf(
        critical_region[0], theta, standard_error
    )
    return power

def read_points(path: str) -> list[float]:
    points = []

    with open(path, newline='') as file:
        reader = csv.reader(file, delimiter=' ', quotechar='|')
        for row in reader:
            points.append(float(''.join(row)))

    return points

def solve(points: list[float]):
    alpha = 0.1
    theta_0 = np.mean(points)
    delta = np.std(points) / np.sqrt(len(points))
    thetas = [theta_0 + i * delta for i in range(1, 6)]

    small_sample_power = np.array(
        [
            get_test_power(
                points, theta_0, alpha, theta, 25
            ) for theta in thetas
        ]
    )
    full_sample_power = np.array(
        [
```

```

        get_test_power(
            points, theta_0, alpha, theta, len(points)
        ) for theta in thetas
    ]
)

table = PrettyTable()
table.add_column('Значение параметра распределения', thetas)
table.add_column('Мощность критерия', small_sample_power)
table.add_column('Ошибка II рода', 1 - small_sample_power)
print('Любые 25 из заданной выборки:')
print(table)

table = PrettyTable()
table.add_column('Значение параметра распределения', thetas)
table.add_column('Мощность критерия', full_sample_power)
table.add_column('Ошибка II рода', 1 - full_sample_power)
print('Все значения выборки:')
print(table)

plt.figure(figsize=(10, 5))
plt.plot(thetas, small_sample_power, label='$Мощность критерия при
N=25$')
plt.plot(
    thetas, full_sample_power, label=f'Мощность критерия
$N={len(points)}$'
)
plt.xlabel('Значение параметра распределения')
plt.ylabel('Мощность критерия')
plt.title('График мощности критерия')
plt.axhline(
    alpha, color='red',
    label=f"$\\alpha={alpha}$"
)
plt.legend()
plt.show()

if __name__ == '__main__':
    parser = argparse.ArgumentParser()
    parser.add_argument('-file')

    args = parser.parse_args()
    file = args.file or './data/Test14.csv'

    points = read_points(file)
    solve(points)

```