

## Лабораторная работа № 1

### «Проверка статистических гипотез»

студента Баранова Александра группы Б22-534. Дата сдачи: 01.11.2024

Ведущий преподаватель: Новиков М. А. оценка: \_\_\_\_\_ подпись: \_\_\_\_\_

### Вариант №2

*Цель работы:* изучение функций Statistics and Machine Learning Toolbox™ MATLAB / Python SciPy.stats для проверки статистических гипотез.

#### 1. Исходные данные

Характеристики наблюдаемых случайных величин:

| СВ    | Распределение | Параметры          | Математическое ожидание, $m_i$  | Дисперсия, $\sigma_i^2$                               |
|-------|---------------|--------------------|---------------------------------|-------------------------------------------------------|
| $X_1$ | $R(2,6)$      | $a_1 = 2, b_1 = 6$ | $m_1 = \frac{a_1 + b_1}{2} = 4$ | $\sigma_1^2 = \frac{(b_1 - a_1)^2}{12} = \frac{4}{3}$ |
| $X_2$ | $R(2,6)$      | $a_2 = 2, b_2 = 6$ | $m_2 = \frac{a_2 + b_2}{2} = 4$ | $\sigma_2^2 = \frac{(b_2 - a_2)^2}{12} = \frac{4}{3}$ |

*Указание:* для генерации случайных чисел использовать функции **rand**, **randn**, **chi2rnd** (**scipy.stats: uniform.rvs, norm.rvs, chi2.rvs**)

Выборочные характеристики:

| СВ            | Среднее, $\bar{x}_i$ | Оценка дисперсии, $s_i^2$ | Оценка с.к.о., $s_i$ | Объем выборки, $n_i$ |
|---------------|----------------------|---------------------------|----------------------|----------------------|
| $X_1$         | 4.00                 | 1.34                      | 1.16                 | 50                   |
| $X_2$         | 3.85                 | 1.70                      | 1.31                 | 50                   |
| <i>Pooled</i> | 3.92                 | 1.51                      | 1.23                 | 100                  |

**Указание:** для расчета использовать функции **mean**, **var**, **std** (**scipy.stats.describe**)

## 2. Однопараметрические критерии

Для случайной величины  $X_1$ :

| Тест                              | Стат. гипотеза,<br>$H_0$                             | Выборочное<br>значение<br>статистики<br>критерия | $p$ -value | Стат.<br>решение при<br>$\alpha = 0.05$ | Ошибка<br>стат.<br>решения |
|-----------------------------------|------------------------------------------------------|--------------------------------------------------|------------|-----------------------------------------|----------------------------|
| z-test                            | $m_1 = 4$<br>( $\sigma$ известна)                    | -0.01                                            | 0.99       | $H_0$ принимается                       | Нет                        |
| t-test                            | $m_1 = 4$<br>( $\sigma$ неизвестна)                  | -0.01                                            | 0.99       | $H_0$ принимается                       | Нет                        |
| $\chi^2$ -test<br>(m – изв)       | $\sigma_1 = \sqrt{\frac{4}{3}}$<br>( $m$ известно)   | 49.23                                            | 0.50       | $H_0$ принимается                       | Нет                        |
| $\chi^2$ -test<br>(m – не<br>изв) | $\sigma_1 = \sqrt{\frac{4}{3}}$<br>( $m$ неизвестно) | 49.23                                            | 0.46       | $H_0$ принимается                       | Нет                        |

**Указание:** для проверки гипотез использовать функции **ztest**, **ttest**, **vartest** (**scipy.stats: ttest\_1samp, chisquare**)

## 3. Двухвыборочные критерии

Для случайных величин  $X_1, X_2$ :

| Тест               | Стат. гипотеза,<br>$H_0$ | Выборочное<br>значение<br>статистики<br>критерия | $p$ -value | Стат.<br>решение при<br>$\alpha = 0.05$ | Ошибка<br>стат.<br>решения |
|--------------------|--------------------------|--------------------------------------------------|------------|-----------------------------------------|----------------------------|
| 2-sample<br>t-test | $m_1 = m_2$              | 0.62                                             | 0.54       | $H_0$ принимается                       | Нет                        |

|                                       |                       |      |      |                   |     |
|---------------------------------------|-----------------------|------|------|-------------------|-----|
| 2-sample<br>F-test<br>(m – изв)       | $\sigma_1 = \sigma_2$ | 0.78 | 0.81 | $H_0$ принимается | Нет |
| 2-sample<br>F-test<br>(m – не<br>изв) | $\sigma_1 = \sigma_2$ | 0.77 | 0.82 | $H_0$ принимается | Нет |

Указание: для проверки гипотез использовать функции **ttest2**, **vartest2** (**scipy.stats: ttest\_ind, chisquare**)

#### 4. Исследование распределений статистик критерия

Статистическая гипотеза:  $H_0: m_1 = 4(\sigma_1\text{- известна})$

Формула расчёта статистики критерия  $Z$ : 
$$Z = \frac{\bar{X} - m_0}{\frac{\sigma}{\sqrt{n}}}$$

Формула расчёта статистики  $P\text{-value}$ :

$$p = 2\min(F_Z(z_{\text{выб}} | H_0), 1 - F_Z(z_{\text{выб}} | H_0))$$

Число серий экспериментов  $N = 1000$

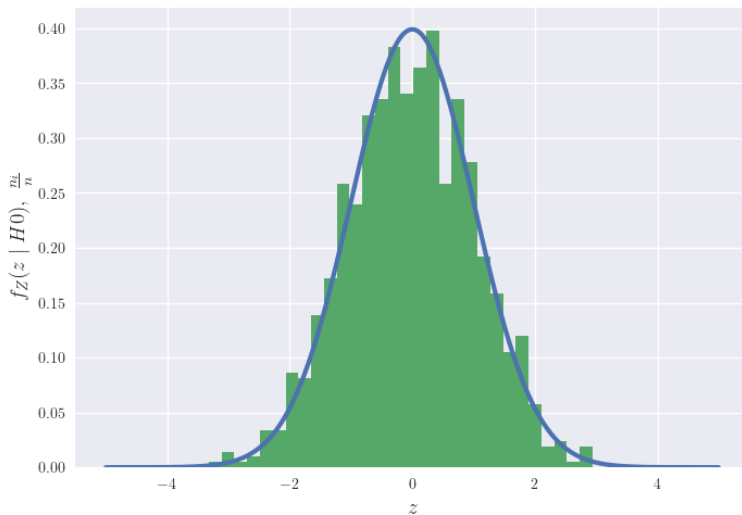
Теоретические характеристики:

| СВ               | Распределение<br>в условиях $H_0$ | Параметры | Математическое<br>ожидание | Дисперсия                 | С.к.о.                           |
|------------------|-----------------------------------|-----------|----------------------------|---------------------------|----------------------------------|
| $Z$              | $N(0, 1)$                         |           | 0                          | 1                         | 1                                |
| $P\text{-value}$ | $R(0, 1)$                         |           | $\frac{1}{2}$              | $\frac{1}{12} \sim 0.083$ | $\frac{1}{\sqrt{12}} \sim 0.288$ |

Выборочные характеристики:

| СВ | Среднее | Оценка дисперсии | Оценка с.к.о. |
|----|---------|------------------|---------------|
|----|---------|------------------|---------------|

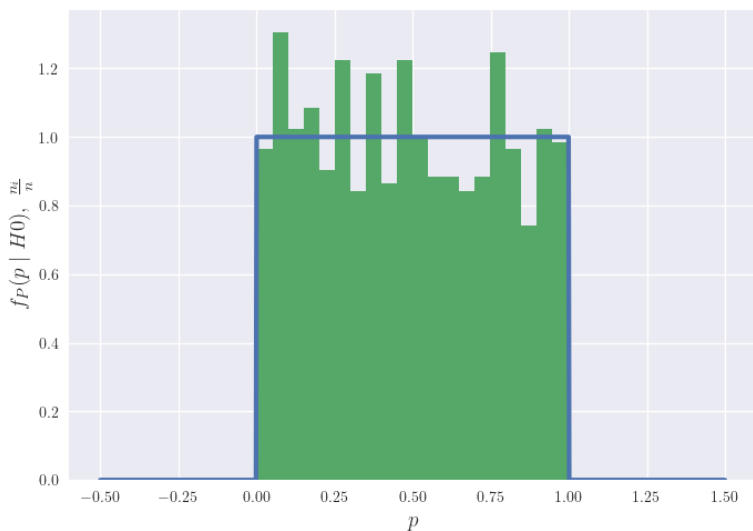
|                  |       |      |      |
|------------------|-------|------|------|
| $Z$              | -0.03 | 1.05 | 1.03 |
| $P\text{-value}$ | 0.49  | 0.08 | 0.29 |



**Указание:** при расчете выборочных значений статистики критерия использовать функции **norminv**, **tinvt**, **chi2inf**, **finvt** (**scipy.stats**: **norm.ppf**, **t.ppf**, **chi2.ppf**, **f.ppf**)

Гистограмма частот статистики  $Z$  и теоретическая функция  $f_Z(z | H_0)$ :

Гистограмма частот статистики  $P\text{-value}$  и теоретическая функция  $f_p(p|H_0)$ :



**Указание:** для построения гистограмм и теоретических функций плотности использовать функции **hist**, **normpdf**, **tpdf**, **chi2pdf**, **fpdf** (**scipy.stats: norm.pdf, t.pdf, chi2.pdf, f.pdf, histogram; matplotlib.pyplot: hist**)