

Лабораторная работа № 1
«Проверка статистических гипотез»

студента Розинко Е.Д. группы Б21-524. Дата сдачи: 05.11.23
Ведущий преподаватель: Трофимов А.Г. оценка: _____ подпись: _____

Вариант № 4

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, σ_i^2
X_1	R(5,15)	R(a, b)	10	8,3
X_2	N(10, 5)	N(m, σ)	10	25

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	9.959	8.555	2.925	300
X_2	10.179	24.118	4.911	100
<i>Pooled</i>	10.014	12.426	6.213	400

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

2. Однопараметрические критерииДля случайной величины X_1 :

Тест	Стат. гипотеза, H_0	Выборочное значение статистики критерия	p -value	Стат. решение при $\alpha = 0,05$	Ошибка стат. решения
z-test	$m = 10,02$	-0.375	0.708	H_0 принимается	нет
t-test	$m = 9,89$	0.411	0.681	H_0 принимается	нет
χ^2 -test (m – изв)	$\sigma = 3$	285.205	0.557	H_0 принимается	нет
χ^2 -test (m – не изв)	$\sigma = 2,8$	326.250	0.267	H_0 принимается	нет

Указание: для проверки гипотез использовать функции **ztest**, **ttest**, **vartest** (**scipy.stats**: **ttest_1samp**, **chisquare**)

3. Двухвыборочные критерииДля случайных величин X_1, X_2 :

Тест	Стат. гипотеза, H_0	Выборочное значение статистики критерия	p -value	Стат. решение при $\alpha = 0,05$	Ошибка стат. решения
2-sample t-test	$m_1 = m_2$	-0.538	0.590	H_0 принимается	нет
2-sample F-test (m – изв)	$\sigma_1 = \sigma_2$	0.354	0	H_0 отклоняется	нет
2-sample F-test (m – не изв)	$\sigma_1 = \sigma_2$	0.355	0	H_0 отклоняется	нет

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

4. Исследование распределений статистик критерия

Статистическая гипотеза: $H_0: m_1 = m_2$ (σ_1, σ_2 – не изв.)

Формула расчёта статистики критерия Z :
$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Формула расчёта статистики P -value: $p = 2 * \min(F_Z(z), 1 - F_Z(z))$

Число серий экспериментов $N = 1000$

Теоретические характеристики:

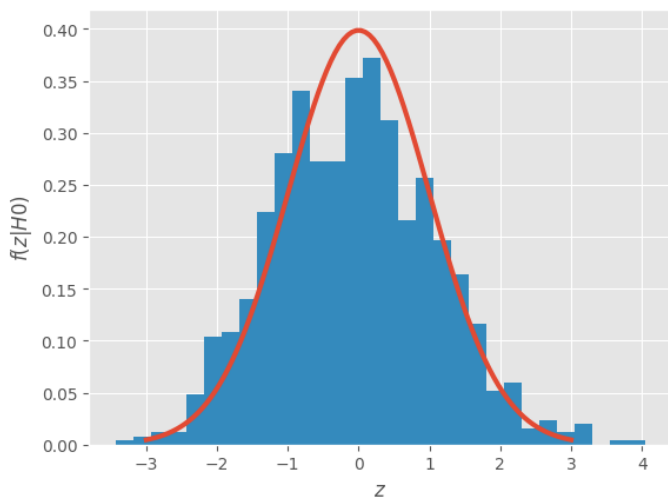
СВ	Распределение в условиях H_0	Параметры	Математическое ожидание	Дисперсия	С.к.о.
Z	$T(398)$	$T(n_1 + n_2 - 2)$	$m = 0$	$\sigma^2 = 0,000192$	$\sigma = 0,0138$
P -value	$R(0, 1)$	$R(a,b)$	$m = 0,5$	$\sigma^2 = 0,08$	$\sigma = 0,29$

Выборочные характеристики:

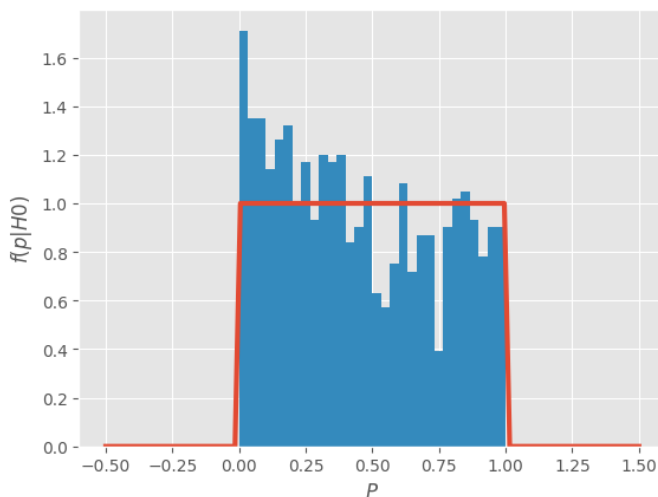
СВ	Среднее	Оценка дисперсии	Оценка с.к.о.
Z	-0.066	1.299	1.14
P -value	0.45	0.089	0.298

Указание: при расчете выборочных значений статистики критерия использовать функции **norminv**, **tinv**, **chi2inf**, **finv** (**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 -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**)