

Лабораторная работа № 1
«Проверка статистических гипотез»

студент Когановский Г.И. группы Б22-534.

Дата сдачи: 31.10.24

Ведущий преподаватель: Трофимов А.Г.

оценка:

подпись: _____

Вариант №7

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	$N(-1, 2)$	$m_1 = -1, \sigma_1 = 2$	$m_1 = -1$	$\sigma_1^2 = 4$
X_2	$R(-2, 0)$	$a = -2, b = 0$	$m_2 = -1$	$\sigma_2^2 = \frac{1}{3}$

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	$m_1 = -0.78$	$s_1 = 3.98$	$s_1 = 1.99$	$n_1 = 100$
X_2	$m_2 = -0.97$	$s_2 = 0.33$	$s_2 = 0.57$	$n_2 = 100$
<i>Pooled</i>	-0.88	2.16	1.47	200

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

2. Однопараметрические критерии

Для случайной величины X_1 :

Тест	Стат. гипотеза, H_0	Выборочное значение статистики критерия	p -value	Стат. решение при $\alpha = 0.01$	Ошибка стат. решения
z-test	$X_1 \sim N(m_1, \sigma_1)$	-3.87	0.99	Принимается H_0	-
t-test	$M(X_1) = m_1$	1.11	0.13	Принимается H_0	-
χ^2 -test (m – изв)	$X_1 \sim N(m_1, s_1)$	10.6	0.23	Принимается H_0	-
χ^2 -test (m – не изв)	$X_1 \sim N(\bar{X}_1, s_1)$	8.91	0.26	Принимается H_0	-

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, H_0	Выборочное значение статистики критерия	p -value	Стат. решение при $\alpha = 0.01$	Ошибка а стат. решения
2-sample t-test	$M(X_1) = M(X_2)$	0.93	0.18	Принимается H_0	-
2-sample F-test (m – изв)	$D(X_1) = D(X_2)$	12.33	$1.11 \cdot 10^{-16}$	Принимается H_0	-
2-sample F-test (m – не изв)	$D(X_1) = D(X_2)$	12.21	$1.11 \cdot 10^{-16}$	Принимается H_0	-

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

4. Исследование распределений статистик критерия

Статистическая гипотеза: $H_0: \sigma_1 = 2$ (m_1 – не изв.)

Формула расчёта статистики критерия $Z: \chi^2 = \frac{(n-1)s^2}{\sigma^2}$

Формула расчёта статистики P -value: $f(x) = \begin{cases} \frac{x^{k/2-1}e^{-x/2}}{2^{k/2}\Gamma(k/2)}, & x \geq 0 \\ 0, & x < 0 \end{cases}$

Число серий экспериментов $N = 1000$

Теоретические характеристики:

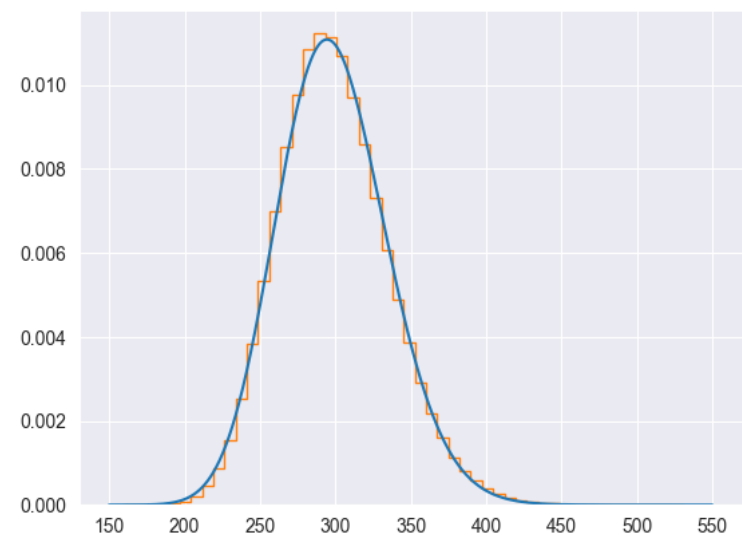
СВ	Распределение в условиях H_0	Параметры	Математическое ожидание	Дисперсия	С.к.о.
Z	$Z \sim \chi^2(n-1)$		$n-1 = 99$	$2n-2 = 198$	$\sqrt{2n-2} = 14.07$
P -value	P -value $\sim R(0,1)$		$1/2 = 0.5$	$1/12 \sim 0.083$	$1/\sqrt{12} \sim 0.288$

Выборочные характеристики:

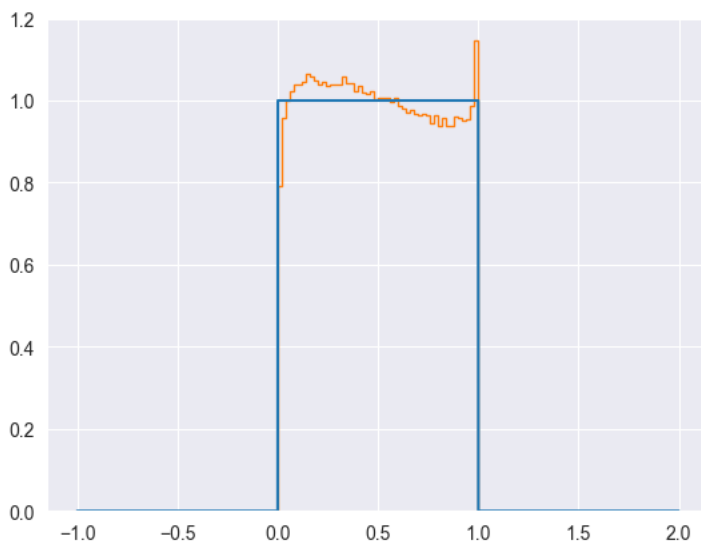
СВ	Среднее	Оценка дисперсии	Оценка с.к.о.
Z	98.5	202.0	14.21
P -value	0.49	0.086	0.293

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 $P\text{-value}$ 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**)