

Лабораторная работа № 1
«Проверка статистических гипотез»

студента Грачева Игоря группы Б23-504 . Дата сдачи: _____
Ведущий преподаватель: _____ оценка: _____ подпись: _____

Вариант №5

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	$N(5, 2)$	$\mu=5, \sigma=2$	5	4
X_2	$N(5, 2)$	$\mu=5, \sigma=2$	5	4

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	4.92795	5.28676	2.29929	100
X_2	4.93663	4.69015	2.16568	100
<i>Pooled</i>	4.93229	4.98845	2.23348	200

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	$m = 5$	-0.36026	0.71865	Не отвергаем H_0	нет
t-test	$m = 5$	-0.31337	0.75466	Не отвергаем H_0	нет
χ^2 -test (m – изв)	$\sigma^2 = 4$	130.97704	0.04095	Отвергаем H_0	Ошибка 1 рода
χ^2 -test (m – не изв)	$\sigma^2 = 4$	130.84725	0.03536	Отвергаем H_0	Ошибка 1 рода

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	$m_1 = m_2$	-0.02747	0.97811	Не отвергаем H_0	нет
2-sample F-test (m – изв)	$\sigma_1^2 = \sigma_2^2$	1.12735	0.55007	Не отвергаем H_0	нет
2-sample F-test (m – не изв)	$\sigma_1^2 = \sigma_2^2$	1.12720	0.55250	Не отвергаем H_0	нет

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

4. Исследование распределений статистик критерия

Статистическая гипотеза: $H_0: \sigma_1 = \sigma_2$

Формула расчёта статистики критерия Z : $Z \sim F(99, 99)$ при H_0

Формула расчёта статистики P -value: $P\text{-value} = 2(1 - F(Z))$

Число серий экспериментов $N = 10000$

Теоретические характеристики:

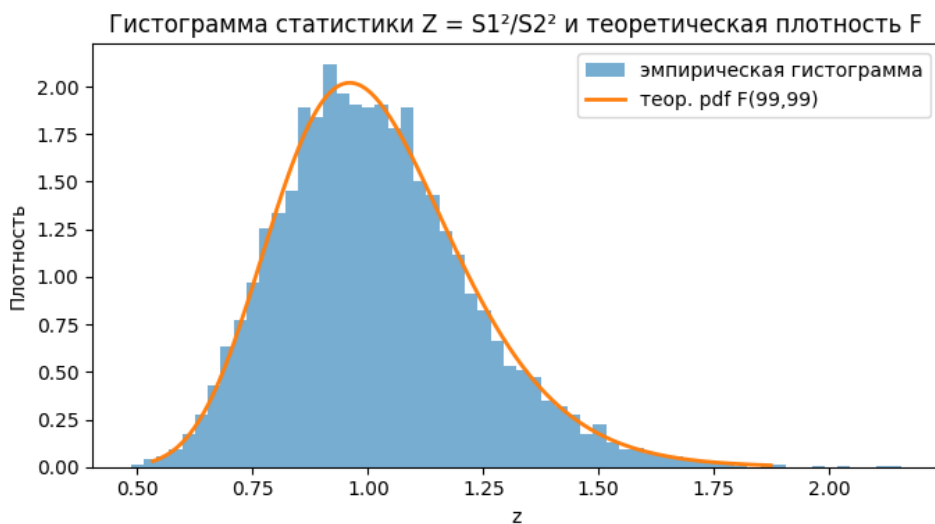
СВ	Распределение в условиях H_0	Параметры	Математическое ожидание	Дисперсия	С.к.о.
Z	$F(99, 99)$	$Df1 = 99, df2 = 99$	1.02062	0.04342	0.20837
$P\text{-value}$	$U(0, 1)$	$A = 0, b = 1$	0.50000	0.08333	0.28868

Выборочные характеристики:

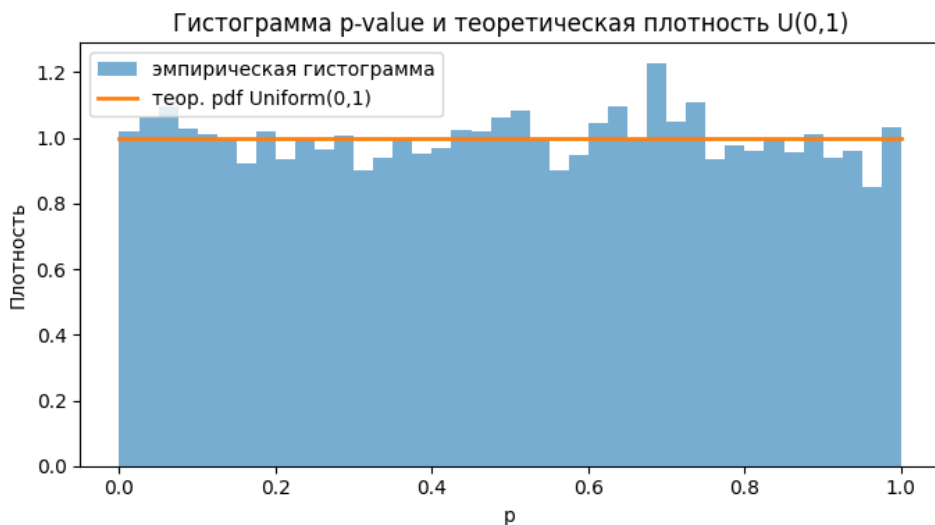
СВ	Среднее	Оценка дисперсии	Оценка с.к.о.
Z	1.01696	0.04334	0.20818
$P\text{-value}$	0.49837	0.08288	0.28788

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 $f_P(p|H_0)$:



Лабораторный практикум по курсу «Математическая статистика»

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**)