

Лабораторная работа № 1
«Проверка статистических гипотез»

студента Моисеенко О.И. группы Б20-514. Дата сдачи: _____
Ведущий преподаватель: Трофимов А.Г. оценка: _____ подпись: _____

Вариант №15

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	χ^2	15	15	30
X_2	R	5, 25	$(5 + 25)/2 = 15$	$(25 - 5)^2 / 12 = 33, (3) = 100/3$

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	15.124385	25.014042	5.001404	100
X_2	14.851133	36.878963	6.072805	100
<i>Pooled</i>	14.987759	30.965169	5.564636	

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	$m_1 = 15$	0.2474535993 268442	0.80455719 73480057	H_0 принимается	нет
t-test	$m_1 = 15$	0.2474535993 268442	0.80506967 16182391	H_0 принимается	нет
χ^2 -test (m – изв)	$\sigma = 5.477226$	83.431711622 50556	0.23189681 682241764	H_0 принимается	нет
χ^2 -test (m – не изв)	$\sigma = 5.477226$	83.380139503 84718	0.06499655 778255227	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	$m_1 = m_2$	0.3455887363 152111	0.73001870 11064814	H_0 принимается	нет
2-sample F-test (m – изв)	$\sigma_1 = \sigma_2$	0.6782859375 489307	0.01339942 9742594738	H_0 отклоняется	нет
2-sample F-test (m – не изв)	$\sigma_1 = \sigma_2$	0.6782740076 869322	0.01370067 7212369387	H_0 отклоняется	нет

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

4. Исследование распределений статистик критерия

Статистическая гипотеза: $H_0: m_I = 15$ (σ_1 – известна)

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

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

Число серий экспериментов $N = 1000$

Теоретические характеристики:

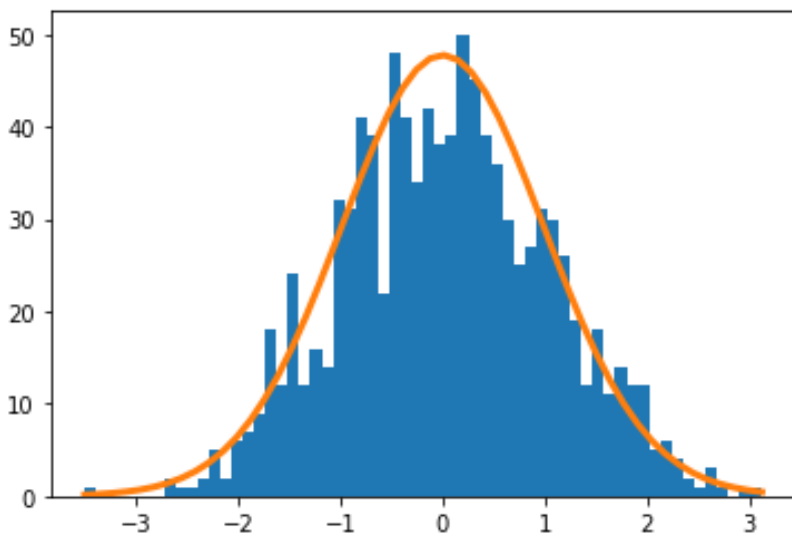
СВ	Распределение в условиях H_0	Параметры	Математическое ожидание	Дисперсия	С.к.о.
Z	$N(0,1)$	$m = 0$ $\sigma^2 = 1$	0	1	1
P -value	$R(0,1)$	$a = 0$ $b = 1$	0.5	0.0833	0.2887

Выборочные характеристики:

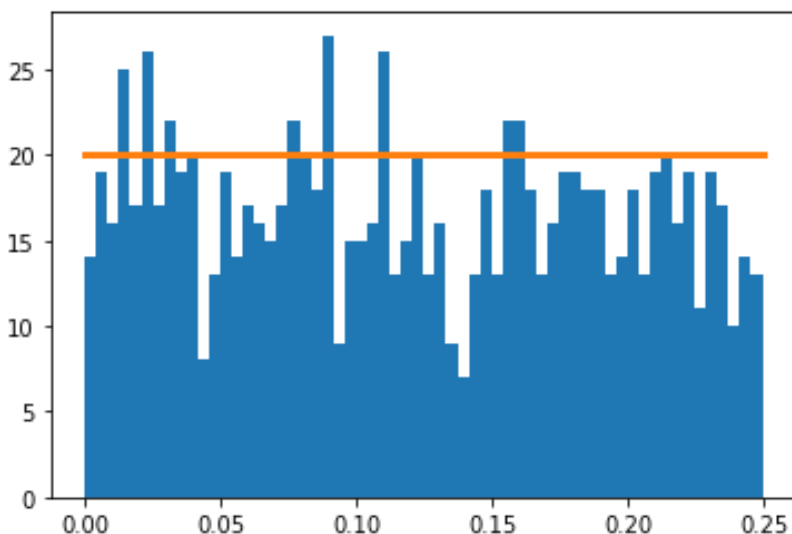
СВ	Среднее	Оценка дисперсии	Оценка с.к.о.
Z	0.034713191568054344	1.0509120563751322	1.0251400179366388
P -value	0.12099650636525473	0.0052127261673303015	0.07219921168080924

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**)