

Лабораторная работа № 2

«Критерии согласия и однородности выборок»

студента Никитина А.Д. группы Б23-504. Дата сдачи: 10.11.2025

Ведущий преподаватель: _____ оценка: _____ подпись: _____

Вариант №13

Цель работы: изучение функций Statistics and Machine Learning Toolbox™ MATLAB / Python SciPy.stats для проверки критериев согласия (*goodness-of-fit tests*) и однородности выборок.

1. Исходные данные

Характеристики наблюдаемой случайной величины X :

Распределение	Параметры	Математическое ожидание, m	Дисперсия, σ^2
$X \sim R(-5, 5)$	$a = -5,$ $b = 5$	$m_1 = \frac{a_1 + b_1}{2} = 0$	$\sigma_1^2 = \frac{(b_1 - a_1)^2}{12}$ $= \frac{25}{3}$

Объём выборки $n_1 = 300$

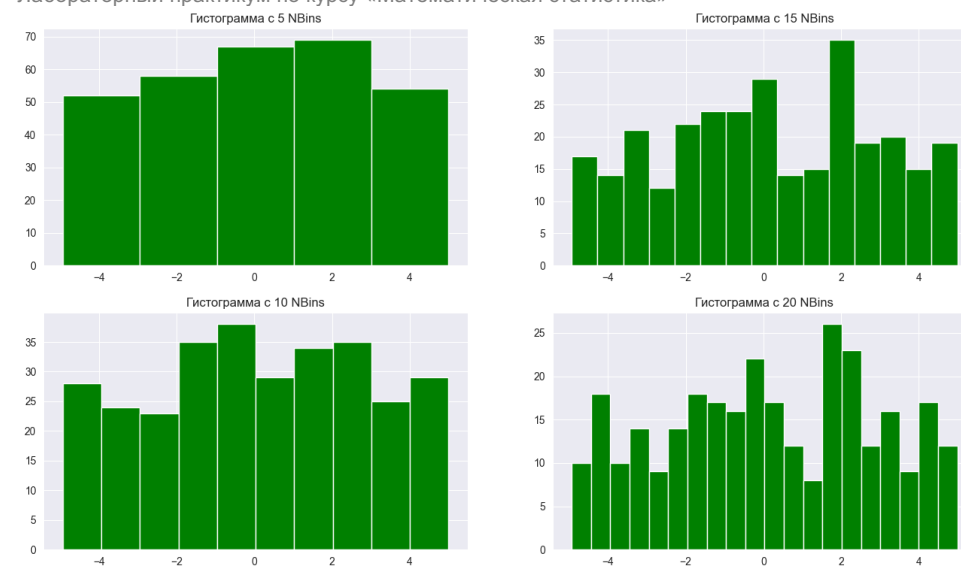
Примечание: для генерации случайных чисел использовать функции **rand**, **randn**, **chi2rnd** (**scipy.stats: uniform.rvs, norm.rvs, chi2.rvs**)

Выборочные характеристики:

Среднее, \bar{x}	Оценка дисперсии, s^2	Оценка с.к.о., s
0.14	7.46	2.73

2. Визуальное представление выборки

Гистограммы частот:



Примечание: для построения гистограмм использовать функцию **hist** (**scipy.stats: histogram; matplotlib.pyplot: hist**)

3. Критерий хи-квадрат

а) Статистическая гипотеза: $H_0: X \sim N(m, \sigma)$

Число интервалов группировки	Выборочное значение статистики критерия	p -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
5	157.33	0.00	H_0 принимается	2-го рода
10	166.46	0.00	H_0 принимается	2-го рода
15	200.85	0.00	H_0 принимается	2-го рода
20	210.00	0.00	H_0 принимается	2-го рода

б) Статистическая гипотеза: $H_0: X \sim R$

Число интервалов группировки	Выборочное значение статистики критерия	p -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
5	3.90	0.14	H_0 принимается	нет
10	8.20	0.31	H_0 принимается	нет

15	26.99	0.007	H_0 отклоняется	1-го рода
20	31.06	0.019	H_0 отклоняется	1-го рода

в) Статистическая гипотеза: $H_0: X \sim \chi^2(5)$

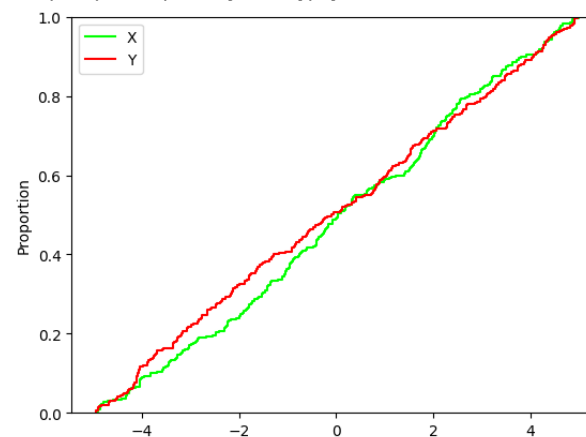
Число интервалов группировки	Выборочное значение статистики критерия	p -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
5	Inf	0.00	H_0 отклоняется	нет
10	Inf	0.00	H_0 отклоняется	нет
15	Inf	0.00	H_0 отклоняется	Нет
20	Inf	0.00	H_0 отклоняется	нет

Примечание: при расчетах использовать функции **chi2gof**, **fitdist** (**scipy.stats: histogram, chisquare**)

4. Критерий Колмогорова

Статистическая гипотеза, H_0	Выборочное значение статистики критерия	p -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
$X \sim N(m, \sigma)$	0.5	7.06	H_0 принимается	нет
$X \sim R$	0.05	0.26	H_0 принимается	нет
$X \sim \chi^2(5)$	0.56	3.87	H_0 принимается	Нет

Примечание: при расчетах использовать функции **kstest**, **lillietest**, **fitdist** (**scipy.stats: kstest**)



Эмпирическая и теоретические функции распределения

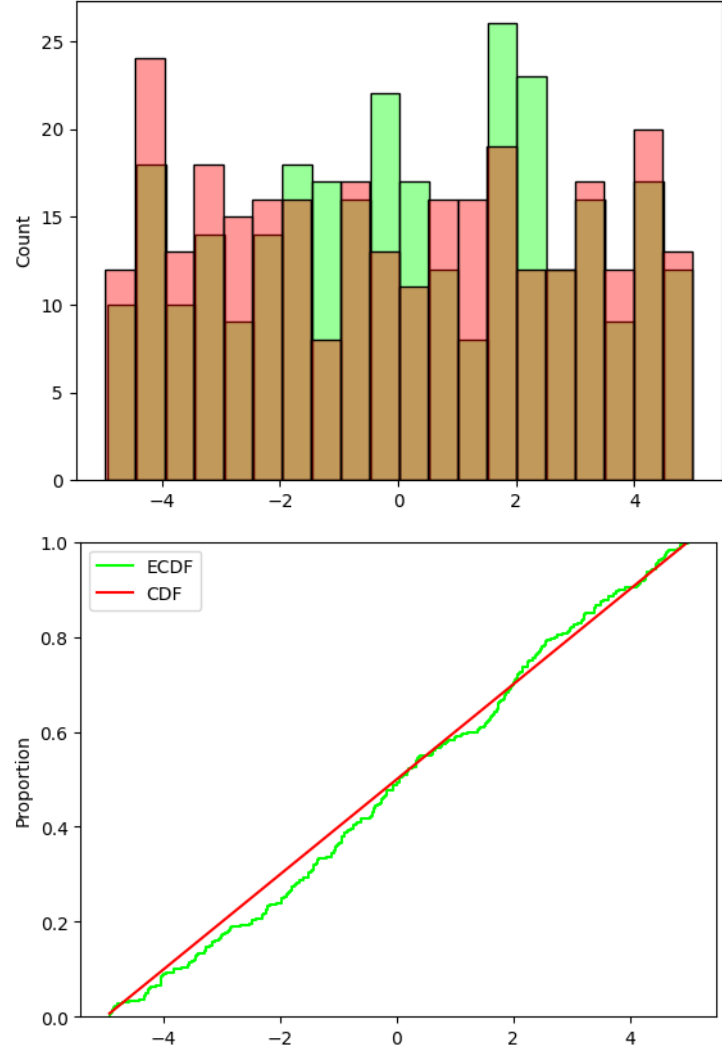
Примечание: для построения графиков использовать функции **ecdf**, **cdf** (**scipy.stats: uniform.cdf, norm.cdf, chi2.cdf**; **statsmodels.distributions.empirical_distribution: ECDF**)

5. Двухвыборочные критерии

Характеристики наблюдаемой случайной величины Y :

Распределение	Параметры	Математическое ожидание	Дисперсия
$Y \sim R(-5, 5)$	$a = -5,$ $b = 5$	$m_2 = \frac{a_2 + b_2}{2} = 0$	$\sigma_2^2 = \frac{(b_2 - a_2)^2}{12}$ $= \frac{25}{3}$

Объём выборки $n_2 = 300$



Критерий	Стат. гипотеза, H_0	Выборочное значение статистики критерия	p -value	Стат. решение при $\alpha = 0.05$	Ошибка стат. решения
Chi-squared	$F_x(x)=F_y(x)$	-0.000847	1.00	H_0 принимает ся	нет
KS-test	$F_x(x)=F_y(x)$	0.0866	0.21	H_0 принимает ся	нет

Sign test	$F_x(x)=F_y(x)$	4.00	0.686	H_0 принимает ся	нет
U-test	$F_x(x)=F_y(x)$	0.869	0.384	H_0 принимает ся	нет

Примечание: при расчетах использовать функции **chi2gof**, **kstest2**, **signest**, **ranksum** (scipy.stats: **chisquare**, **ks_2samp**; statsmodels.stats.descriptivestats: **sign_test**, **ranksums**)