

Лабораторная работа № 2

«Критерии согласия и однородности выборок»

студента Когановского Григория группы Б22-534. Дата сдачи: 29.11.2024

Ведущий преподаватель: Новиков М.А. оценка: _____ подпись: _____

Вариант №7

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
$N(-1,2)$	$m_1 = -1, \sigma_1 = 2$	$m_1 = -1$	$\sigma_1^2 = 4$

Объём выборки $n_1 = 100$

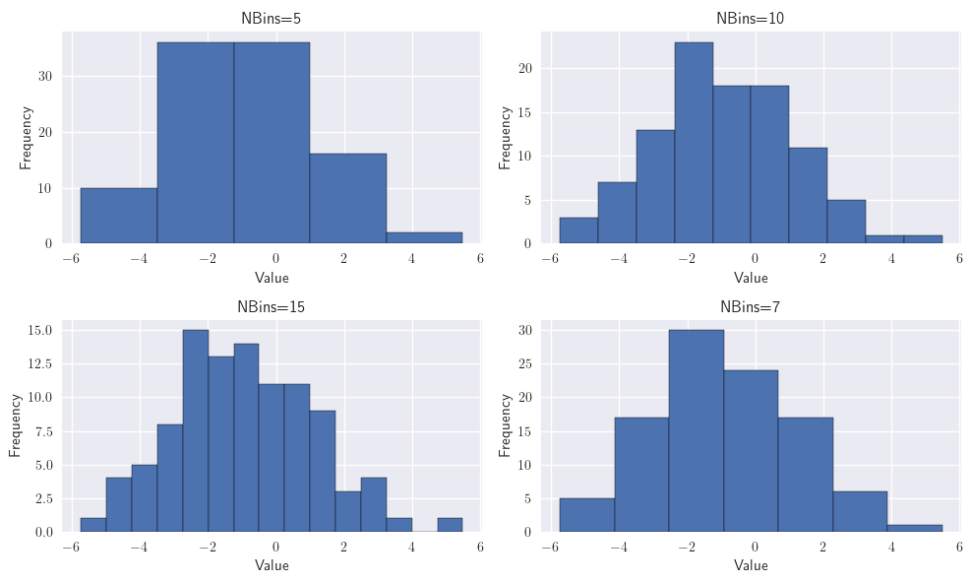
Примечание: для генерации случайных чисел использовать функции **rand**, **randn**, **chi2rnd** (**scipy.stats: uniform.rvs, norm.rvs, chi2.rvs**)

Выборочные характеристики:

Среднее, \bar{x}	Оценка дисперсии, s^2	Оценка с.к.о., s
-0.91	4.37	2.09

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	0.48	0.79	H_0 принимается	Нет
10	2.13	0.95	H_0 принимается	Нет
15	6.94	0.86	H_0 принимается	Нет
7	0.99	0.91	H_0 принимается	Нет

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	47.60	0.00	H_0 отклоняется	Нет
10	55.20	0.00	H_0 отклоняется	Нет
15	56.90	0.00	H_0 отклоняется	Нет
7	48.12	0.00	H_0 отклоняется	Нет

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

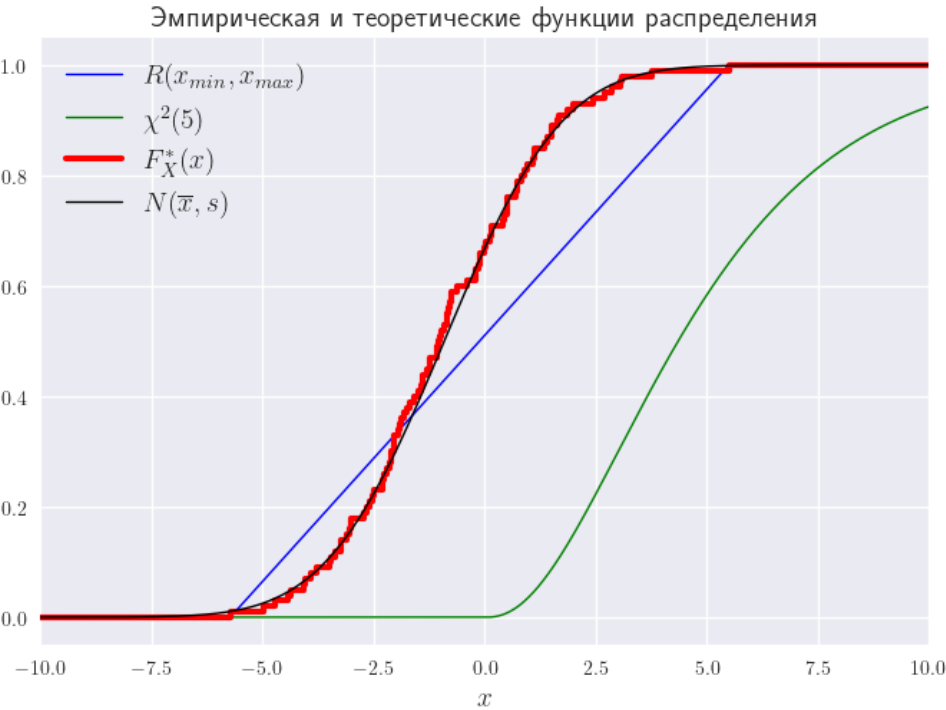
Число интервалов группировки	Выборочное значение статистики критерия	p -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
5	$+\infty$	0.00	H_0 отклоняется	Нет
10	$+\infty$	0.00	H_0 отклоняется	Нет
15	$+\infty$	0.00	H_0 отклоняется	Нет
7	$+\infty$	0.00	H_0 отклоняется	Нет

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
$X \sim N(m, \sigma)$	0.06	0.86	H_0 принимается	Нет
$X \sim R$	0.25	0.00	H_0 отклоняется	Нет
$X \sim \chi^2(5)$	0.80	0.00	H_0 отклоняется	Нет

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



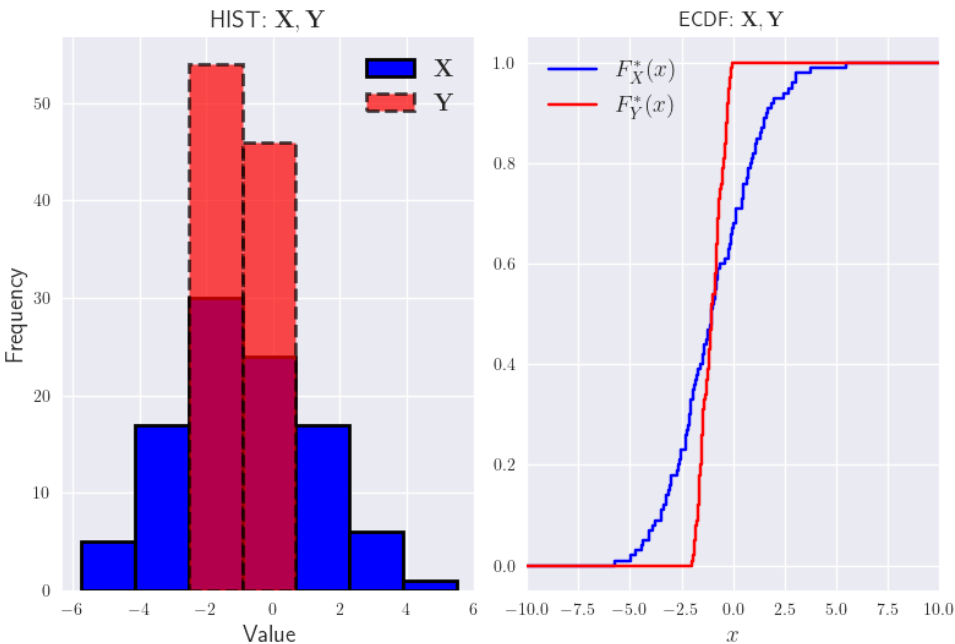
Осенний семестр 2021/2022. Лабораторный практикум по курсу «Математическая статистика»
Примечание: для построения графиков использовать функции **ecdf**, **cdf** (**scipy.stats: uniform.cdf, norm.cdf, chi2.cdf**; **statsmodels.distributions.empirical_distribution: ECDF**)

5. Двухвыборочные критерии

Характеристики наблюдаемой случайной величины Y :

Распределение	Параметры	Математическое ожидание	Дисперсия
$R(-2,0)$	$a_2 = -2, b_2 = 0$	$m_2 = \frac{a_2 + b_2}{2} = -1$	$\sigma_2^2 = \frac{(b_2 - a_2)^2}{12} = \frac{1}{3}$

Объём выборки $n_2 = 100$



Критерий	Статистическая гипотеза, H_0	Выборочное значение статистики критерия	p -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
Chi-squared	$F_X(x) = F_Y(x)$	59.77	0.00	H_0 отклоняется	Нет
KS-test	$F_X(x) = F_Y(x)$	0.34	0.00	H_0 отклоняется	Нет
Sign test	$F_X(x) = F_Y(x)$	1	0.92	H_0 принимается	Да
U-test	$F_X(x) = F_Y(x)$	4989	0.98	H_0 принимается	Да

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