

Лабораторная работа № 2
«Критерии согласия и однородности выборок»

студента Розинко Е.Д. группы Б21-524. Дата сдачи: 16.11.2023
Ведущий преподаватель: Трофимов А.Г. оценка: _____ подпись: _____

Вариант № 4

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
$R(5, 15)$	$R(a, b)$	10	8.3

Объём выборки $n_1 = 300$

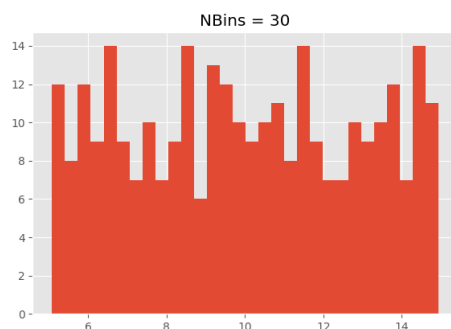
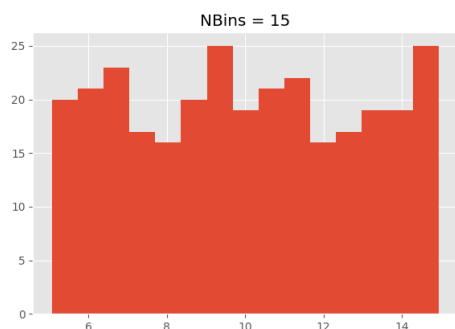
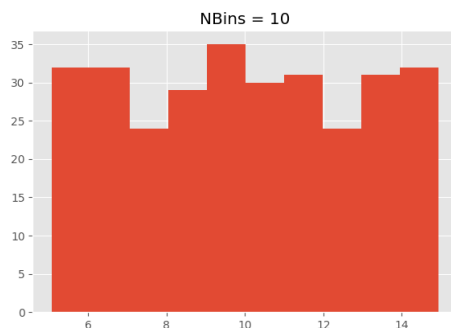
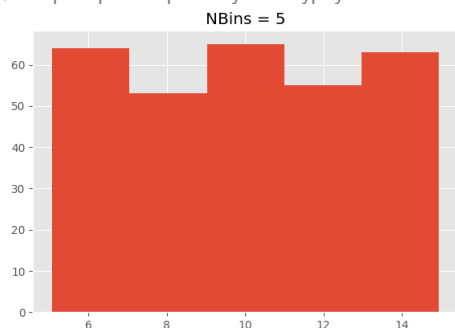
Примечание: для генерации случайных чисел использовать функции **rand**, **randn**, **chi2rnd** (**scipy.stats: uniform.rvs, norm.rvs, chi2.rvs**)

Выборочные характеристики:

Среднее, \bar{x}	Оценка дисперсии, s^2	Оценка с.к.о., s
10.009	8.278	2.877

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	68.071	0	H_0 отвергается	Нет
10	81.177	0	H_0 отвергается	Нет
15	90.323	0	H_0 отвергается	Нет
30	89.817	0	H_0 отвергается	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	2.162	0.679	H_0 принимается	Нет

10	3.853	0.407	H_0 принимается	Нет
15	6.052	0.173	H_0 принимается	Нет
30	16.910	0.133	H_0 принимается	нет

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

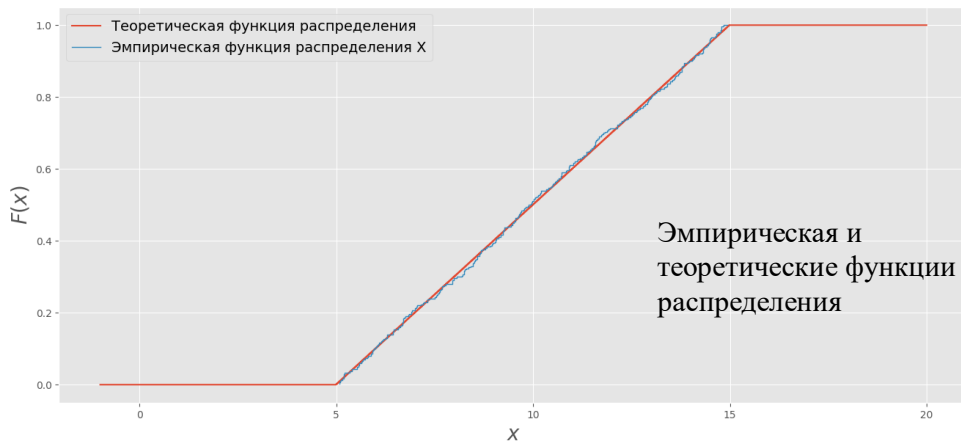
Число интервалов группировки	Выборочное значение статистики критерия	p -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
5	60.079	0	H_0 отвергается	Нет
10	71.553	0	H_0 отвергается	Нет
15	95.824	0	H_0 отвергается	Нет
30	95.824	0	H_0 отвергается	нет

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
$X \sim N(m, \sigma)$	0.066	0.136	H_0 принимается	нет
$X \sim R$	0.022	0.997	H_0 принимается	нет
$X \sim \chi^2(5)$	0.135	0.000	H_0 отвергается	нет

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



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

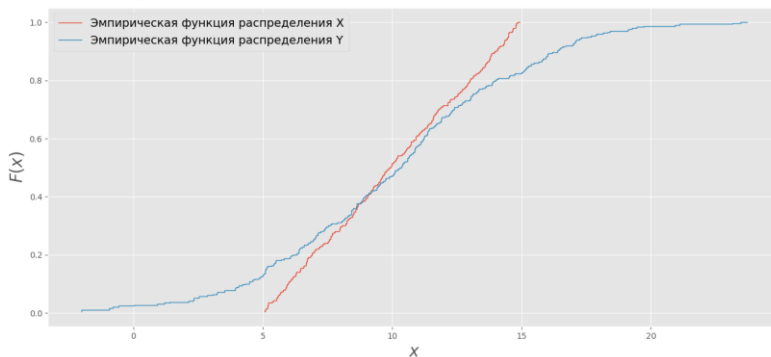
5. Двухвыборочные критерии

Характеристики наблюдаемой случайной величины Y :

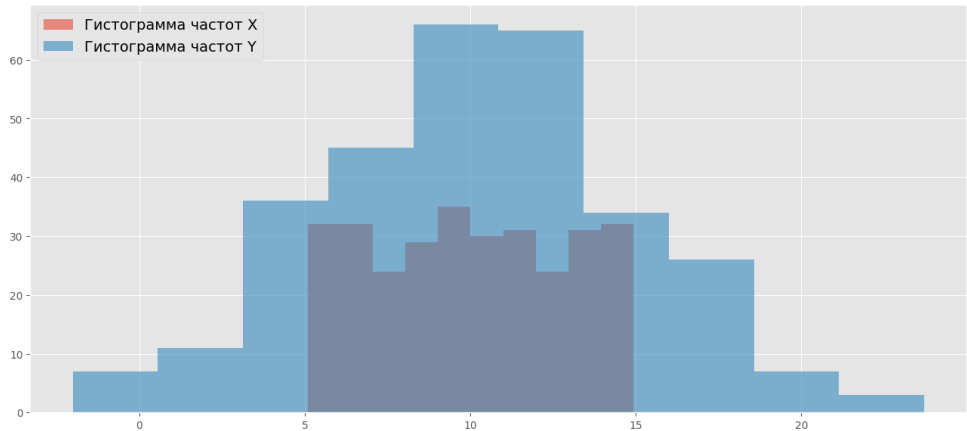
Распределение	Параметры	Математическое ожидание	Дисперсия
$N(10, 5)$	$N(m, \sigma)$	10	25

Объём выборки $n_2 = 100$

ECDF: X, Y



HIST: X, Y



Критерий	Стат. гипотеза, H_0	Выборочное значение статистики критерия	p -value	Стат. решение при $\alpha = 0.05$	Ошибка стат. решения
Chi-squared	$FX(\xi) = FY(\xi)$	535.510	0	H_0 отвергается	нет
KS-test	$FX(\xi) = FY(\xi)$	0.180	0	H_0 отвергается	Нет
Sign test	$FX(\xi) = FY(\xi)$	0.951	0.149	H_0 принимается	Нет
U-test	$FX(\xi) = FY(\xi)$	-0.490	0.624	H_0 принимается	нет

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