

## Лабораторная работа № 2

### «Критерии согласия и однородности выборок»

студента Моисеенко Олеся группы Б20-514. Дата сдачи: 08.11.22

Ведущий преподаватель: Сорока А.А. оценка: \_\_\_\_\_ подпись: \_\_\_\_\_

#### Вариант №15

*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, $\sigma^2$
$\chi^2(15)$	$k = 15$	15	30

Объём выборки  $n_1 = 100$

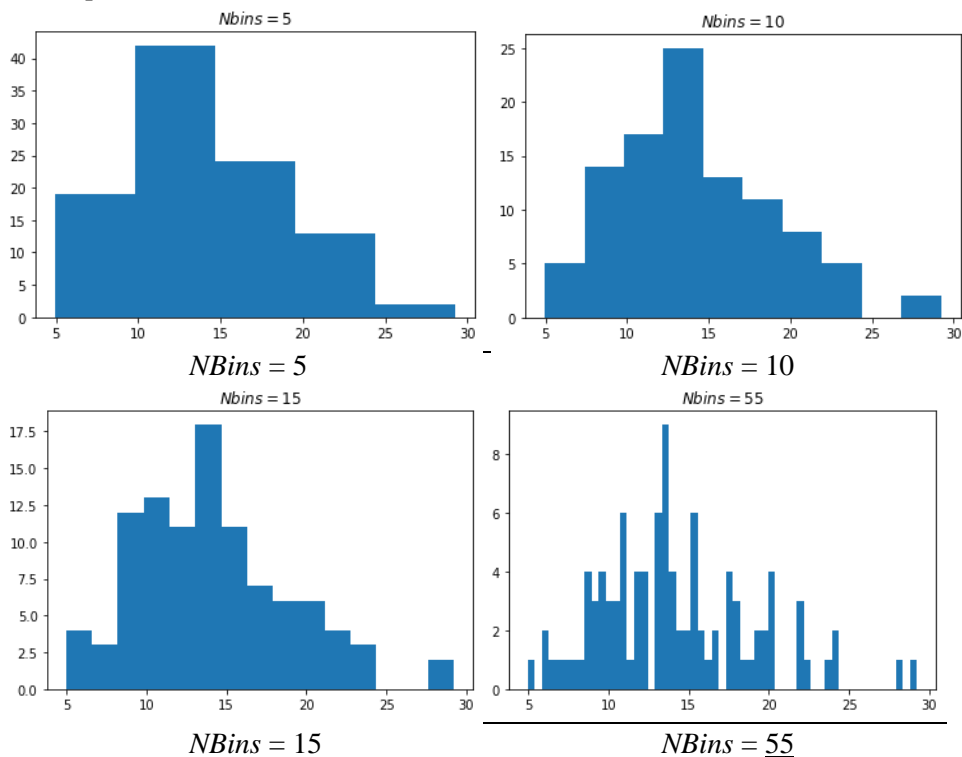
*Примечание:* для генерации случайных чисел использовать функции **rand**, **randn**, **chi2rnd** (scipy.stats: **uniform.rvs**, **norm.rvs**, **chi2.rvs**)

Выборочные характеристики:

Среднее, $\bar{x}$	Оценка дисперсии, $s^2$	Оценка с.к.о., $s$
14.150636	23.027153	4.798662

## 2. Визуальное представление выборки

Гистограммы частот:



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

3. Критерий хи-квадрата) Статистическая гипотеза:  $H_0: X \sim N(m, \sigma)$ 

Число интервалов группировки	Выборочное значение статистики критерия	$p\text{-value}$	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
5	3.847	0.292	$H_0$ принимается	2-го рода
10	14.719	0.080	$H_0$ принимается	2-го рода
15	30.407	0.005	$H_0$ отклоняется	нет
55	101.515	0.000	$H_0$ отклоняется	нет

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

Число интервалов группировки	Выборочное значение статистики критерия	$p\text{-value}$	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
5	195.281	0.000	$H_0$ отклоняется	нет
10	195.174	0.000	$H_0$ отклоняется	нет
15	213.148	0.000	$H_0$ отклоняется	нет
55	207.699	0.000	$H_0$ отклоняется	нет

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

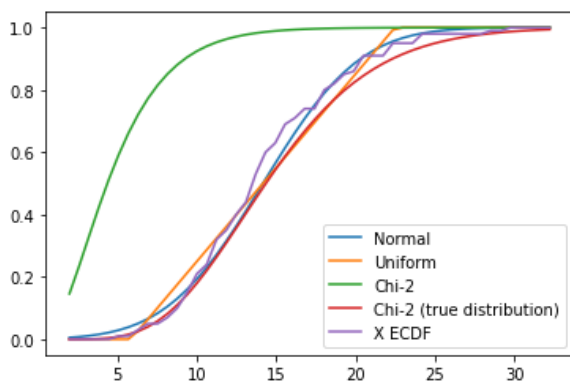
Число интервалов группировки	Выборочное значение статистики критерия	<i>p-value</i>	Статистическое решение при $\alpha = 0.05$ _____	Ошибка стат. решения
5	666.901	0.000	$H_0$ отклоняется	нет
10	674.787	0.000	$H_0$ отклоняется	нет
15	680.228	0.000	$H_0$ отклоняется	нет
55	615.754	0.000	$H_0$ отклоняется	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
$X \sim N(m, \sigma)$	0.09349420356012017	0.3258728427990727	$H_0$ принимается	2-го рода
$X \sim R$	0.11208030185253381	0.15016672700506925	$H_0$ принимается	2-го рода
$X \sim \chi^2(5)$	0.7941438588590355	4.903529972331366e-68	$H_0$ отклоняется	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**)

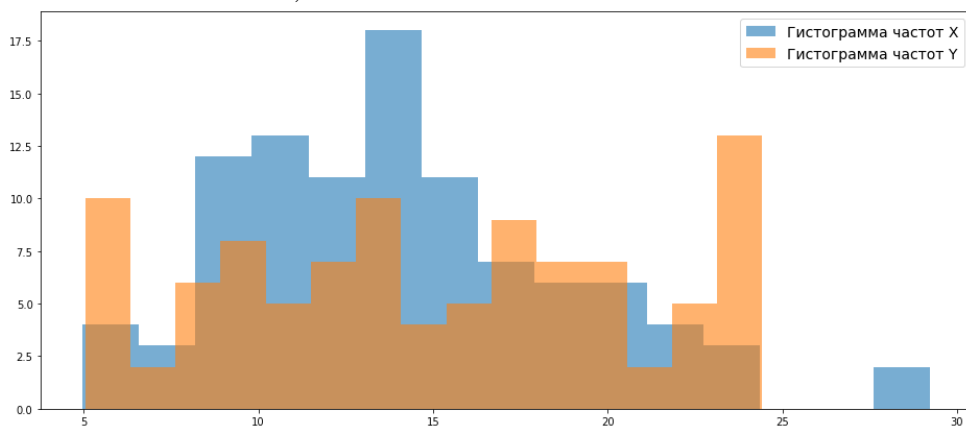
## 5. Двухвыборочные критерии

Характеристики наблюдаемой случайной величины  $Y$ :

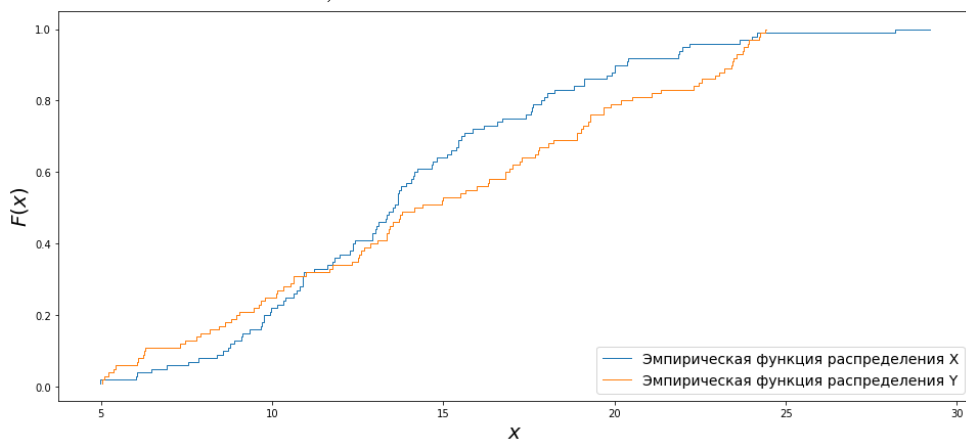
Распределение	Параметры	Математическое ожидание	Дисперсия
$R(5, 25)$	$a = 5$ $b = 25$	15	$100/3$

Объём выборки  $n_2 = 100$

HIST: X, Y



ECDF: X, Y



Критерий	Стат. гипотеза, $H_0$	Выборочное значение статистики критерия	$p$ -value	Стат. решение при $\alpha = 0.05$	Ошибка стат. решения
Chi-squared	$F_x(\xi)=F_y(\xi)$	68.729	0.000	$H_0$ отклоняется	нет
KS-test	$F_x(\xi)=F_y(\xi)$	0.17	0.1111952 605382919 2	$H_0$ принимается	2-го рода
Sign test	$F_x(\xi)=F_y(\xi)$	-7.000	0.193	$H_0$ принимается	2-го рода
U-test	$F_x(\xi)=F_y(\xi)$	4551.0	0.2731400 830994096	$H_0$ принимается	2-го рода

*Примечание:* при расчетах использовать функции **chi2gof**, **kstest2**, **signtest**, **ranksum** (**scipy.stats.chisquare**, **ks\_2samp**; **statsmodels.stats.descriptivestats.sign\_test**, **ranksums**)