**Лабораторная работа № 2**

**«Критерии согласия и однородности выборок»**

студента Васильева Феодосия группы Б21-524 . Дата сдачи: 18.11.2023

Ведущий преподаватель: Трофимов А.Г. оценка: подпись:\_\_\_\_\_\_\_

Вариант № 6

*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, σ2 |
|  |  | 2 | 4 |

Объём выборки *n*1 = 150

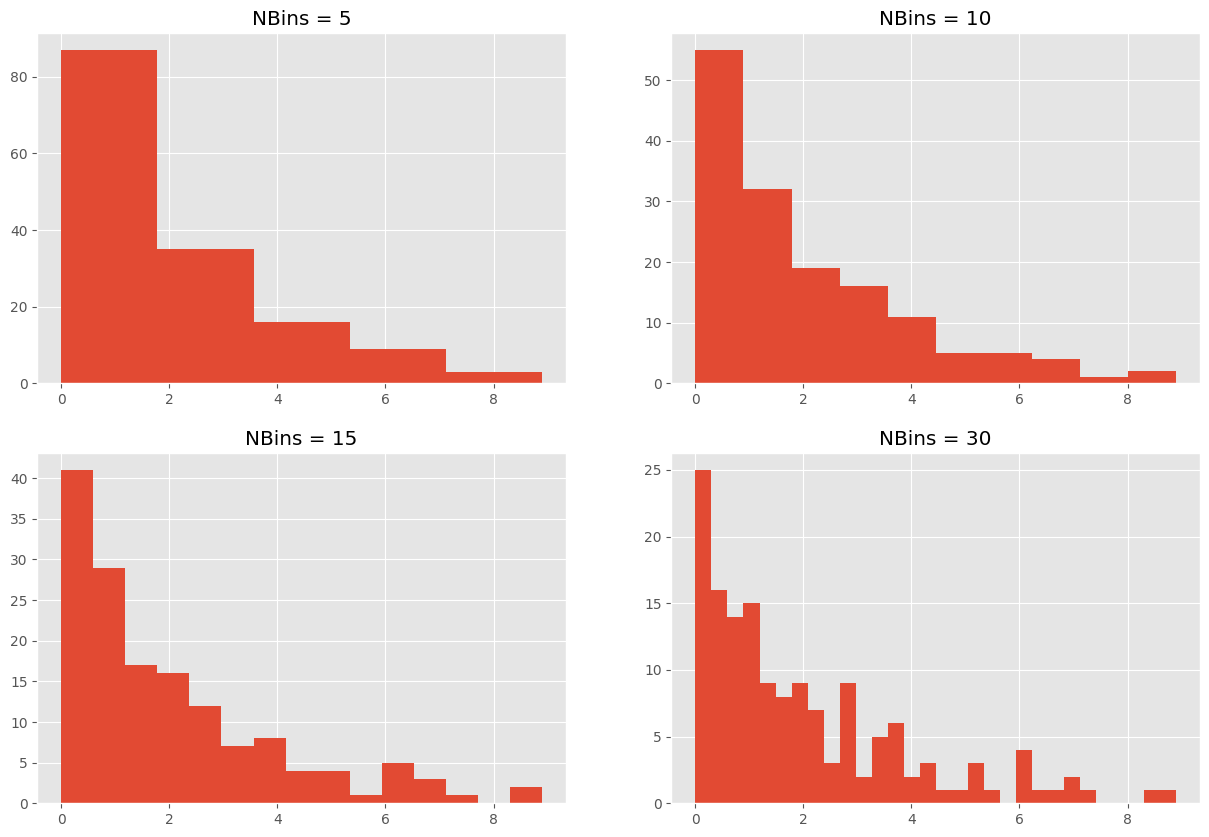
*Примечание*: для генерации случайных чисел использовать функции **rand**, **randn, chi2rnd** **(scipy.stats: uniform.rvs, norm.rvs, chi2.rvs)**

Выборочные характеристики:

|  |  |  |
| --- | --- | --- |
| Среднее, | Оценка дисперсии, s2 | Оценка с.к.о., s |
| 2.011 | 3.744 | 1.935 |

2. Визуальное представление выборки

Гистограммы частот:



3. Критерий хи-квадрат

*а*) Статистическая гипотеза: *H*0: *X* ~ *N*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Число интервалов группировки | Выборочное значение статистики критерия | *p-value* | Статистическое решение при  α = \_\_0.05 \_\_\_\_ | Ошибка стат. решения |
| 5 | 67.742 | 0 |  | нет |
| 10 | 80.469 | 0 |  | нет |
| 15 | 92.952 | 0 |  | нет |
| 30 | 107.859 | 0 |  | нет |

*б*) Статистическая гипотеза: *H*0: *X* ~ *R*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Число интервалов группировки | Выборочное значение статистики критерия | *p-value* | Статистическое решение при  α = \_0.05\_\_\_\_\_ | Ошибка стат. решения |
| 5 | 270.484 | 0 |  | нет |
| 10 | 305.402 | 0 |  | нет |
| 15 | 333.389 | 0 |  | нет |
| 30 | 335.946 | 0 |  | нет |

*в*) Статистическая гипотеза: *H*0: *X* ~ χ2(5)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Число интервалов группировки | Выборочное значение статистики критерия | *p-value* | Статистическое решение при  α = \_0.05\_\_\_\_\_ | Ошибка стат. решения |
| 5 | 1.440 | 0.974 |  | нет |
| 10 | 2.699 | 0.507 |  | нет |
| 15 | 3.054 | 0.240 |  | нет |
| 30 | 9.722 | 0.719 |  | нет |

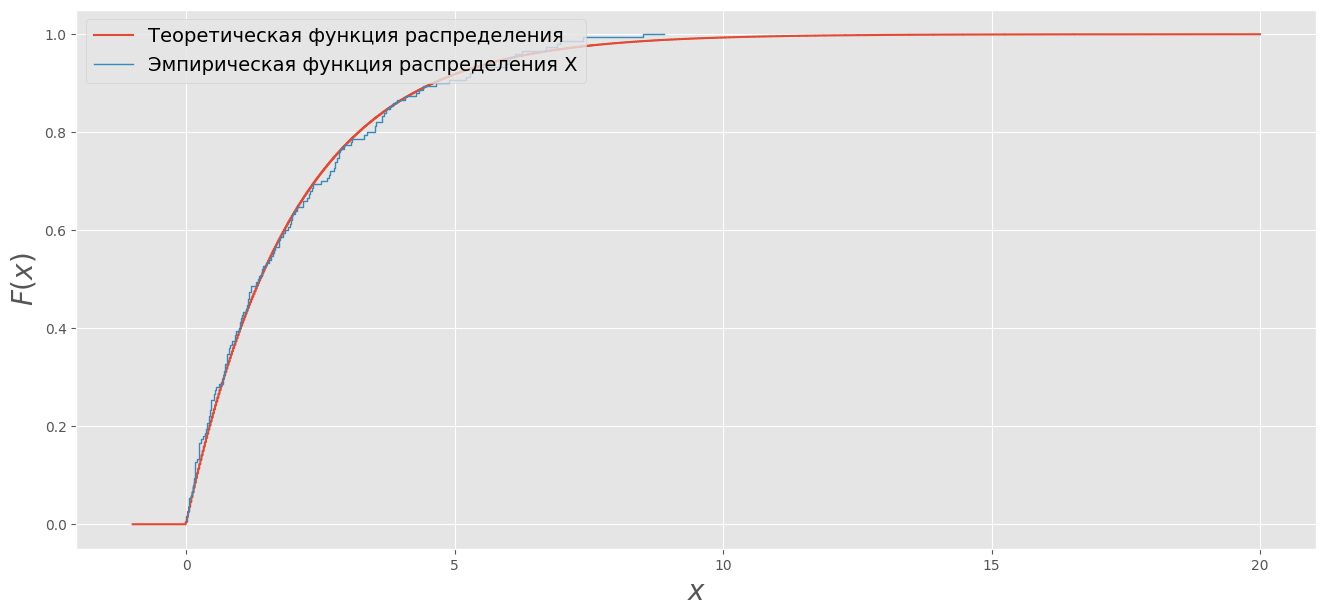
*Примечание*: при расчетах использовать функции **chi2gof**, **fitdist (scipy.stats: histogram, chisquare)**

4. Критерий Колмогорова

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Статистическая гипотеза, *H*0 | Выборочное значение статистики критерия | *p-value* | Статистическое решение при  α = \_0.05\_\_\_\_\_ | Ошибка стат. решения |
| *X* ~ *N*(*m*, σ) | 0.150 | 0.002 |  | нет |
| *X* ~ *R* | 0.201 | 0 |  | нет |
| *X* ~ χ2(5) | 0.483 | 0 |  | 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)**



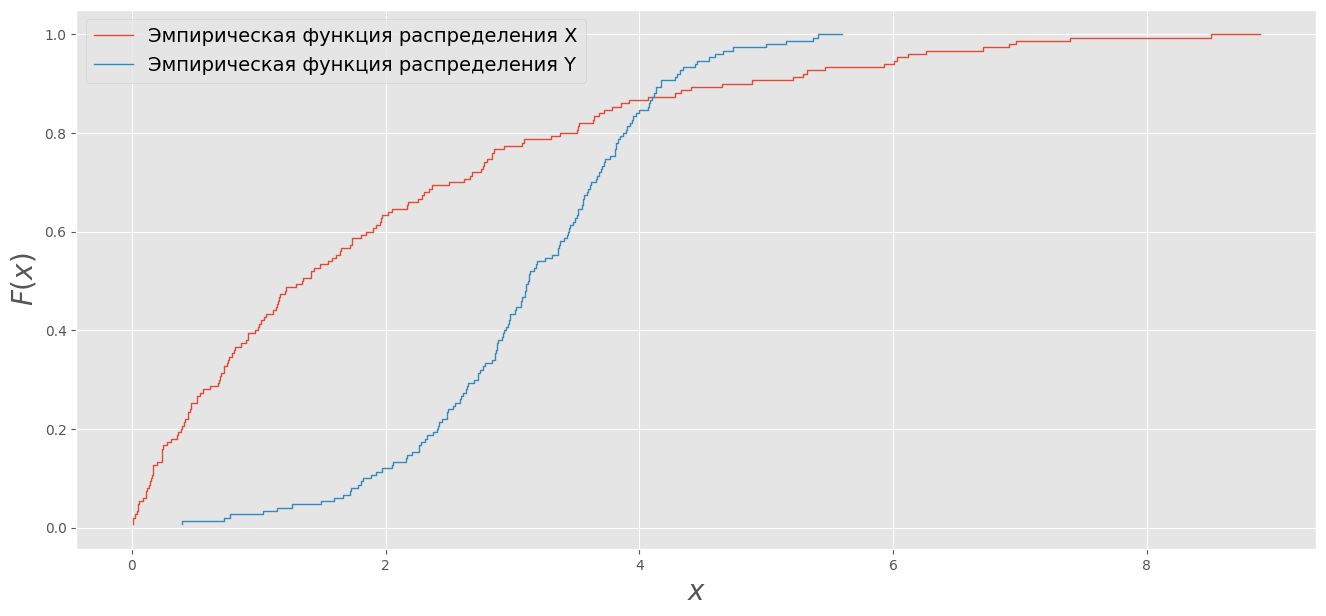
5. Двухвыборочные критерии

Характеристики наблюдаемой случайной величины *Y*:

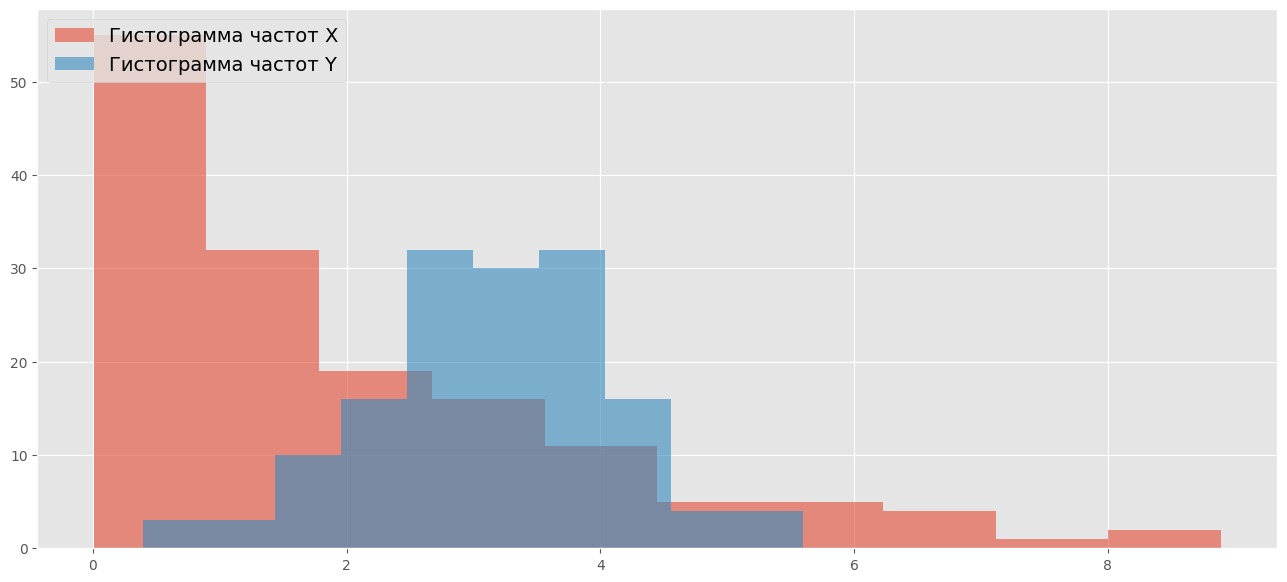
|  |  |  |  |
| --- | --- | --- | --- |
| Распределение | Параметры | Математическое ожидание | Дисперсия |
|  |  | 3 |  |

Объём выборки *n*2 = 150

ECDF: X, Y



HIST: X, Y



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Критерий | Стат. гипотеза, *H*0 | Выборочное значение статистики критерия | *p-value* | Стат. решение при  α = 0.05 | Ошибка стат. решения |
| Chi-squared |  | 756.545 | 0 |  | 1 рода |
| KS-test |  | 0.513 | 0 |  | нет |
| Sign test |  | -33.000 | 0 |  | нет |
| U-test |  | -7.634 | 0 |  | нет |

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