**Лабораторная работа № 2**

**«Критерии согласия и однородности выборок»**

студента Кевролетина Алексея группы Б19-504 . Дата сдачи: 9.11.2021

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

Вариант № 10

*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 |
| --- | --- | --- | --- |
|  |  |  |  |

Объём выборки *n*1 = 300

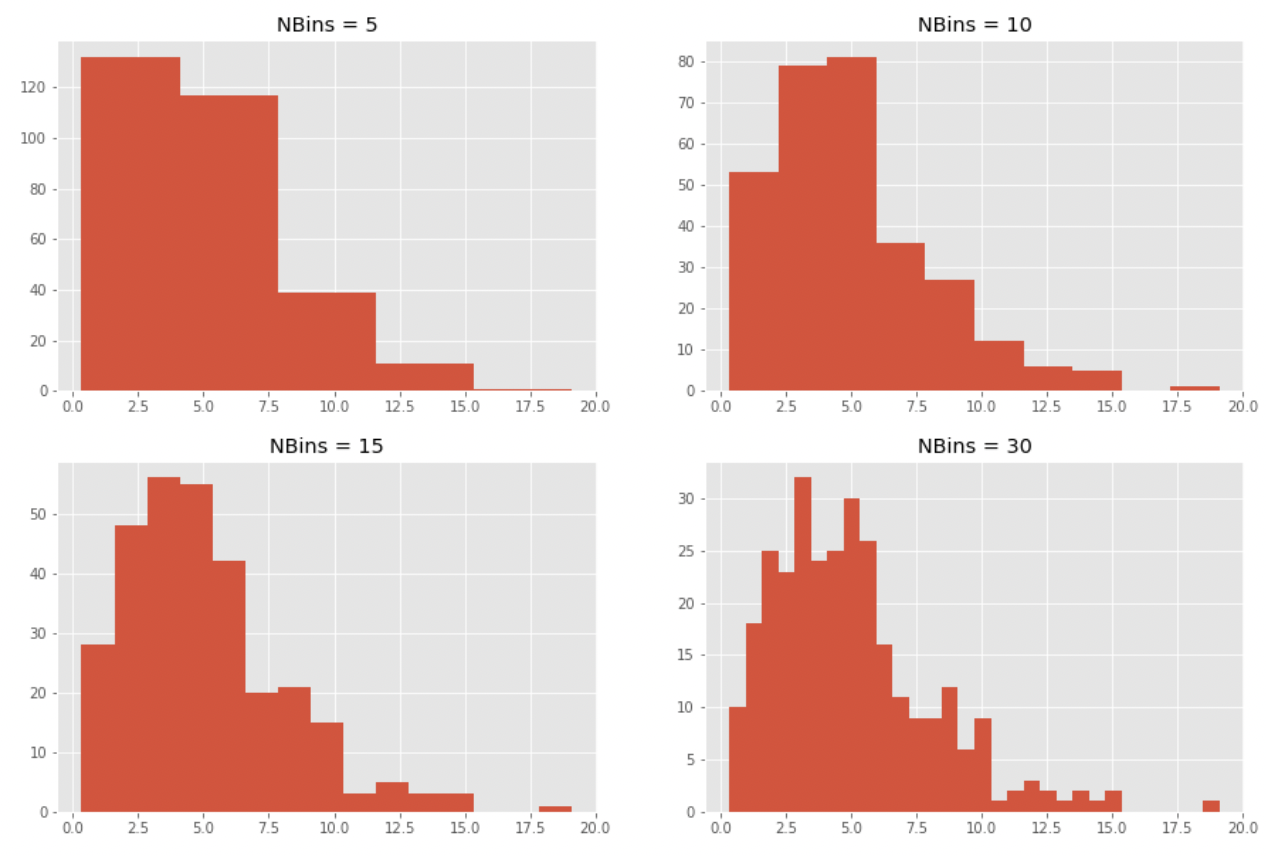
*Примечание*: для генерации случайных чисел использовать функции **rand**, **randn, chi2rnd** **(scipy.stats: uniform.rvs, norm.rvs, chi2.rvs)**

Выборочные характеристики:

| Среднее, | Оценка дисперсии, s2 | Оценка с.к.о., s |
| --- | --- | --- |
| 4.993 | 9.393 | 3.065 |

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 | 30.977 | 0 |  | нет |
| 10 | 28.457 | 0 |  | нет |
| 15 | 25.753 | 0 |  | нет |
| 30 | 35.719 | 0.002 |  | 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 | 403.691 | 0 |  | нет |
| 10 | 461.232 | 0 |  | нет |
| 15 | 469.577 | 0 |  | нет |
| 30 | 484.099 | 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 | 0.632 | 0.542 |  | нет |
| 10 | 4.515 | 0.785 |  | нет |
| 15 | 7.882 | 0.890 |  | нет |
| 30 | 15.261 | 0.847 |  | нет |

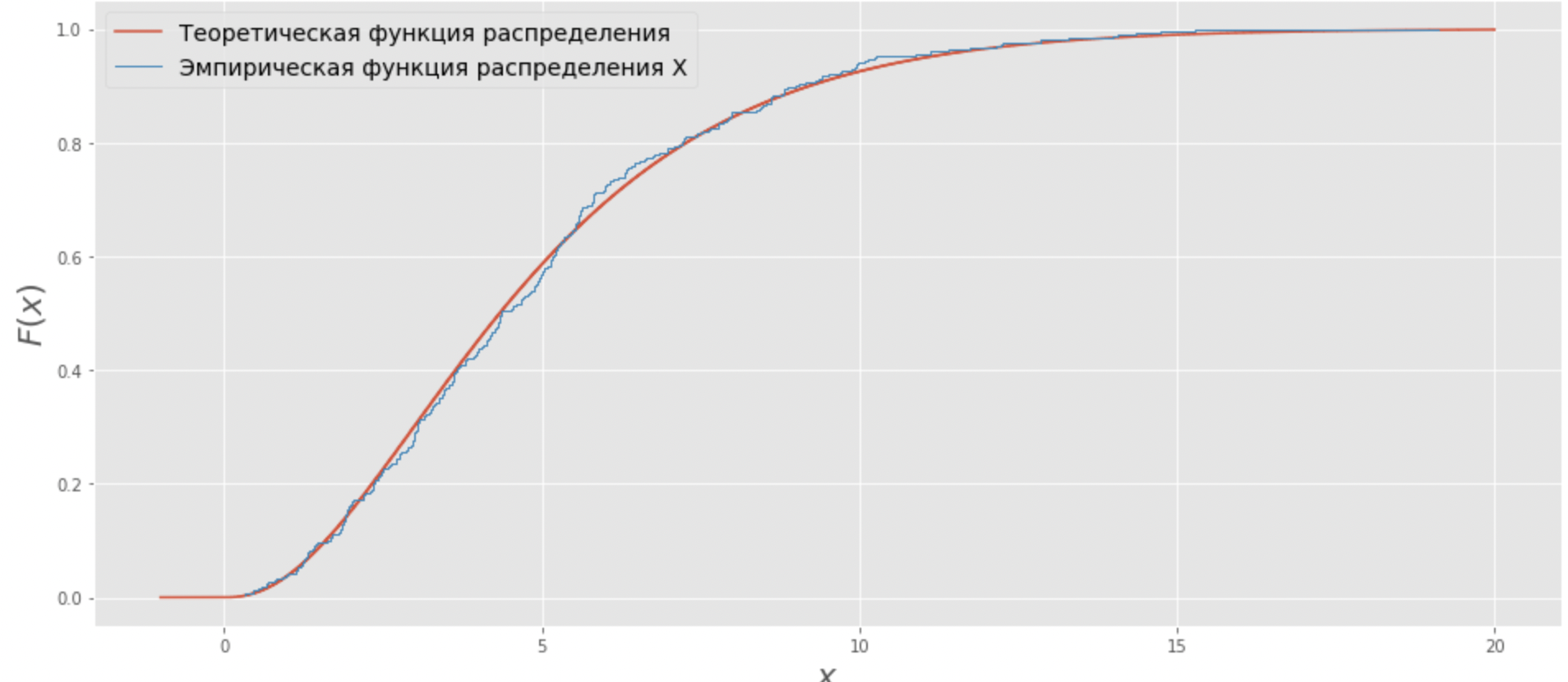
*Примечание*: при расчетах использовать функции **chi2gof**, **fitdist (scipy.stats: histogram, chisquare)**

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) | 0.102 | 0.004 |  | нет |
| *X* ~ *R* | 0.131 | 0.000 |  | нет |
| *X* ~ χ2(5) | 0.033 | 0.905 |  | 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)**



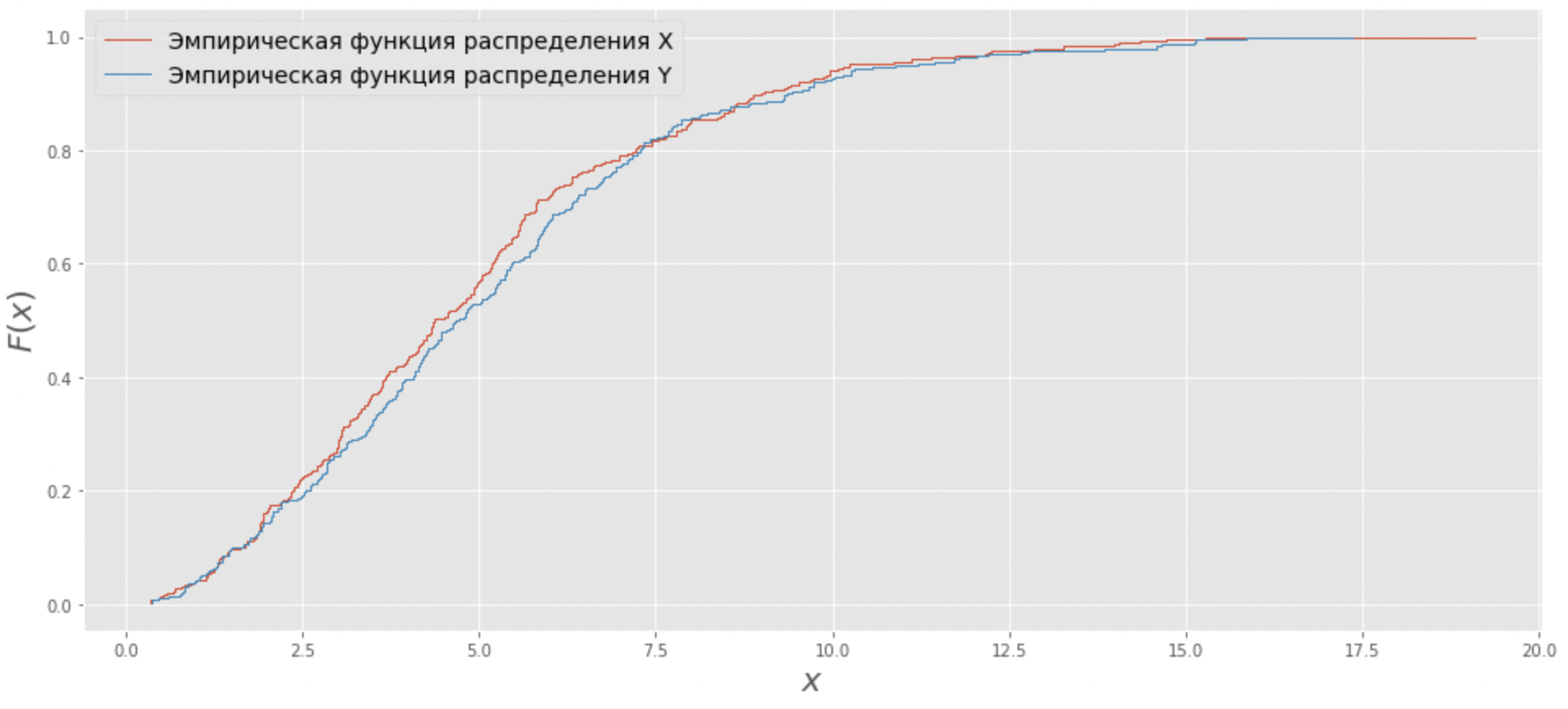
5. Двухвыборочные критерии

Характеристики наблюдаемой случайной величины *Y*:

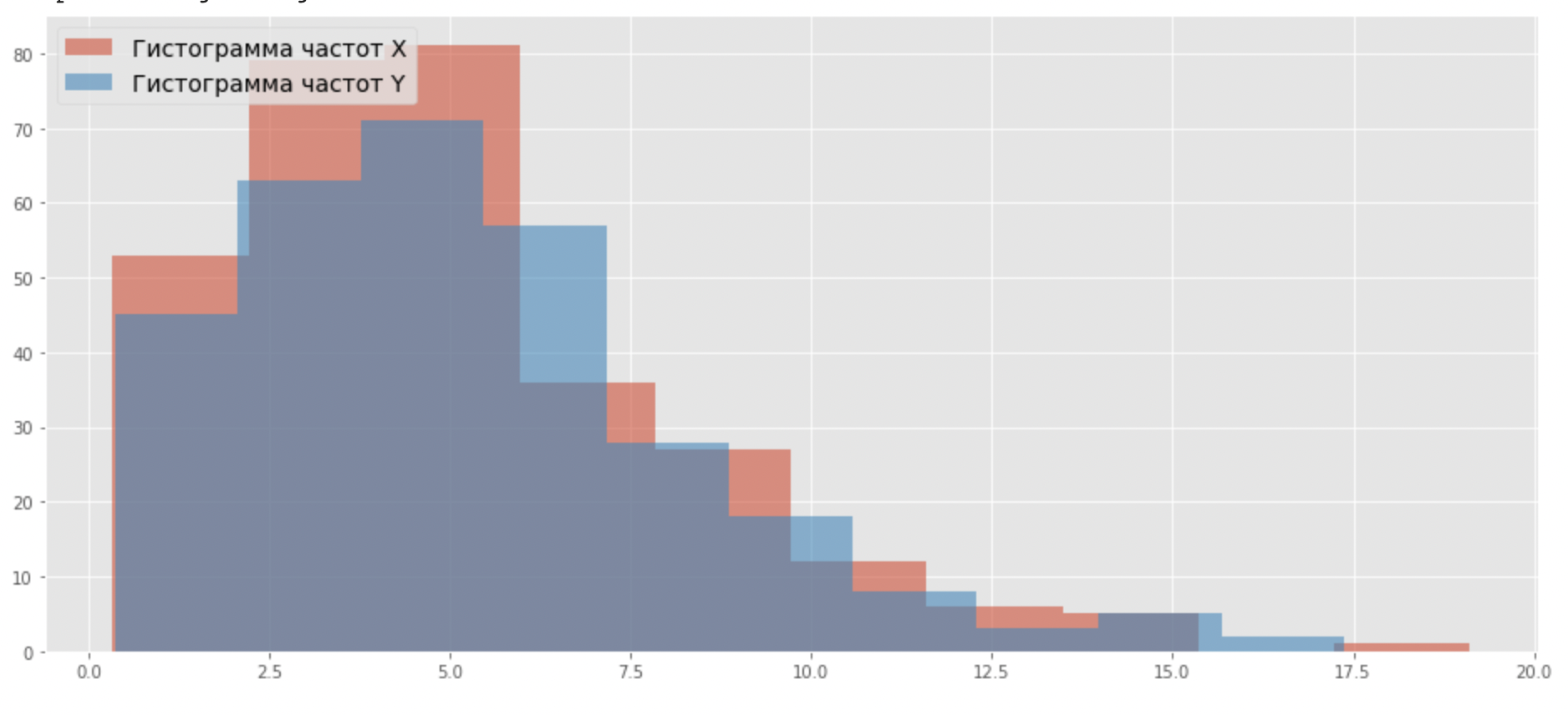
| Распределение | Параметры | Математическое ожидание | Дисперсия |
| --- | --- | --- | --- |
|  |  |  |  |

Объём выборки *n*2 = 300

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 |  | 24.000 | 0.004 |  | 1 рода |
| KS-test |  | 0.077 | 0.342 |  | нет |
| Sign test |  | -6.000 | 0.525 |  | нет |
| U-test |  | 21312.0 | 0.400 |  | нет |

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