

### Лабораторная работа № 3 «Однофакторный дисперсионный анализ»

студента Смирнова Даниила группы Б18-501. Дата сдачи: 23.11.2020

Ведущий преподаватель: Трофимов А.Г. оценка:

подпись: \_\_\_\_\_

#### Вариант № 6

*Цель работы:* изучение функций Statistics and Machine Learning Toolbox™ MATLAB / Python SciPy.stats для проведения однофакторного дисперсионного анализа (*One-Way ANOVA*).

#### 1. Исходные данные

Характеристики наблюдаемых случайных величин:

СВ	Распределение	Параметры	Математическое ожидание, $m_i$	Дисперсия, $\sigma_i^2$	Объем выборки, $n_i$
$X_1$	chi2	2	2	4	100
$X_2$	N	3,1	3	1	150
$X_3$	R	-2,2	0	4/3	200
$X_4$	N	5,1	5	1	100

Количество случайных величин  $k = 4$

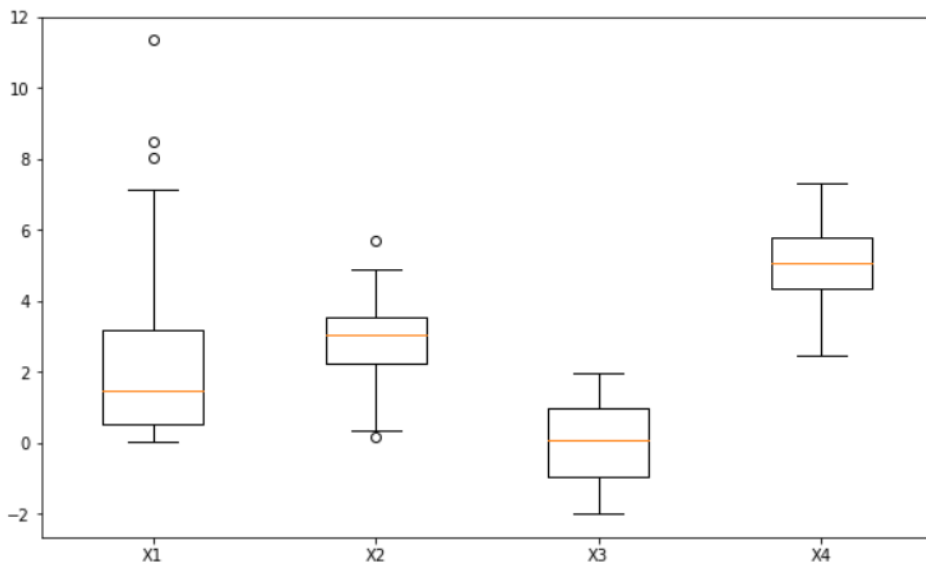
*Примечание:* для генерации случайных чисел использовать функции **rand**, **randn**, **chi2rnd** (scipy.stats: **uniform.rvs**, **norm.rvs**, **chi2.rvs**)

Выборочные характеристики:

СВ	Среднее, $\bar{x}_i$	Оценка дисперсии, $s_i^2$	Оценка с.к.о., $s_i$
$X_1$	2.15	4.43	2.1
$X_2$	2.87	1.03	1,01
$X_3$	0.06	1.32	1.15
$X_4$	5.06	1.08	1.04
<i>Pooled</i>	2.12	5.04	2.24

## 2. Визуальное представление выборок

Диаграммы *Box-and-Whisker*:



*Примечание:* для построения диаграмм использовать функции **boxplot**, **vartestn** (**matplotlib.pyplot.boxplot**)

## 3. Проверка условия применимости дисперсионного анализа

Статистическая гипотеза:  $H_0 : \sigma_1^2 = \dots = \sigma_k^2$

Критерий Бартлетта:

Выборочное значение статистики критерия	<i>p-value</i>	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
94.47	2.3898230345310784e-20	Гипотеза не верна	Гипотеза $H_0$ отклоняется (Решение верно)

*Примечание:* для проверки гипотезы использовать функцию **vartestn** (**scipy.stats.bartlett**)

#### 4. Однофакторный дисперсионный анализ

Таблица дисперсионного анализа:

Источник вариации	Показатель вариации	Число степеней свободы	Несмещённая оценка
Группировочный признак	3.27	3	4.36
Остаточные признаки	1.76	546	1,77
Все признаки	5.03	549	5.03

Эмпирический коэффициент детерминации  $\eta^2 = 0.65$

Эмпирическое корреляционное отношение  $\eta = 0.8$

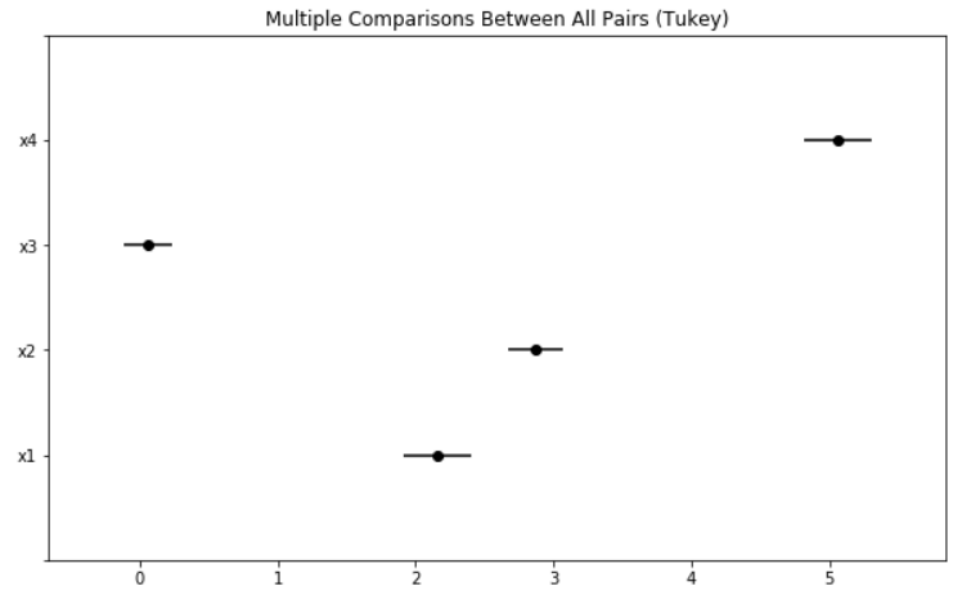
Статистическая гипотеза:  $H_0 : m_1 = \dots = m_k$

Выборочное значение статистики критерия	<i>p-value</i>	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
338.27	4.416596093257432e-124	Гипотеза не верна	Гипотеза Н0 отклоняется (Решение верно)

**Примечание:** при расчетах использовать функцию **anova1** (**scipy.stats.f\_oneway**)

5. Метод линейных контрастов

Доверительные интервалы для  $m_1, \dots, m_k$ :



Попарные сравнения  $m_i$  и  $m_j$ :

Гипотеза	Выборочное значение статистики критерия	<i>p-value</i>	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
$H_0: m_1 = m_2$	0.7112	0.001	Гипотеза не верна	Гипотеза Н0 отклоняется (Решение верно)
$H_0: m_1 = m_3$	-2.0961	0.001	Гипотеза не верна	Гипотеза Н0 отклоняется (Решение верно)
$H_0: m_1 = m_4$	2.9105	0.001	Гипотеза не верна	Гипотеза Н0 отклоняется

				(Решение верно)
$H_0: m_2 = m_3$	-2.8073	0.001	Гипотеза не верна	Гипотеза $H_0$ отклоняется (Решение верно)
$H_0: m_2 = m_4$	2.1993	0.001	Гипотеза не верна	Гипотеза $H_0$ отклоняется (Решение верно)
$H_0: m_3 = m_4$	5.0066	0.001	Гипотеза не верна	Гипотеза $H_0$ отклоняется (Решение верно)

*Примечание:* при расчетах использовать функцию **multcompare**  
(**statsmodels.stats.multicomp.pairwise\_tukeyhsd**)