

## Лабораторная работа № 3

### «Однофакторный дисперсионный анализ»

студента Моисеенко Олеся группы Б20-514. Дата сдачи: \_\_\_\_\_

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

#### Вариант №15

*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, $n_i$
$X_1$	$\chi^2(15)$	$k = 15$	15	30	50
$X_2$	$R(5, 25)$	$a = 5$ $b = 25$	15	$33,333333 = 100/3$	100
$X_3$	$N(5, 2)$	$m = 5$ $\sigma^2 = 2$	5	2	100

Количество случайных величин  $k = \underline{3}$

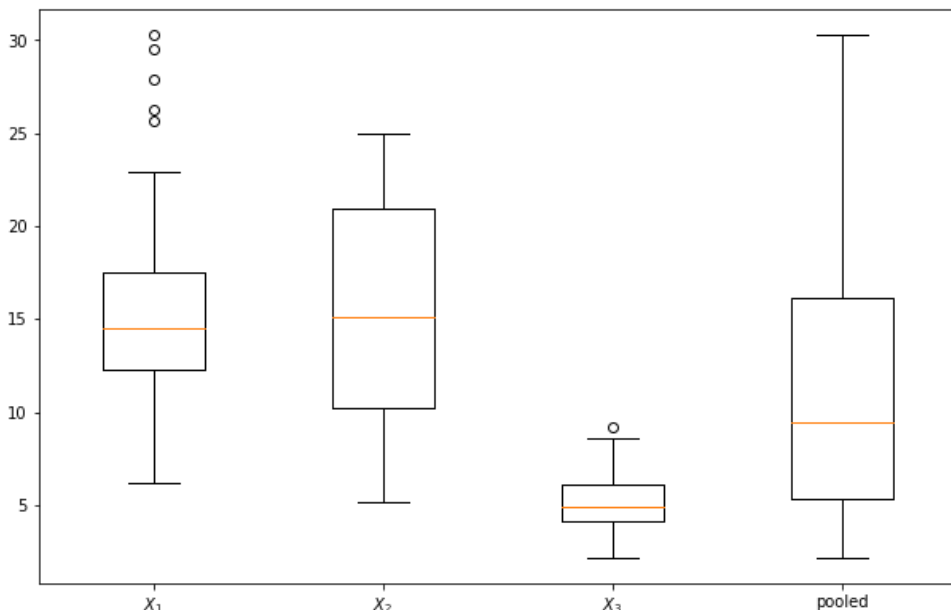
*Примечание:* для генерации случайных чисел использовать функции **rand**, **randn**, **chi2rnd** (scipy.stats: **uniform.rvs**, **norm.rvs**, **chi2.rvs**)

Выборочные характеристики:

СВ	Среднее, $\bar{x}_i$	Оценка дисперсии, $s_i^2$	Оценка с.к.о., $s_i$
$X_1$	15.601281	31.689094	5.629307
$X_2$	15.425819	35.128094	5.926896
$X_3$	5.076484	1.970757	1.403837
<i>Pooled</i>	11.321177	47.178926	6.868692

## 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	$p$ -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
164.22850946678074	2.1788748384198263e-36	$H_0$ отклоняется	нет

*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	$D_b^* = 26.001567114822492$	$K - 1 = 2$	$n * D_b^* / (K - 1) = 3250.1958893528117$
Остаточные признаки	$D_w^* = 21.177359296548417$	$n - K = 247$	$n * D_w^* / (n - K) = 21.434574186789895$
Все признаки	$D_x^* = 47.17892641137091$	$n - 1 = 249$	$n * D_x^* / (n - 1) = 47.368400011416576$

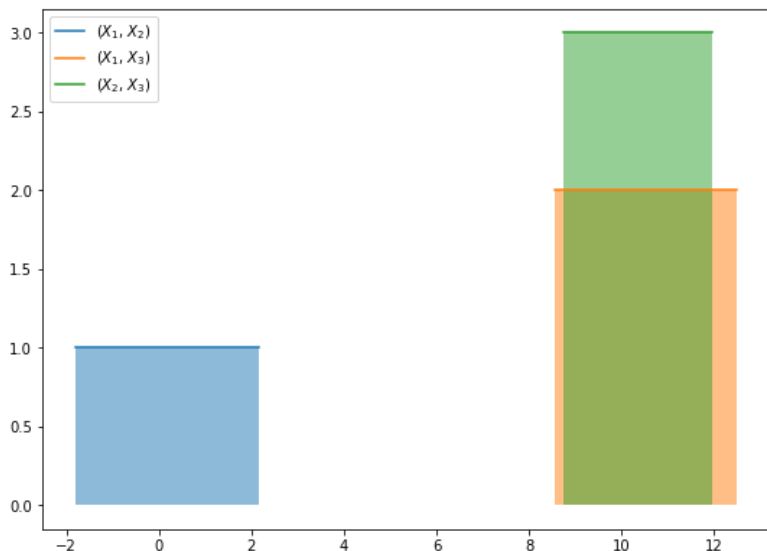
Эмпирический коэффициент детерминации  $\eta^2 = \underline{0.5511267231497595}$ Эмпирическое корреляционное отношение  $\eta = \underline{0.7423790966546402}$ Статистическая гипотеза:  $H_0 : m_1 = \dots = m_k$ 

Выборочное значение статистики критерия	$p\text{-value}$	Статистическое решение при $\alpha = \underline{0.05}$	Ошибка стат. решения
151.633331319262	1.0896405147776392e-43	$H_0$ отклоняется	нет

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	$p$ -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
$H_0: m_1 = m_2$	0.8524	0.6353	$H_0$ принимается	нет
$H_0: m_1 = m_3$	-10.619	0.001	$H_0$ отклоняется	нет
$H_0: m_2 = m_3$	-11.4715	0.001	$H_0$ отклоняется	нет

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