

# Лабораторная работа № 3

## «Однофакторный дисперсионный анализ»

студента Мельниковой М.Н. Дата сдачи: \_\_\_\_\_  
 Ведущий преподаватель: Трофимов А.Г. оценка: \_\_\_\_\_ подпись: \_\_\_\_\_

### Вариант №9

*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, $n_i$
$X_1$	$N(10, 2)$	$m = 10$ $\sigma = 2$	10	4	50
$X_2$	$N(5, 2)$	$m = 5$ $\sigma = 2$	5	4	50
$X_3$	$N(10, 5)$	$m = 10$ $\sigma = 5$	10	25	50

Количество случайных величин  $k = 3$

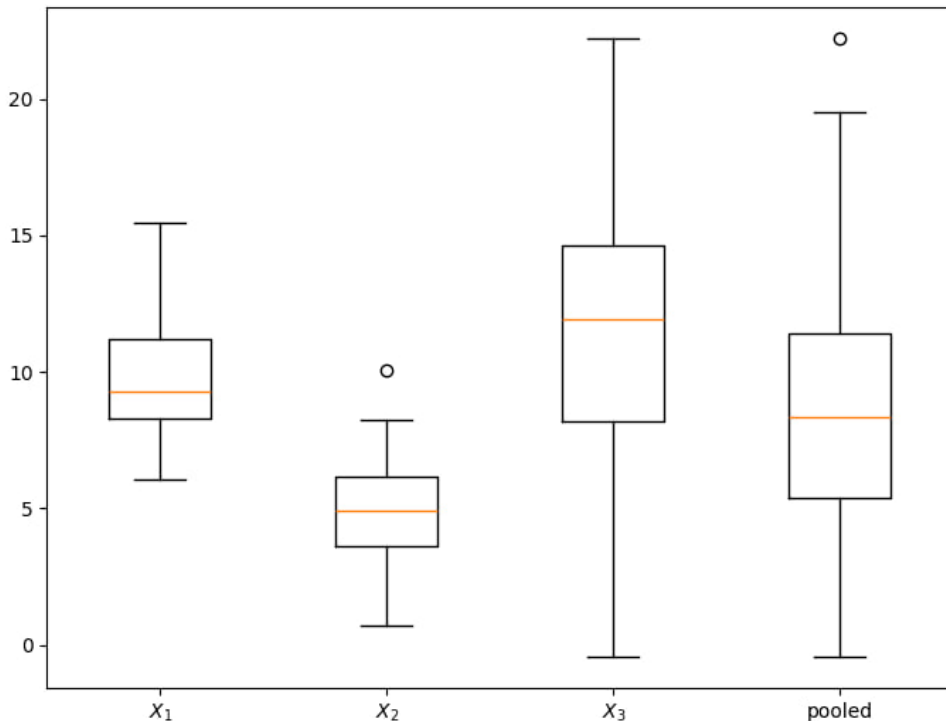
*Примечание:* для генерации случайных чисел использовать функции **rand**, **randn**, **chi2rnd** (**scipy.stats: uniform.rvs, norm.rvs, chi2.rvs**)

Выборочные характеристики:

СВ	Среднее, $\bar{x}_i$	Оценка дисперсии, $s_i^2$	Оценка с.к.о., $s_i$
$X_1$	9.714443023899921	4.723387428367179	2.173335553559822
$X_2$	4.828114676459174	3.971797523457107	1.9929369090508378
$X_3$	11.386237826833872	23.18653468642451	4.8152398368538725
<i>Pooled</i>	8.642931842397653	18.36947119470521	4.28596210840754

## 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
49.285264254844634	1.9853672219890e-11	$H_0$ 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	$D_b^* = 130.662450811$	$K-1=2$	$n * D_b^* / (K - 1) = 9799.683810845048$
Остаточные признаки	$D_w^* = 10.6272398794$	$n-K=148$	$n * D_w^* / (n - K) = 10.844122325934963$
Все признаки	$D_x^* = 141.289690690$	$n-1=149$	$n * D_x^* / (n - 1) = 142.23794364833918$

Эмпирический коэффициент детерминации  $\eta^2 = 0.9247840388957904$

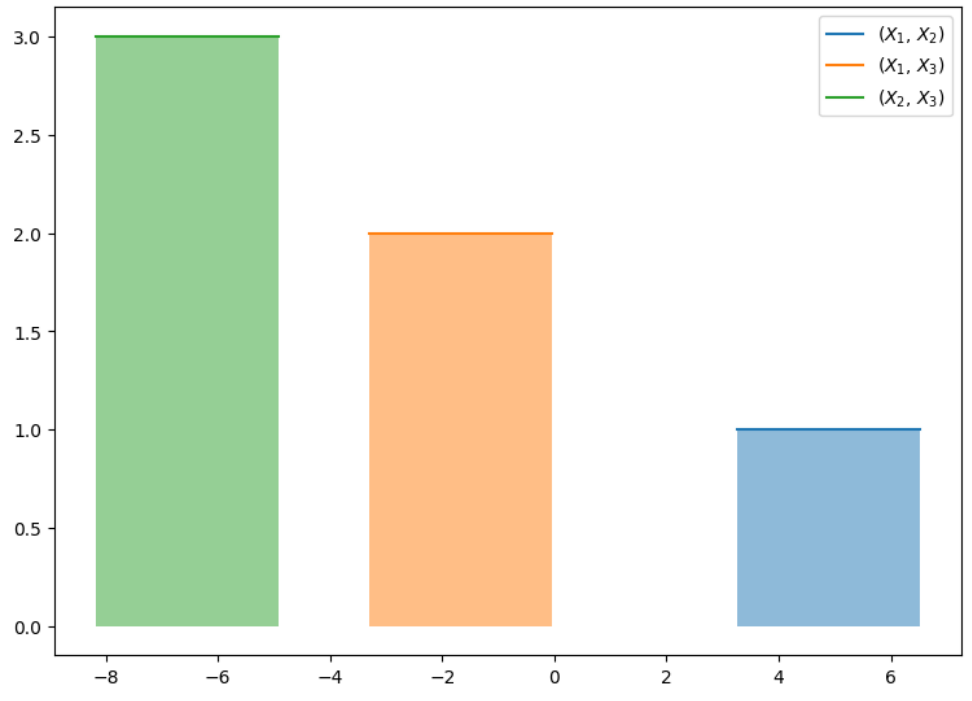
Эмпирическое корреляционное отношение  $\eta = 0.9616569236977345$

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

Выборочное значение статистики критерия	$p\text{-value}$	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
53.54673538290311	3.3947003972580522e-18	$H_0$ 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	<i>p-value</i>	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
$H_0: m_1 = m_2$	-4.8863	0.0	$H_0$ отклоняется	нет
$H_0: m_1 = m_3$	1.6718	0.0324	$H_0$ отклоняется	1-го рода
$H_0: m_2 = m_3$	6.5581	0.0	$H_0$ отклоняется	нет

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