

Лабораторная работа № 3
«Однофакторный дисперсионный анализ»

студента Грачева Игоря группы Б23-504. Дата сдачи: _____
Ведущий преподаватель: _____ оценка: _____ подпись: _____

Вариант №5

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, σ_i^2	Объем выборки, n_i
X_1	$N(5, 2)$	$m = 5, \sigma = 2$	5	4	50
X_2	$N(5, 2)$	$m = 5, \sigma = 2$	5	4	200
X_3	$N(5, 5)$	$m = 5, \sigma = 5$	5	25	100

Количество случайных величин $k = 3$

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

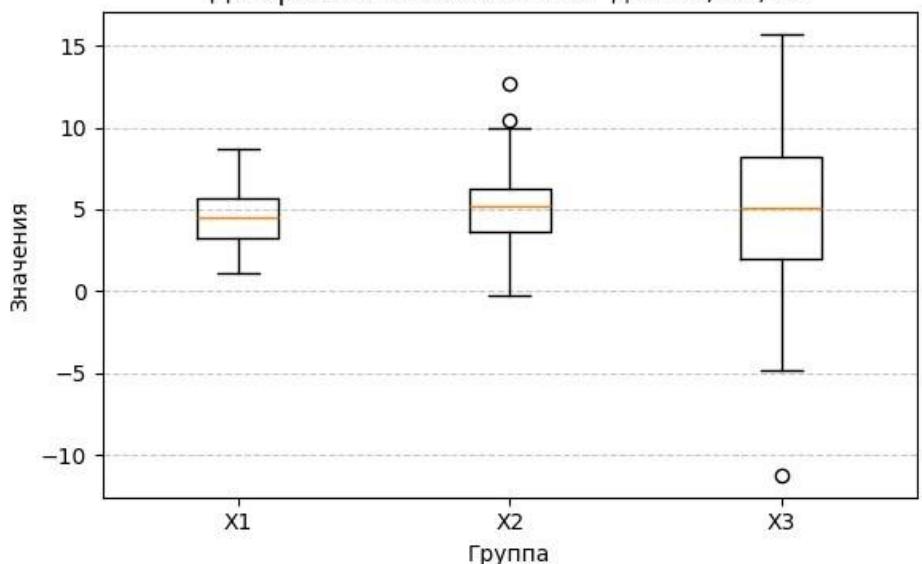
Выборочные характеристики:

СВ	Среднее, \bar{x}_i	Оценка дисперсии, s_i^2	Оценка с.к.о., s_i
X_1	4.5491	3.4869	1.8673
X_2	5.1067	3.7499	1.9365
X_3	5.0275	21.8770	4.6773
<i>Pooled</i>	5.0044	8.8845	2.9807

2. Визуальное представление выборок

Диаграммы *Box-and-Whisker*:

Диаграммы Box-and-Whisker для X1, X2, X3



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
127.5754	$1.98 \cdot 10^{-28}$	Отвергается	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	Несмешённая оценка
Группировочный признак	12.5127	2	6.2564
Остаточные признаки	3082.9111	347	8.8845
Все признаки	3095.4238	349	8.8694

Эмпирический коэффициент детерминации $\eta^2 = 0.0040$

Эмпирическое корреляционное отношение $\eta = 0.0636$

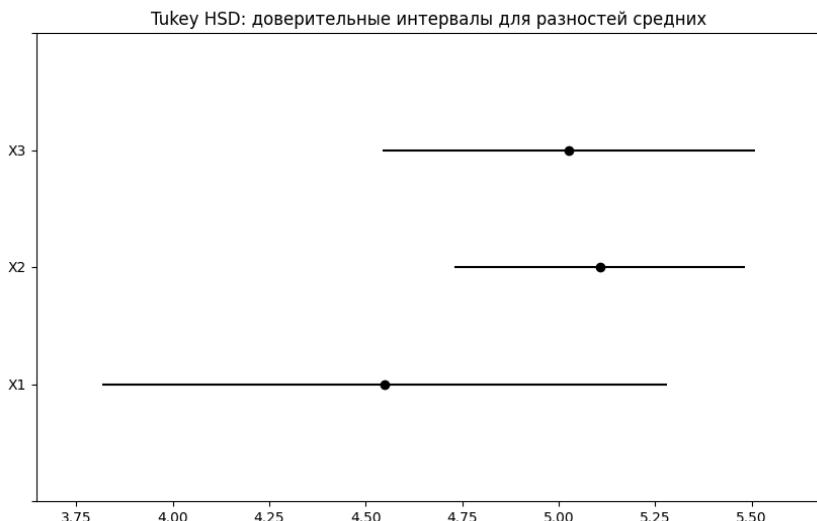
Статистическая гипотеза: $H_0 : m_1 = \dots = m_k$

Выборочное значение статистики критерия	p-value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
0.7042	0.495214	Принимается	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
$H_0: m_1 = m_2$	$t = -1.1831$	0.24	Не отвергается	Нет
$H_0: m_1 = m_3$	$t = -0.9266$	0.36	Не отвергается	Нет
$H_0: m_2 = m_3$	$t = -0.2170$	0.83	Не отвергается	Нет

Примечание: при расчетах использовать функцию **multcompare** (`statsmodels.stats.multicomp.pairwise_tukeyhsd`)