

Лабораторная работа № 3

«Однофакторный дисперсионный анализ»

студента Никитина Александра группы Б23-504. Дата сдачи: 23.11.2025

Ведущий преподаватель: Басакова А.В. оценка: _____ подпись: _____

Вариант №13

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	$R(-5,5)$	$a_1 = -5,$ $b_1 = 5$	$m_1 = \frac{a_1 + b_1}{2} = 0$	$\sigma_1^2 = \frac{(b_1 - a_1)^2}{12}$ $= \frac{25}{3}$	200
X_2	$R(-5,5)$	$a_2 = -5,$ $b_2 = 5$	$m_2 = \frac{a_2 + b_2}{2} = 0$	$\sigma_2^2 = \frac{(b_2 - a_2)^2}{12}$ $= \frac{25}{3}$	100
X_3	$N(-5,3)$	$m_3 = -5,$ $\sigma_3 = 3$	$m_3 = -5$	$\sigma_3^2 = 3$	200

Количество случайных величин $k = 3$

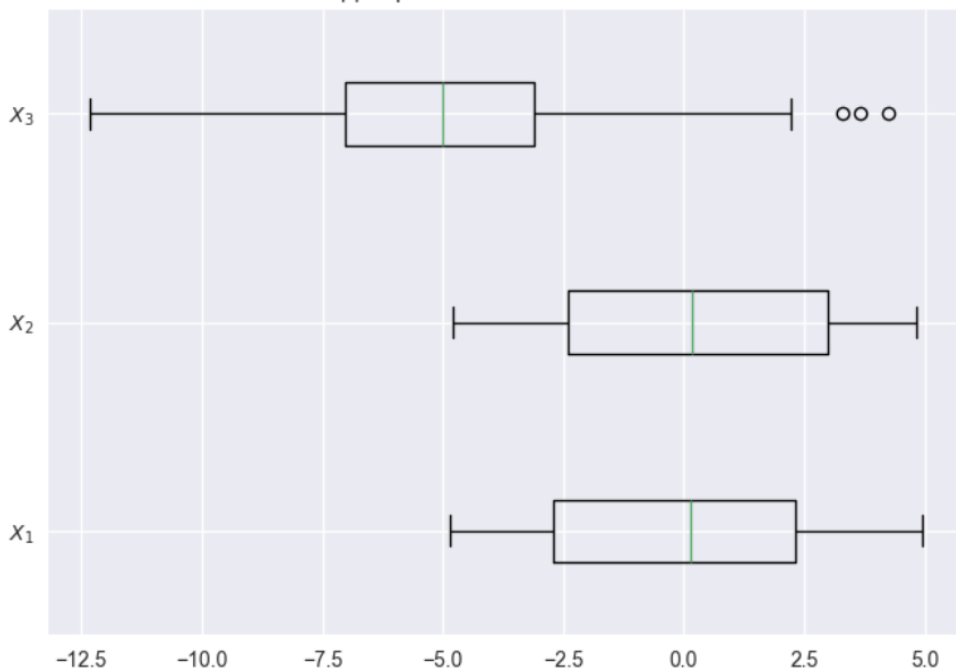
Примечание: для генерации случайных чисел использовать функции **rand**, **randn**, **chi2rnd** (scipy.stats: **uniform.rvs**, **norm.rvs**, **chi2.rvs**)

Выборочные характеристики:

СВ	Среднее, \bar{x}_i	Оценка дисперсии, s_i^2	Оценка с.к.о., s_i
X_1	-0.06	8.39	2.90
X_2	8.54	8.54	2.92
X_3	-5.07	9.44	3.07
<i>Pooled</i>	-2.02	15.03	3.88

2. Визуальное представление выборок

Диаграммы Box-and-Whisker



Диаграммы *Box-and-Whisker*:

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

3. Проверка условия применимости дисперсионного анализа

Статистическая гипотеза: $H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$

Критерий Бартлетта:

Выборочное значение статистики критерия	p -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
0.75	0.00	H_0 принимается	2-го рода

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	$\tilde{D}_{\text{межгр}} = 6.21$	$k - 1 = 2$	$\frac{n}{k - 1} \tilde{D}_{\text{межгр}} = 1552.80$
Остаточные признаки	$\tilde{D}_{\text{внутр}} = 8.79$	$n - k = 497$	$\frac{n}{n - k} \tilde{D}_{\text{внутр}} = 8.84$
Все признаки	$\tilde{D}_{\text{общ}} = 15.00$	$n - 1 = 499$	$\frac{n}{n - 1} \tilde{D}_{\text{общ}} = 15.03$

Эмпирический коэффициент детерминации $\eta^2 = 0.41$

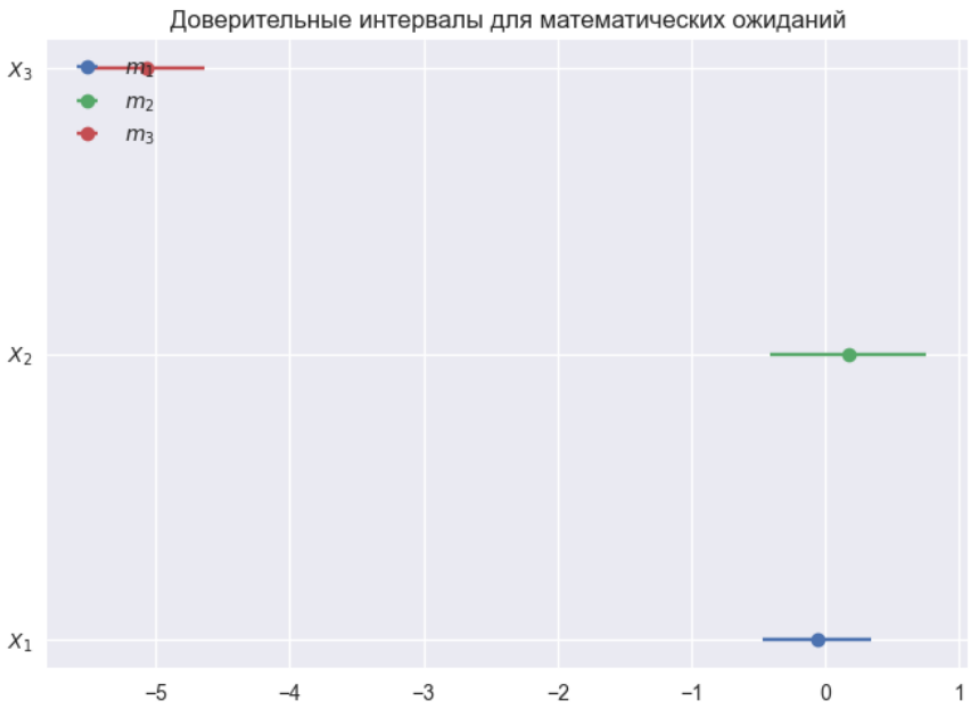
Эмпирическое корреляционное отношение $\eta = 0.64$

Статистическая гипотеза: $H_0: m_1 = m_2 = m_3$

Выборочное значение статистики критерия	p -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
175.65	0.00	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
$H_0: m_1 = m_2$	0.23	0.80	H_0 принимается	Нет
$H_0: m_1 = m_3$	-5.01	0.00	H_0 отклоняется	Нет
$H_0: m_2 = m_3$	-5.24	0.00	H_0 отклоняется	Нет

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