

Лабораторная работа № 3
«Однофакторный дисперсионный анализ»

студента Шамаева Сергея группы Б21-514 . Дата сдачи: _____

Ведущий преподаватель: Трофимов А.Г. оценка: _____ подпись: _____

Вариант № 4(19)

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, 15)	5, 15	10	2.886	100
X_2	N(10, 5)	10, 5	10	25	100
X_3	N(10, 2)	10, 2	10	4	100
X_4	N(10, 5)	10, 5	10	25	100

Количество случайных величин $k = 4$

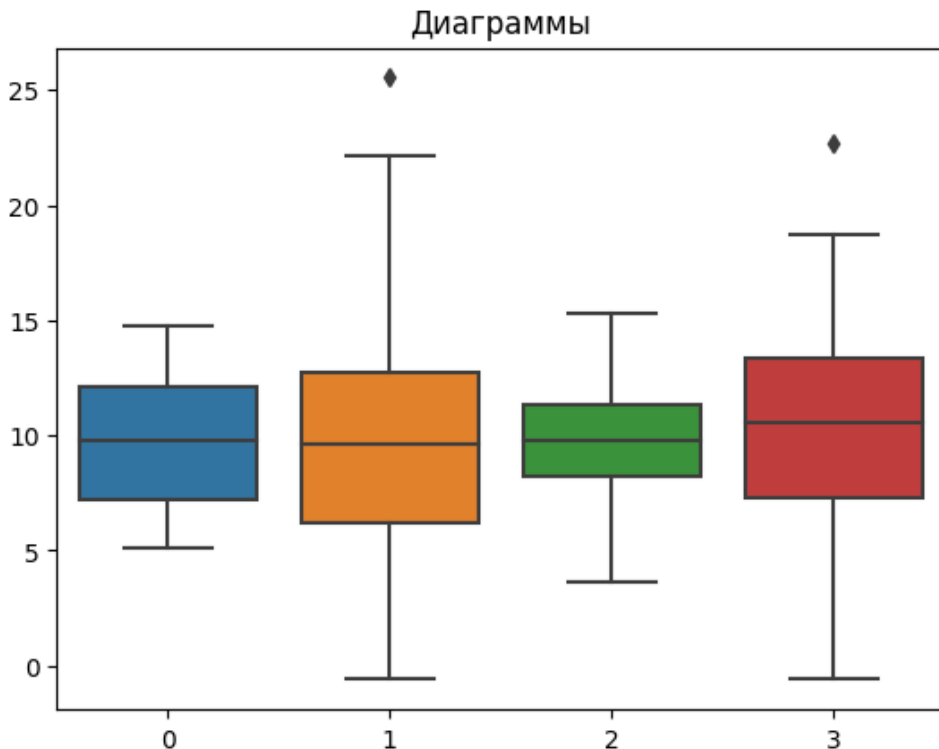
Примечание: для генерации случайных чисел использовать функции **rand**, **randn**, **chi2rnd** (**scipy.stats: uniform.rvs, norm.rvs, chi2.rvs**)

Выборочные характеристики:

СВ	Среднее, \bar{x}_i	Оценка дисперсии, s_i^2	Оценка с.к.о., s_i
X_1	9.782	8.129	2.851
X_2	9.922	27.092	5.205
X_3	9.658	4.934	2.221
X_4	10.210	20.739	4.554
Pooled	9.893	15.152	3.893

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
85.627	1.903e-18	H_0 не принимается	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	Дмежгр = 0.042	$K - 1 = 3$	$\text{Дмежгр} * n / (k - 1) = 5.643$
Остаточные признаки	Двнгр = 15.224	$n - k = 396$	$\text{Двнгр} * n / (n - k) = 15.377$
Все признаки	Добщ = 15.151	$n - 1 = 399$	$\text{Добщ} * n / (n - 1) = 15.189$

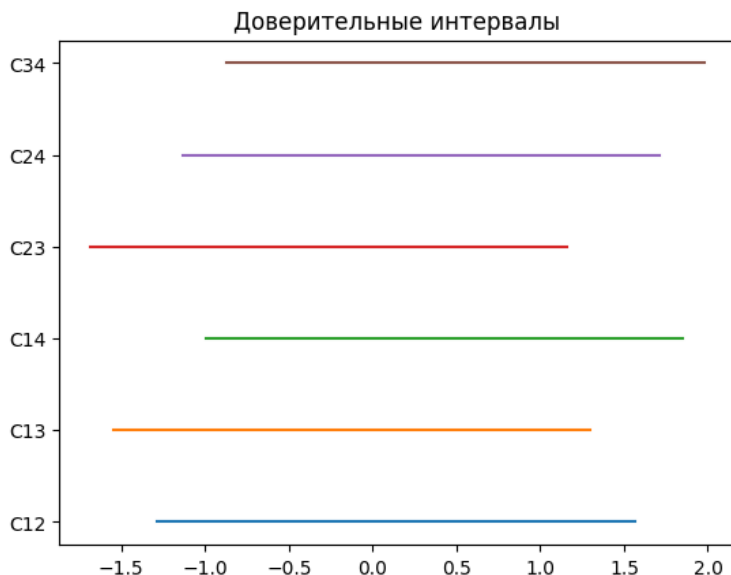
Эмпирический коэффициент детерминации $\eta^2 = \text{Дмежгр} / \text{Добщ} = 0.003$

Эмпирическое корреляционное отношение $\eta = 0.053$

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
0.371	0.774	Н0 принимается	нет

Примечание: при расчетах использовать функцию **anova1** (**scipy.stats.f_oneway**)

5. Метод линейных контрастовДоверительные 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
$H_0: m_1 = m_2$	0.140	0.994	H_0 принимаем	нет
$m_1 = m_3$	-0.124	0.996	H_0 принимаем	нет
$m_1 = m_4$	0.429	0.865	H_0 принимаем	нет
$m_2 = m_3$	-0.264	0.963	H_0 принимаем	нет
$m_2 = m_4$	0.289	0.953	H_0 принимаем	нет
$m_3 = m_4$	0.553	0.749	H_0 принимаем	нет

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