

Лабораторная работа № 3
«Однофакторный дисперсионный анализ»

студента Розинко Е.Д. группы Б21-524 Дата сдачи: 01.12.2023
Ведущий преподаватель: Трофимов А.Г. оценка: _____ подпись: _____

Вариант № 4

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	$R(5, 15)$	$R(a, b)$	10	8,3	100
X_2	$N(10, 5)$	$N(m, \sigma)$	10	25	100
X_3	$N(10, 2)$	$N(m, \sigma)$	10	4	100
X_4	$N(10, 5)$	$N(m, \sigma)$	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**)

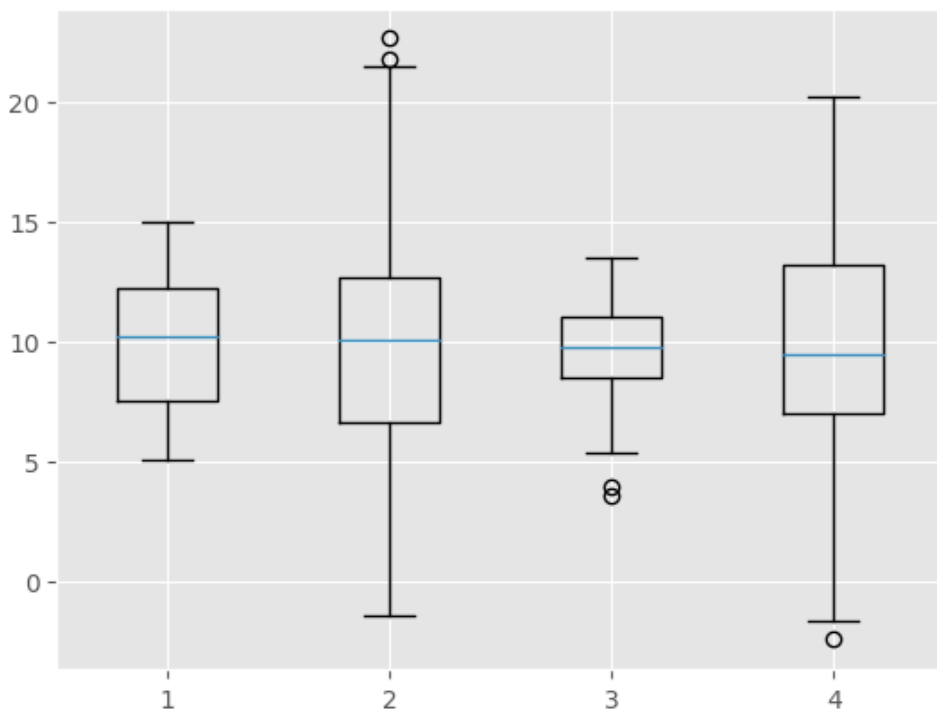
Выборочные характеристики:

СВ	Среднее, \bar{x}_i	Оценка дисперсии, s_i^2	Оценка с.к.о., s_i
X_1	9.967	8.699	2.935
X_2	9.812	19.188	4.358
X_3	9.650	3.975	1.984
X_4	10.031	23.111	4.783
<i>Pooled</i>	9.865	13.743	6.872

2. Визуальное представление выборок

Диаграммы *Box-and-Whisker*:

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
83.236	6.203	H_0 принимается	2 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	$D_b^* = 0.022$	$k - 1 = 3$	$(n/(k-1))D_b^* = 2.895$
Остаточные признаки	$D_w^* = 13.606$	$n - k = 396$	$(n/(n-k))D_w^* = 13.743$
Все признаки	$D_x^* = 13.628$	$n - 1 = 399$	$(n/(n-1))D_x^* = 13.662$

Эмпирический коэффициент детерминации $\eta^2 = 0.002$

Эмпирическое корреляционное отношение $\eta = 0.040$

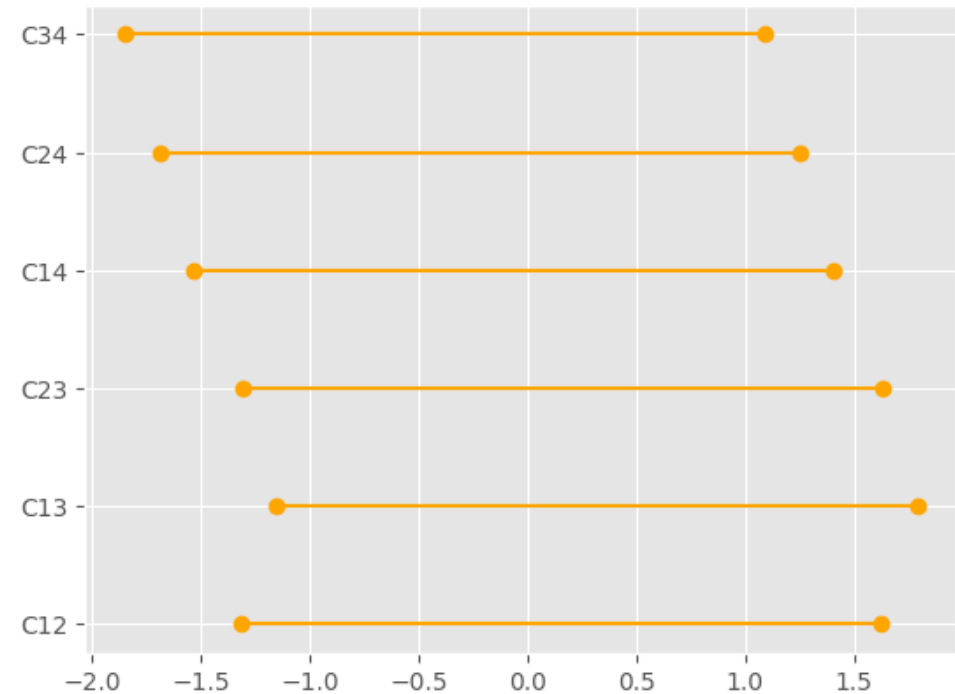
Статистическая гипотеза: $H_0 : m_1 = \dots = m_k$

Выборочное значение статистики критерия	p -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
0.211	0.889	H_0 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
$H_0: m_1 = m_2$	-0.1545	0.9911	H_0 принимается	Нет
$H_0: m_1 = m_3$	-0.3165	0.9308	H_0 принимается	Нет
$H_0: m_2 = m_3$	-0.1621	0.9897	H_0 принимается	Нет
$H_0: m_1 = m_4$	0.0643	0.9993	H_0 принимается	Нет
$H_0: m_2 = m_4$	0.2188	0.9755	H_0 принимается	Нет
$H_0: m_3 = m_4$	0.3809	0.8865	H_0 принимается	Нет

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