

Лабораторная работа № 3 «Однофакторный дисперсионный анализ»

студента Баранова Александра группы Б22-534. Дата сдачи: 29.11.2024

Ведущий преподаватель: Новиков М.А. оценка: _____ подпись: _____

Вариант №2

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	$R(2,6)$	$a_1 = 2, b_1 = 6$	$m_1 = \frac{a_1 + b_1}{2} = 4$	$\sigma_1^2 = \frac{(b_1 - a_1)^2}{12} = \frac{4}{3}$	100
X_2	$R(2,6)$	$a_2 = 2, b_2 = 6$	$m_2 = \frac{a_2 + b_2}{2} = 4$	$\sigma_2^2 = \frac{(b_2 - a_2)^2}{12} = \frac{4}{3}$	250
X_3	$R(0,8)$	$a_3 = 0, b_3 = 8$	$m_3 = \frac{a_3 + b_3}{2} = 4$	$\sigma_3^2 = \frac{(b_3 - a_3)^2}{12} = \frac{16}{3}$	100
X_4	$N(4,1)$	$m_4 = 4, \sigma_4 = 1$	$m_4 = 4$	$\sigma_4^2 = 1$	100

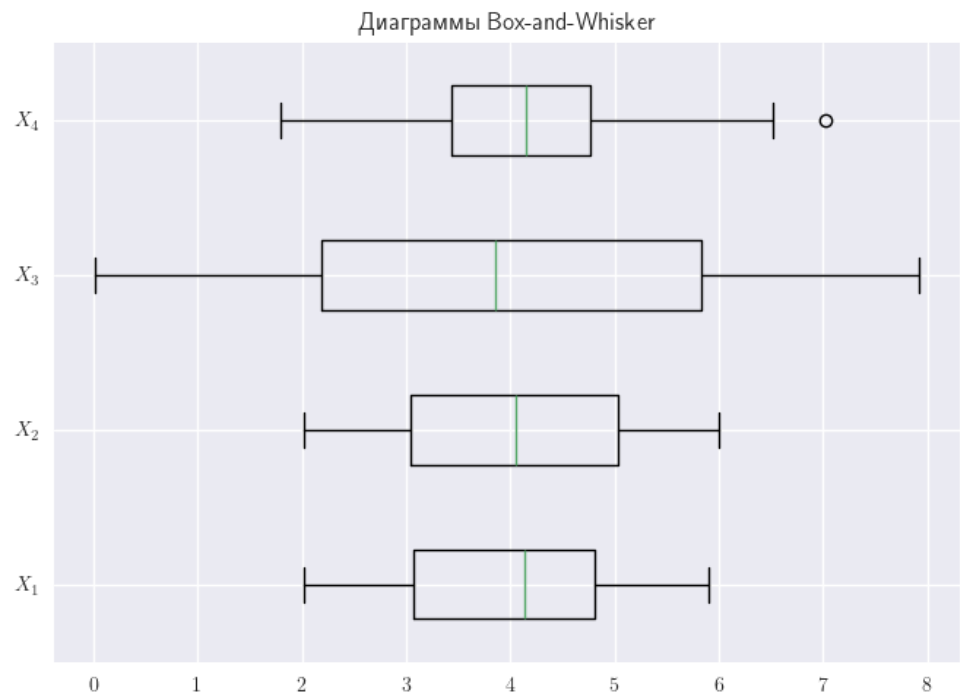
Количество случайных величин $k = 4$

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	3.99	1.23	1.11
X_2	4.04	1.37	1.17
X_3	3.90	5.22	2.28
X_4	4.16	1.02	1.01
<i>Pooled</i>	4.03	1.97	1.41

2. Визуальное представление выборок



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

3. Проверка условия применимости дисперсионного анализа

Статистическая гипотеза: $H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$

Критерий Барлетта:

Выборочное значение статистики критерия	<i>p-value</i>	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
107.99	0.00	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	$\tilde{D}_{\text{межгр}} = 0.01$	$k - 1 = 3$	$\frac{n}{k - 1} \tilde{D}_{\text{межгр}} = 1.19$
Остаточные признаки	$\tilde{D}_{\text{внутр}} = 1.96$	$n - k = 546$	$\frac{n}{n - k} \tilde{D}_{\text{внутр}} = 1.98$
Все признаки	$\tilde{D}_{\text{общ}} = 1.97$	$n - 1 = 549$	$\frac{n}{n - 1} \tilde{D}_{\text{общ}} = 1.97$

Эмпирический коэффициент детерминации $\eta^2 = 0.00$

Эмпирическое корреляционное отношение $\eta = 0.06$

Статистическая гипотеза: $H_0 : m_1 = m_2 = m_3 = m_4$

Выборочное значение статистики критерия	<i>p-value</i>	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
0.60	0.61	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	p -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
$H_0 : m_1 = m_2$	0.06	0.99	H_0 принимается	Нет
$H_0 : m_1 = m_3$	-0.08	0.97	H_0 принимается	Нет
$H_0 : m_1 = m_4$	0.17	0.82	H_0 принимается	Нет
$H_0 : m_2 = m_3$	-0.14	0.83	H_0 принимается	Нет
$H_0 : m_2 = m_4$	0.12	0.90	H_0 принимается	Нет
$H_0 : m_3 = m_4$	0.26	0.56	H_0 принимается	Нет

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