

**Лабораторная работа № 3**  
**«Однофакторный дисперсионный анализ»**

студента Васильев группы Б21 - 524. Дата сдачи: 28.11.2023  
Ведущий преподаватель: Трофимов А.Г. оценка:  
подпись: \_\_\_\_\_

Вариант № 6

*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, $\sigma_i^2$	Объем выборки, $n_i$
$X_1$	$\chi^2(2)$	$n$	2	4	100
$X_2$	$N(3, 1)$	$N(m, \sigma)$	3	1	150
$X_3$	$R(-2, 2)$	$R(a, b)$	0	1.33	200
$X_4$	$N(5, 1)$	$N(m, \sigma)$	5	1	100

Количество случайных величин  $k = \underline{\quad 4 \quad}$

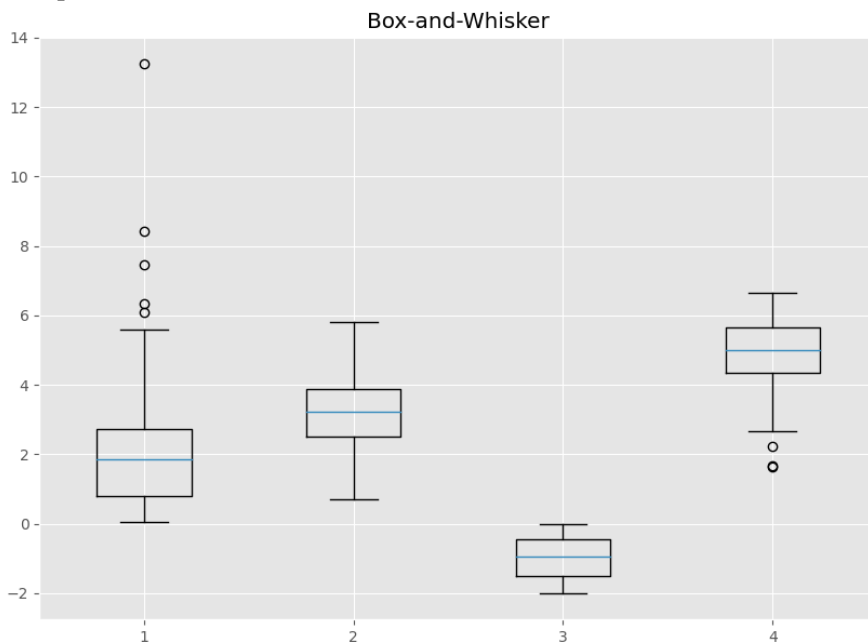
*Примечание:* для генерации случайных чисел использовать функции **rand**,  
**randn**, **chi2rnd** (scipy.stats: **uniform.rvs**, **norm.rvs**, **chi2.rvs**)

Выборочные характеристики:

СВ	Среднее, $\bar{x}_i$	Оценка дисперсии, $s_i^2$	Оценка с.к.о., $s_i$
$X_1$	2.175	4.194	2.038
$X_2$	3.197	0.959	0.976
$X_3$	-0.975	0.333	0.575
$X_4$	4.902	0.956	0.973
<i>Pooled</i>	1.804	1.317	0.658

## 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 $\alpha = \underline{\hspace{1cm}0.05\hspace{1cm}}$	Ошибка стат. решения
237.251351	3.740133931374756e-51	$H_0$ 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	$D_b^*=5.109$	$K - 1 = 3$	$\frac{n}{K-1}D_b^*=936.661$
Остаточные признаки	$D_w^*=1.307$	$n - K = 546$	$\frac{n}{n-K}D_w^*=1.317$
Все признаки	$D_x^*=6.416$	$n - 1 = 549$	$\frac{n}{n-1}D_x^*=6.428$

Эмпирический коэффициент детерминации  $\eta^2 = 0.796$

Эмпирическое корреляционное отношение  $\eta = 0.892$

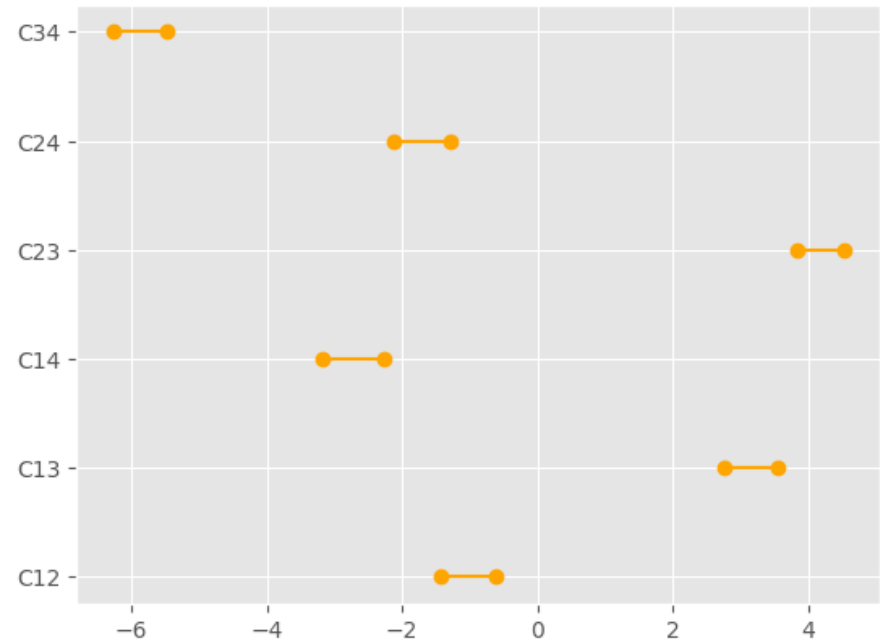
Статистическая гипотеза:  $H_0 : m_1 = \dots = m_k$

Выборочное значение статистики критерия	<i>p-value</i>	Статистическое решение при $\alpha = \underline{\underline{0.05}}$	Ошибка стат. решения
711.3621236866518	3.897966969970542e-188	$H_0$ 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
$H_0: m_1 = m_2$	-1.7672	0	$H_0$ принимается	нет

$H_0: m_1 = m_3$	-0.2057	0.8917	$H_0$ принимается	нет
$H_0: m_1 = m_4$	1.0216	0.012	$H_0$ принимается	нет
$H_0: m_2 = m_3$	1.5615	0	$H_0$ принимается	нет
$H_0: m_2 = m_4$	2.7887	0	$H_0$ принимается	нет
$H_0: m_3 = m_4$	1.2273	0.0001	$H_0$ принимается	нет

*Примечание:* при расчетах использовать функцию **multcompare** (**statsmodels.stats.multicomp.pairwise\_tukeyhsd**)