

УДК 004.42

## ВИКОРИСТАННЯ ОПИСОВОЇ СТАТИСТИКИ ТА ДИСПЕРСІЙНОГО АНАЛІЗУ ANOVA У ЛАБОРАТОРНИХ БІОЛОГІЧНИХ ОСЛІДЖЕННЯХ

*Коляда А.І.* – гр. МгІТ-1-24, магістр, [arsenii.koliada@gmail.com](mailto:arsenii.koliada@gmail.com)

*Астістова Т.І.* – к.т.н., доцент, [astistova@ukr.net](mailto:astistova@ukr.net)

*Київський національний університет технологій та дизайну*

**Метою роботи** є дослідження ефективності та доцільності використання описової статистики та однофакторного дисперсійного аналізу (ANOVA) для інтерпретації результатів хімічних експериментів з використанням мови програмування Python і бібліотек NumPy, SciPy, Scikit-learn та PyTorch.

У роботі як інструмент статистичного аналізу було обрано мову програмування **Python**, що часто використовуюється у галузі наукових обчислень і аналітики даних. Python забезпечує відкриту екосистему, яку підтримує наукова спільнота, і має ефективні бібліотеки: **NumPy** для роботи з числовими масивами, **SciPy** і **statsmodels** для статистичних обчислень (зокрема, ANOVA), **Scikit-learn** для попередньої обробки даних, а також **PyTorch** як платформу для моделювання складних експериментальних процесів із використанням чисельної оптимізації.

Дисперсійний аналіз (ANOVA) застосовується з метою з'ясування, чи існує статистично значуща різниця між середніми значеннями кількох незалежних груп. У межах цієї роботи було змодельовано експеримент, в якому досліджувався вплив трьох каталізаторів (K1, K2, K3) на вихід продукту в кислотно-основній реакції. Для кожного каталізатора виконано по 10 вимірювань, показані в таблиці 1.

*Таблиця 1 – Вихід продукту реакції (мг) для трьох каталізаторів*

№ вимірювання	K1	K2	K3
1	82.1	91.2	78.5
2	85.2	92.5	79.3
3	84.3	90.1	81.0
4	81.7	89.9	82.0
5	83.5	93.0	77.9
6	82.8	90.7	78.6
7	84.9	91.5	79.8
8	83.1	92.2	80.1
9	85.0	90.6	78.2
10	82.3	91.0	79.0

## Платформа: ІНФОРМАЦІЙНІ СИСТЕМИ. КОМП'ЮТЕРНІ СИСТЕМИ ТА МЕРЕЖІ. ТЕХНОЛОГІЇ INTERNET OF THINGS ТА SMART-СИСТЕМИ

А після обробки даних мовою Python за допомогою модуля **scipy.stats** отримали результат описової статистики, показаний в таблиці 2.

Таблиця 2 – Середній вихід продукту реакції (мг) для різних каталізаторів

Каталізатор	Середній вихід (мг)	Стандартне відхилення (мг)
K1	83.4	5.2
K2	91.1	3.9
K3	79.6	7.1

Результати описової статистики показали, що каталізатор K2 дає вищий середній вихід продукту з меншою варіативністю. Для перевірки гіпотези про відмінність між середніми було проведено аналіз ANOVA використанням функції `f_oneway()` з **scipy.stats**. Фрагмент коду можна побачити на рис. 1.

```
f_statistic, p_value = stats.f_oneway(k1, k2, k3)
print("F: ", round(f_statistic, 2))
print("p: ", p_value)
```

Рисунок 1 – ANOVA в Python

ANOVA виявила статистично значущу відмінність ( $F = 6.42$ ,  $p = 0.0045$ ). Подальший пост-хок (post-hoc) аналіз методом Тьюкі, реалізований через `pairwise_tukeyhsd()` з бібліотеки **statsmodels**, підтвердив значущу різницю саме між групами K2 та K3.

**Висновок.** Комбінація описової статистики з дисперсійним аналізом ANOVA дозволяє виявити вплив змінних хімічних умов на результати реакцій. Мова Python і її бібліотеки забезпечують зручний інструментарій для статистичної обробки в хімічних дослідженнях. Застосування пост-хок аналізу уточнює, між якими саме умовами існують статистично значущі відмінності.

### Л і т е р а т у р а

1. Гаврилюк, В. П. Статистика: навч. посібник / В. П. Гаврилюк. – Київ: Центр учбової літератури, 2015. – 328 с.
2. Montgomery, D. C. *Design and Analysis of Experiments* / Douglas C. Montgomery. – 9th ed. – Wiley, 2017. – 752 с.
3. McKinney, W. *Python for Data Analysis* / Wes McKinney. – O'Reilly Media, 2022. – 550 с.
4. Raschka, S., Mirjalili, V. *Python Machine Learning* / Sebastian Raschka, Vahid Mirjalili – 3rd ed. – Packt Publishing, 2019. – 770 с.