

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА
ФАКУЛЬТЕТ КОМП'ЮТЕРНИХ НАУК ТА КІБЕРНЕТИКИ
Кафедра прикладної статистики**



**РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ
«Комп'ютерна статистика»**

для студентів

галузь знань	12 «Інформаційні технології»
спеціальність	124 «Системний аналіз»
освітній рівень	бакалавр
освітня програма	«Системний аналіз»
вид дисципліни	за вибором

Форма навчання	денна
Навчальний рік	2023/2024
Семестр	8
Кількість кредитів ECTS	3
Мова викладання, навчання та оцінювання	українська
Форма заключного контролю	залік

Викладачі: **д.ф.-м.н, доц. Ірина РОЗОРА**


Пролонговано: на 20__/20__ н.р. _____ (_____) «__» 20__ р.
на 20__/20__ н.р. _____ (_____) «__» 20__ р.

КИЇВ – 2022

Розробник: Ірина РОЗОРА док. фіз.-мат. н., доцент,
доцент кафедри Прикладної Статистики


ЗАТВЕРДЖЕНО

Завідувач кафедри прикладної статистики


_____ (Ірина РОЗОРА)

Протокол № 1 від «27» серпня 2022 р.

Схвалено Гарантом освітньо-професійної програми «Системний аналіз»



_____ к.ф.-м. н., доцент Михайло ШАРАПОВ
(підпис)

«31» серпня 2022 року

Схвалено науково-методичною комісією факультету комп'ютерних наук та кібернетики

Протокол № 1 від «31» серпня 2022 року

Голова науково-методичної комісії


_____ к.ф.-м. н., доцент Людмила ОМЕЛЬЧУК
(підпис)

1 Мета дисципліни – ознайомлення студентів зі статистичними методами аналізу даних за допомогою мови програмування та середовища для статистичних обчислень, аналізу та зображення даних в графічному вигляді R, а також практична підготовка студентів до проведення статистичних досліджень, їх аналізу та обробки.

2 Попередні вимоги до опанування або вибору навчальної дисципліни

Знати: основи теорії ймовірностей та математичної статистики

Вміти: формалізувати умови задач та скласти план розв'язку

Володіти елементарними навичками: розв'язувати типові задачі з теорії ймовірностей, математичної статистики.

3 Анотація навчальної дисципліни

Дисципліна «Комп'ютерна статистика» є складовою частиною циклу професійної підготовки фахівців освітньо-кваліфікаційного рівня „бакалавр”; вона включає вивчення програмного середовища для статистичних обчислень R на основі статистичних задач, а саме візуалізація даних, побудова довірчих інтервалів, перевірки статистичних критеріїв тощо. Викладається у 8-му семестрі, обсяг 90 год. (3 кредити ECTS), з них лекції – 26 год., практичних – 10 год., самостійна робота – 54 год. Передбачено 1 змістовий модуль та залік.

4 Завдання (навчальні цілі)

набуття знань, умінь та навичок (компетентностей) відповідно до освітньої кваліфікації бакалавра з системного аналізу. Зокрема, розвивати:

- K02. Здатність застосовувати знання у практичних ситуаціях
- K18. Здатність формалізувати проблеми, описані природною мовою, у тому числі за допомогою математичних методів, застосовувати загальні підходи до математичного моделювання конкретних процесів.
- ФКСАС 2. Здатність проводити аналітично обґрунтоване планування експериментів і спостережень, здійснювати статистичний аналіз отриманих результатів та коректно їх інтерпретувати

5 Результати навчання за дисципліною

Результат навчання (РН) (1 – знати; 2 – вміти; 3. комунікація; 4. автономність та відповідальність)		Форми викладання та навчання	Методи оцінювання	Відсоток у підсумковій оцінці з дисципліни
Код	Результат навчання			
РН.1	Знати і розуміти основні можливості програмного середовища R для аналізу та візуалізації даних	Лекції, практичні, самостійна робота	Поточне оцінювання (ПО), індивідуальний проєкт (ІП)	40
РН.2	Вміти обчислювати чи оцінювати основні числові показники у вибірці, реалізувати їх в R.			
РН.4.1	Виявляти здатність до самонавчання та продовження професійного розвитку	Самостійна робота	ПО, ІП	50
РН.4.2	Уміти організувати власну діяльність та одержувати результат у рамках обмеженого часу	Самостійна робота	ПО, ІП	10
РН.3	Демонструвати навички взаємодії з іншими людьми, уміння працювати в командах	Самостійна робота	ПО, ІП	

6 Співвідношення результатів навчання дисципліни з програмними результатами навчання

Результати навчання дисципліни					
Програмні результати навчання	РН1	РН2	РН3	РН4.1	РН4.2
<i>(з опису освітньої програми)</i>					
ПРСАС 1. Проводити статистичне оцінювання невизначених параметрів розподілів стохастичних факторів досліджуваних процесів, формалізувати стохастичні фактори у вигляді випадкових величин, векторів, процесів.	+	+			
ПРСАС 2. Застосовувати вивчені методи системного і статистичного аналізу, обробки даних та імітаційного моделювання.	+	+	+	+	+

7 Схема формування оцінки

7.1 Форми оцінювання студентів:

- семестрове оцінювання:

1. Контрольна робота (РН.1, РН.2): 30 балів
2. Поточне оцінювання (РН.1, РН.2): 40 балів
3. Індивідуальний проєкт (РН.1, РН.2): 30 балів

Контрольні запитання до змістового модуля I

Що таке R, де його взяти і як встановити. Робота з RStudio

Типи даних та елементарні функції в R

Статистики середнього положення та їх реалізація в R

Статистики розкиду і їх реалізація в R

Гістограми та полігон частот.

Стовпцеві та кругові діаграми.

Гафічна перевірка узгодженості розподілу. P-P і Q-Q діаграми

Скринька з вусами

Діаграми розсіювання

Коефіцієнт кореляції Пірсона. Його візуалізація

Ранги та рангові кореляції

Дискретні розподіли в R

Неперервні розподіли в R

Оцінки узагальненого методу моментів

оцінки методу максимальної вірогідності

Побудова довірчих інтервалів

Тести для даних з нормальним розподілом. t-тест перевірка середнього для однієї вибірки.

Тести для даних з нормальним розподілом. F-тест для перевірки дисперсії

Хі-2 критерії.

Критерій Колмогорова. Критерій Шапіро перевірки нормальності.

p-value. (P значення). Рівень значущості та потужність критерію.

Критерій Вілкоксона (Wilcoxon)

Проста множинна лінійна регресія в R

7.2 Організація оцінювання

Терміни проведення оцінювання

Модуль № 1 – до 7 тижня.

За відсутності студента з поважних причин перездача екзамену здійснюється відповідно до «Положення про порядок оцінювання знань студентів при кредитно-модульній системі організації навчального процесу» від 1 жовтня 2010 року.

7.3 Шкала відповідності оцінок

Відмінно / Excellent	90-100
Добре / Good	75-89
Задовільно / Satisfactory	60-74
Незадовільно / Fail	0-59

8. Структура навчальної дисципліни. Тематичний план лекцій і лабораторних занять

№	Назва лекції	Кількість годин		
		Лекції	Практичні	Самост. Робота
Змістовий модуль 1.				
1	Мова програмування R, де її взяти і як встановити. Робота з RStudio	2		4
2	Типи даних та елементарні функції в R. Неперервні та дискретні розподіли в R	2	2	4
3	Частотні розподіли та гістограми в R: Візуалізація даних.	2		4
4	Міри центральної тенденції: середнє, медіана, мода, піврозмах. Міри розсіювання: розмах, середнє абсолютне відхилення, дисперсія та стандартне відхилення. Коефіцієнт варіації.	2	2	4
5	Міри позиції: квантилі, квартилі, міжквартильний розмах. Міра форми або асиметрії: коефіцієнт асиметрії Пірсона.	2		4
6	Побудова довірчих інтервалів.	2		4
7	Побудова оцінок методом моментів та максимальної вірогідності	2		4
8	Критерій Стьюдента (t-критерій) та F-критерій для нормальних спостережень, їх різновиди в R	2		4
9	Критерій Колмогорова та Шапіро в R.	2	2	4
10	Реалізація χ^2 критеріїв незалежності,	2		4

	однорідності, перевірки розподілу в R			
11	Непараметричні критерії. Критерій Вілкоксона.	2		4
12	Лінійна та множинна регресія в R	2	2	4
13	Презентація та захист індивідуальних проєктів	2		6
	Модульна контрольна робота 1		2	
	Залік			
	ВСЬОГО	26	10	54

Загальний обсяг 90 год.¹, в тому числі:

Лекцій – 26 год.

Практичні – 10 год.

Самостійна робота - 54 год.

9. Рекомендовані джерела

1. Р.Є. Майборода. Комп'ютерна статистика, Навч. посібник. – К.: Видавничо-поліграф. центр “Київський університет”, 2019.
2. Карташов М.В. Ймовірність. Статистика. Процеси. – К.: Видавничо-поліграф. центр “Київський університет”, 2008.
3. В.М. Турчин. Теорія ймовірностей та математична статистика, 2-е видання, Підручник. – Дніпро: “Ліра”, 2018.

¹ Загальна кількість годин, відведених на дану дисципліну згідно навчального плану.