

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ПРИКАРПАТСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВАСИЛЯ СТЕФАНІКА**



Факультет математики та інформатики

Кафедра математичного і функціонального аналізу

СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

СТАТИСТИЧНИЙ АНАЛІЗ З R

Рівень вищої освіти – другий (магістерський)

Освітня програма Прикладна та теоретична статистика

Спеціалізація (за наявності) _____

Спеціальність 112 Статистика

Галузь знань 11 Математика та статистика

Затверджено на засіданні кафедри
Протокол № 1 від “25” серпня 2023 р.

м. Івано-Франківськ – 2023 р.

1. Загальна інформація

Назва дисципліни	Статистичний аналіз з R
Викладач (і)	Осипчук Михайло Михайлович
Контактний телефон викладача	380503732451
E-mail викладача	mykhailo.osypchuk@pnu.edu.ua
Формат дисципліни	Очний
Обсяг дисципліни	6 кредити ЄКТС, 180 год.
Посилання на сайт дистанційного навчання	https://d-learn.pnu.edu.ua/course/subscription/through/url/5bc49ea470983c1c42ae
Консультації	На заняттях та перед підсумковим контролем згідно розкладу занять.

2. Анотація до навчальної дисципліни

Предметом вивчення навчальної дисципліни є теоретичні відомості по роботі з інтерактивними компонентами R; опис мови і базових графічних можливостей середовища; опис поширених процедур обробки даних і побудови статистичних моделей; просторовий аналіз і графічне зображення результатів.

3. Мета та цілі навчальної дисципліни

Мета та цілі курсу полягають у формуванні в студентів нових теоретичних знань і практичних навичок з аналізу та візуалізації даних з використанням програмного середовища статистичних розрахунків R.

4. Програмні компетентності та результати навчання

Загальні компетентності:

ЗК-4. Здатність використовувати комп'ютерні та інформаційні технології в професійній діяльності.

Фахові компетентності:

ФК-1. Знання на рівні новітніх досягнень, необхідні для дослідницької або практичної діяльності у сфері математики, статистики та їхніх практичних застосувань.

ФК-7. Здатність аналізувати статистичні алгоритми, оцінювати їхню

обґрунтованість та ефективність.

ФК-8. Здатність використовувати спеціалізовані мови програмування та програмне забезпечення для розв'язання задач дослідницького або практичного характеру.

Результати навчання:

ПРН-9. Уміти використовувати в практичній діяльності та розробляти спеціалізоване статистичне програмне забезпечення.

ПРН-13. Уміти інтегрувати знання з різних галузей для розв'язання теоретичних та практичних задач і проблем.

5. Організація навчання

Обсяг навчальної дисципліни	
Вид заняття	Загальна кількість годин
лекції	28
семінарські заняття / практичні / лабораторні	0/16/16
самостійна робота	120

Ознаки навчальної дисципліни			
Семестр	Спеціальність	Курс (рік навчання)	Нормативний / вибірковий
1	112 Статистика	1	Нормативна

Тематика навчальної дисципліни			
Тема	кількість год.		
	лекції	заняття	сам. роб
Змістовий модуль 1. Програмне середовище R			
Тема 1. Основні компоненти статистичного середовища R. Історія появи та основні принципи організації середовища R. Робота з командною панеллю інтерфейсу R. Об'єкти, пакети, функції та пристрої.	2		4
Тема 2. Опис мови R. Типи даних мови R. Вектори і матриці. Фактори. Списки і таблиці. Імпортуювання даних в R. Зображення даних і часу: часові ряди.	2	4	12

Організація обчислень: функції, розгалуження, цикли. Векторизовані обчислення в R із використанням apply-функцій.			
Тема 3. Графічні можливості R. Діаграми розсіювання plot() і параметри графічних функцій. Гістограми, функції ядерної густини і функція cdfplot(). Діаграми розмахів. Кругові і стовпчикові діаграми. Діаграми Клівленда та одновимірні діаграми розсіювання. Категоризовані графіки.	2	4	12
Змістовий модуль 2. Статистичний аналіз			
Тема 4. Описова статистика і підбір розподілів. Оцінка вибірових параметрів із використанням спеціальних функцій. Використання функцій summary() і додаткових пакетів. Аналіз викидів. Заповнення пропущених значень у таблицях даних. Відтворюваність результатів при використанні генератора випадкових чисел. Закони розподілу ймовірностей, реалізовані в R. Підбір закону і розподілу в R. Перевірка на нормальність розподілу.	4	4	16
Тема 5. Класичні методи і критерії статистики. Гіпотеза про рівність середніх двох генеральних сукупностей. Використання рангових критеріїв Уїлкоксона – Манна – Уїтні. Рандомізація, бутстреп та оцінка статистичної потужності (на прикладі двохвибіркового t-критерія). Гіпотеза про однорідність дисперсій. Вступ в дисперсійний аналіз. Оцінка кореляції двох випадкових величин. Критерій хі-квадрат. Точний тест Фішера, критерії Мак-Немара і Кохрана – Мантеля – Хензеля. Оцінка статистичної потужності при порівнянні часток.	4	4	16
Тема 6. Лінійні моделі в дисперсійному аналізі. Протокол розвідувального аналізу даних. Лінійні моделі дисперсійного аналізу. Структура модельних об'єктів дисперсійного аналізу. Оцінка адекватності моделі дисперсійного аналізу. Дисперсійний аналіз за Краскелом – Воллісом. Моделі дво- і багатофакторного дисперсійного аналізу. Особливості в лінійних моделях, що містять категоріальні предиктори. Проблема численних перевірок статистичних гіпотез. Методи порівняння групових середніх в дисперсійному аналізі.	4	4	16
Тема 7. Регресійні моделі залежності між кількісними змінними. Статистична модель. Проста лінійна регресія: який вік Всесвіту? Моделі регресії при різних видах функції втрат. Критерії вибору моделей оптимальної складності. Поліноміальні і нелінійні моделі регресії. Модель множинної	4	4	16

регресії і вибір її специфікації. Процедури діагностики моделей множинної регресії. Регресія гребеня і ласо: використання головних компонент. Порівняння ефективності різних моделей при прогнозуванні.			
Тема 8. Узагальнені, структурні та інші моделі регресії. Моделі згладжування. Узагальнені моделі регресії. Моделі logit- та probit-регресії. Узагальнені моделі для оцінки показників екологічної толерантності. Коваріаційний аналіз. Моделі зі змішаними ефектами для ієрархічно організованих даних. Індуктивні моделі (метод групового обліку аргументів). Моделювання структурними рівняннями.	4	4	16
Тема 9. Просторовий аналіз і створення криптограм. Проста карта: використання растрового малюнка та обчислення відстаней. Аналіз просторового розміщення точок. Використання сервісів картографічної системи Google Maps. Створення картограм за допомогою R.	2	4	12
ЗАГ.:	28	32	120

6. Система оцінювання навчальної дисципліни

Загальна система оцінювання навчальної дисципліни	Для оцінювання навчальних досягнень студентів при вивченні дисципліни передбачено одну контрольну роботу. Підсумковим контролем є екзамен. Оцінка студента з дисципліни є сумою оцінок за контрольну роботу (50%) та оцінки за екзамен (50%).
Вимоги до письмових робіт	Письмова контрольна робота виконується студентом самостійно без доступу до будь-яких джерел інформації в терміни визначені викладачем на одному з практичних занять. Структуру завдань визначає викладач зважаючи на достатність для контролю знань студентів та можливість виконання добре підготовленим студентом за відведений час.
Семінарські заняття	Практичні заняття призначені для навчання студентів застосовувати теоретичні аспекти дисципліни до розв'язування задач пов'язаних з тематикою курсу. На практичних заняттях здійснюється контроль самостійної роботи студентів та виконання контрольних робіт.
Умови допуску до підсумкового контролю	Всі студенти, які прослухали курс, допускаються до

	підсумкового контролю.
Підсумковий контроль	Екзамен. Форма екзамену комбінована (письмове виконання з теоретичних та практичних завдань з можливим обговорення результатів з екзаменатором). Кожен білет містить два теоретичних та три практичних завдання з питань програми курсу. Кожне завдання оцінюється від 0 до 10 балів в залежності від якості та повноти виконання.

7. Політика навчальної дисципліни

Письмові роботи:

Всі контрольні завдання студент виконує самостійно.

Академічна доброчесність:

Порушення вимоги самостійності виконання завдань курсу призводить до нульової оцінки за відповідний контрольний захід.

Відвідування занять

Пропущене заняття не оцінюється. Пропуски занять відпрацьовуються шляхом демонстрації виконання всіх завдань пропущеного заняття.

Неформальна освіта:

Можливе зарахування результатів неформальної освіти через експертизу викладачем джерела такої освіти.

8. Рекомендована література

1. Майборода Р. Комп'ютерна статистика - професійний старт (з використанням R) (електронний ресурс)
2. Осипчук М.М. Лекції із статистичного аналізу та програмного середовища R (електронний ресурс)
3. Baayen R.H. Analyzing Linguistic Data: A Practical Introduction to Statistics using R. Cambridge University Press, 2008.
4. Bivand R.S., Pebesma E., Gómez-Rubio V. Applied Spatial Data Analysis with R. Springer, 2013.

Викладач Осипчук М.М., професор кафедри математичного і функціонального аналізу