

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

Факультет радіофізики, електроніки та комп'ютерних систем

Кафедра математики та теоретичної радіофізики

«ЗАТВЕРДЖУЮ»

Заступник декана з навчальної роботи

_____ О. Ю. Нечипорук

« ____ » _____ 2021 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Статистичний аналіз даних

для студентів

галузь знань	10 “Природничі науки”
спеціальність	105 “Прикладна фізика і наноматеріали”
рівень вищої освіти	бакалавр
освітньо-наукова програма	“Еконофізика”
вид дисципліни	обов’язкова

Форма навчання	денна
Навчальний рік	2021/2022
Семестр	7
Кількість кредитів ECTS	4
Мова викладання	українська
Форма заключного контролю	залік

Викладач:

Іваненко Дмитро Олександрович, кандидат фіз.-мат. наук, доцент кафедри математики та теоретичної радіофізики

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

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

Розробник:

Іваненко Дмитро Олександрович, кандидат фіз.-мат. наук, доцент кафедри математики та теоретичної радіофізики

«ЗАТВЕРДЖЕНО»

Завідувач кафедри математики та теоретичної
радіофізики

_____ В.І. Висоцький

Протокол № __ від « __ » _____ 2021 р.

Схвалено науково-методичною комісією факультету радіофізики, електроніки та комп'ютерних систем

Протокол № __ від « __ » _____ 2021 р.

Голова науково-методичної комісії

С. П. Радченко

« __ » _____ 2021 року.

ВСТУП

1. Мета дисципліни – ознайомлення студентів зі статистичними методами і застосування їх до економічних і фізичних даних з метою аналізу, оцінювання параметрів і характеристик, моделювання.

2. Попередні вимоги до опанування або вибору навчальної дисципліни:

Навчальна дисципліна “Статистичний аналіз даних” базується на циклі дисциплін професійної та практичної підготовки бакалавра, а саме: “Теорія ймовірностей”, “Прикладна математична статистика”, “Статистична радіофізика”, “Об’єктно-орієнтовне програмування”.

Попередні вимоги:

студент повинен знати: основи теорії ймовірностей і математичної статистики, основи програмування.

студент повинен вміти: оперувати з випадковими величинами і процесами, застосовувати статистичні методи на практиці, програмувати.

3. Анотація навчальної дисципліни:

Дисципліна “Статистичний аналіз даних” дозволяє студенту зорієнтуватись в сучасних проблемах прикладної статистики і вивчити програмну систему, яка є інструментом розв’язку таких задач. Як відомо, статистичні методи є найбільш широкоживаним математичним інструментом для фахівців в галузі економіки.

Курс “ Статистичний аналіз даних ” складається з розділів «Базові поняття мови Python», «Авторегресійні моделі», «Моделі ковзного середнього», «SARIMA моделі», «Ланцюги Маркова», «Дифузійні процеси». Кожен з цих розділів дозволяє вивчити певні статистичні методи аналізу даних і з теоретичної точки зору, і з практичної, на базі програмного середовища (мови програмування) Python. Це дуже потужна і популярна мова програмування, яка поширюється безкоштовно в світі і швидко розбудовується користувачами.

4. Завдання навчальної дисципліни (навчальні цілі):

1. Узагальнити та розширити знання в області математичної і прикладної статистики, в застосуванні методів регресійного аналізу часових рядів, моделей, що задаються лінійними стохастичними диференціальними рівняннями дифузійного типу тощо. Засвоїти базові поняття з теорії і методів аналізу марківських процесів. Вивчити основні підходи в теорії стохастичного моделювання.

2. Навчити використовувати мову Python в аналізі даних

3. Навчити застосовувати отримані знання та уміння на практиці і в моделюванні економічних і фізичних процесів.

5. Результати навчання за дисципліною:

Результат навчання (1, знати; 2, вміти; 3, комунікація; 4, автономність та відповідальність)		Форми (та/або методи і технології) викладання і навчання	Методи оцінювання та пороговий критерій оцінювання (за необхідності)	Відсоток у підсумковій оцінці з дисципліни
Код	Результат навчання			
1	знати:	лекційні заняття, практичні заняття	Лабораторні роботи	до 45
1.1	основні оператори і підходи мови Python, типи даних, візуалізацію даних	лекційні заняття, практичні заняття	Лабораторні роботи	
1.2	робота з модулем stats, пакетами statsmodels і scipy. Інструменти модуля PANDAS	лекційні заняття, практичні заняття	Лабораторні роботи	
1.3	Модель авторегресії AR. Рівняння для OMB, властивості оцінок	лекційні заняття, практичні заняття	Лабораторні роботи	
1.4	модель ковзаючого середнього MA	лекційні заняття, практичні заняття	Лабораторні роботи	
1.5	модель авторегресії з залишками у формі ковзаючого середнього. Формули для коефіцієнтів	лекційні заняття, практичні заняття	Лабораторні роботи	
1.6	Перевірка гіпотез про порядок AR, MA, ARIMA, критерій Фостера-Стюарта, критерій серій, підбір форми тренду.	лекційні заняття, практичні заняття	Лабораторні роботи	
1.7	Інформаційний критерій Акаїке, інші інформаційні критерії, перевірка на	лекційні заняття, практичні заняття	Лабораторні роботи	

	стаціонарність, ACF, PACF, тест Дікі-Фулера.			
2	вміти:	лекційні заняття, практичні заняття	Лабораторні роботи	до 45
2.1	програмувати на мові Python, вміти написати код для реалізації основних статистичних методів аналізу даних і їх візуалізації	лекційні заняття, практичні заняття	Лабораторні роботи	
2.2	застосовувати на практиці побудову гістограм, базисного та ланцюгового темпів росту	лекційні заняття, практичні заняття	Лабораторні роботи	
2.3	застосувати критерій Фостера-Стюарта, критерій серій	лекційні заняття, практичні заняття	Лабораторні роботи	
2.4	дослідити лінійну і поліноміальну моделі (навести відповідні графіки і статистики)	лекційні заняття, практичні заняття	Лабораторні роботи	
2.5	підібрати параметри ARMA процесу за допомогою критерію АКАЙКЕ	лекційні заняття, практичні заняття	Лабораторні роботи	
2.6	побудувати ACF, PACF, застосувати тест Дікі-Фулера, підібрати параметри ARIMA	лекційні заняття, практичні заняття	Лабораторні роботи	
2.7	досліджувати часові ряди, здійснювати їх декомпозицію і прогнозування.	лекційні заняття, практичні заняття	Лабораторні роботи	
2.8	складати дизайн-план вибіркового обстеження, вміти застосовувати інтерпретувати модель SARIMA.	лекційні заняття, практичні заняття	Лабораторні роботи	
3	комунікація:	лекційні заняття, практичні заняття	Лабораторні роботи	до 5
3.1	Здатність грамотно будувати наукову комунікацію як в усній так і письмовій формах, підбирати правильну термінологію	лекційні заняття, практичні заняття	Лабораторні роботи	
4	автономність та відповідальність:	лекційні заняття, практичні заняття	Лабораторні роботи	до 5
4.1	продемонструвати розуміння особистої/персональної відповідальності за професійні та/або управлінські рішення, які базуються на використанні статистичних методів	лекційні заняття, практичні заняття	Лабораторні роботи	

6. Співвідношення результатів навчання дисципліни із програмними результатами навчання

Результати навчання дисципліни (код)	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	2.1	2.2	2.3	2.4	2.5	3.1	4.1
Програмні результати навчання (назва)															
знання															
Відтворювати історичний розвиток математичних знань і парадигм, знати сучасні тенденції в математиці	+		+						+						
Володіти основами правових, етичних відносин і психологічних особливостей поведінки															
Знати аксіоми різних складових частин математики, принципи modus ponens (правило виведення логічних висловлювань) та modus tollens (доведення від супротивного) і використовувати умови, формулювання, висновки, доведення та наслідки математичних тверджень у різних складових частинах Відтворювати базові знання фундаментальних розділів математики в обсязі, необхідному для володіння математичним апаратом відповідної галузі знань і використання математичних методів у обраній професії математики		+		+	+	+	+	+							
Володіти базовими знаннями з розділів диференціальних рівнянь у обсязі, необхідному для засвоєння загальнопрофесійних дисциплін; володіти навичками використання програмних засобів і роботи в комп'ютерних мережах, умінням використовувати інтернет-ресурси	+		+						+						
уміння															
Пояснювати математичні концепції мовою, зрозумілою для нефакхівців у галузі математики															
Усно й письмово спілкуватися рідною мовою із професійних питань; читати спеціальну															

літературу іноземною мовою; знаходити, аналізувати та використовувати інформацію з різних довідкових джерел															
Дотримуватися норм етичної поведінки стосовно інших людей, адаптуватися та комунікувати															
Бути наполегливим у досягненні мети під час вирішення математичної проблеми															
Розв'язувати задачі з математичною строгістю та математичними методами, перевіряти умови виконання математичних тверджень, переносити умови та твердження на нові класи об'єктів, знаходити й аналізувати відповідності між поставленою задачею й існуючими моделями										+	+	+	+	+	+
Розв'язувати конкретні математичні задачі, які сформульовано через терміни відповідної предметної області; здійснювати базові перетворення математичних моделей з метою розв'язування математичних та/або прикладних задач										+	+	+	+	+	+
Використовувати раціональні способи пошуку та використання науково-технічної інформації, включаючи засоби електронних інформаційних мереж; застосовувати інформаційні ресурси, у тому числі електронні, для пошуку існуючих математичних моделей															

7. Схема формування оцінки

7.1. Форми оцінювання

Рівень досягнення всіх запланованих результатів навчання визначається за результатами написання письмових контрольних робіт. Внесок результатів навчання у підсумкову оцінку, за умови їх опанування на належному рівні:

- результати навчання 1.1 – 1.10 [знання] – до 45 %;
- результат навчання 2.1 – 2.3 [вміння] – до 45%;
- результат навчання 3.1-3.2 [комунікація] – до 5%;
- результат навчання 4.1 [автономність та відповідальність] – до 5%;

Форми оцінювання:

- **семестрове оцінювання:** Навчальний семестр має один змістовний модуль. Студент повинний виконати і здати чотири лабораторні роботи. Обов'язковим для допуску до іспиту є: набрати на протязі семестру не менше 36 балів.
- **підсумкове оцінювання (у формі заліку):** форма заліку – письмово-усна. Завдання на іспит складається з 2 питань і двох задач, питання оцінюються по 10 балів. Всього за залік можна отримати від 0 до 40 балів. Умовою досягнення позитивної оцінки за дисципліну є отримання не менш ніж 60 балів, оцінка за іспит не може бути меншою **24 бали**.
- **умови допуску до підсумкового заліку:** умовою допуску до заліку є отримання студентом сумарно не менше, ніж *критично-розрахунковий мінімум* за семестр. студенти, які протягом семестру сумарно набрали меншу кількість балів, ніж критично-розрахунковий мінімум **36 балів**, для одержання допуску до заліку обов'язково повинні написати додаткову контрольну роботу.

У випадку відсутності студента з поважних причин відпрацювання та перездачі лабораторних контрольних робіт здійснюються у відповідності до „Положення про організацію освітнього процесу у Київському національному університеті”

7.2. Організація оцінювання;

Оцінювання за формами контролю:

Семестрова робота	Кількість балів	
	Min. – 36	Max. – 60
Лабораторна робота 1	9	15
Лабораторна робота 2	9	15
Лабораторна робота 3	9	15
Лабораторна робота 4	9	15

Орієнтований графік оцінювання:

Форма оцінювання	Орієнтовний період для здійснення відповідної форми оцінювання
Лабораторна робота 1	вересень
Лабораторна робота 2	вересень
Лабораторна робота 3	жовтень
Лабораторна робота 4	жовтень
Добір балів/додаткова контрольна робота	листопад
Залік	грудень

Розрахунок балів, які отримують при успішній здачі іспиту:

Значення	Змістовний модуль	Залік	Підсумкова оцінка
Мінімум	36	24	60

Максимум	60	40	100
----------	----	----	-----

7.3. Шкала відповідності оцінок

Оцінка (за національною шкалою) / National grade	Рівень досягнень, % / Marks, %
Відмінно / Excellent	90-100%
Добре / Good	75-89%
Задовільно / Satisfactory	60-74%
Незадовільно / Fail	0-59%

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

№ з/п	Назва теми	Кількість годин		
		Лекції	Практичні	Самостійна робота
1	Задачі математичної статистики. Характеристики оцінки: конзистентність, незміщеність, ефективність, робастність.	1	1	2
2	Аналіз часових рядів. Дослідження на стаціонарність лишків.	1	1	2
3	Базові поняття мови програмування Python. Найпростіші операції в мові Python.	1	1	2
4	Імпортування даних в Python	1	1	2
5	Типи даних в Python: вектори, списки, матриці, фрейми.	1	1	2
6	Робота з розподілами: функції розподілу, щільності, квантильна функція, моделювання різних розподілів.	1	1	2
7	Базова графіка в Python. Модуль PANDAS.	2	2	4
8	Моделі AR, MA, ARIMA, SARIMA .	2	2	4
9	Тести і критерії для вибору моделі і кількості параметрів.	2	2	4
10	Ланцюги Маркова зі скінченним простором станів.	2	2	4
11	Процеси з незалежними приростами. Процес Вінера. Процес Пуассона.	1	1	2
12	Загальні відомості зі стохастичного інтегрування.	1	1	2
13	Інтеграл Іто і Стратоновича.	1	1	2

14	Відомості про процеси, задані стохастичними диференціальними рівняннями дифузійного типу	1	1	2
Всього		18	18	36

Загальний обсяг **72** год., в тому числі:

Лекції **18** год.

Лабораторні **18** год.

Самостійна робота **36** год.

9. Рекомендована література:

Основні джерела:

1. Майборода Р.Є. «Комп'ютерна статистика.» Підручник. - К., ВПЦ «Київський університет», 2019.

<https://probability.knu.ua/userfiles/mre/cscolor.pdf>

2. Майборода Р.Є., Сугакова О.В. «Аналіз даних за допомогою пакету R». Навчальний посібник – К., 2015.

http://matphys.rpd.univ.kiev.ua/wp/wp-content/uploads/2016/12/Statistics_with_R.pdf

3. Т. Андерсон. «Статистика часових рядов» – Boca Raton, London, New York: CRC Press, Taylor&Francis Group, 2008. – 700 p.

4. Черняк О.І., Комашко О.В., Ставицький А.В., Баженова О.В. «Економетрика» - ВПЦ «Київський університет», 2009.

5. Василик О.І., Яковенко Т.О. «Лекції з теорії і методів вибірових обстежень» - К., ВПЦ «Київський університет», 2010 – 208 с.

6. ARIMA models

<https://pyflux.readthedocs.io/en/latest/arima.html>

7. Турчин В.Н., Турчин Е.В. «Марковские цепи основные понятия, примеры, задачи» - Дніпро, Lizunoffpress, 2016 – 192с.

https://mst-mmfdnu.dp.ua/download/books/TurchinVN_TurchinEV_MC_2016.pdf

8. Гихман И.И., Скороход А.В. «Стохастические дифференциальные уравнения.» К:, издательство «Науковадумка», 1968 – 354 с.

9. Мішура Ю.С., Ральченко , Шевченко Г.М. «Випадкові процеси. Теорія. Статистика. Застосування. » - К., ВПЦ «Київський університет», 2021 – 496 с.

<https://probability.knu.ua/userfiles/mre/cscolor.pdf>