

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Центральноукраїнський державний педагогічний університет імені Володимира Винниченка

Факультет математики, природничих наук та технологій
Кафедра математики та методики її навчання

«ЗАТВЕРДЖУЮ»

В.о. завідувача кафедри

 Ботузова Ю.В.

«4» серпня 2022 року



РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

ОК 4 Аналіз даних та комп'ютерна статистика

Рівень вищої освіти: другий (магістерський)

Галузь знань: 11 Математика та статистика

Спеціальність: 112 Статистика

Освітня програма: Статистика (Фінансова, страхова та комп'ютерна статистика)

Форма навчання: денна

2022 – 2023 навчальний рік

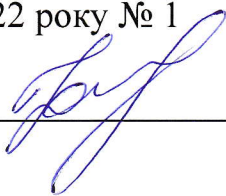
Робоча програма навчальної дисципліни Аналіз даних та комп'ютерна статистика розроблена на основі освітньо-професійної програми Статистика (Фінансова, страхова та комп'ютерна статистика), навчального плану підготовки здобувачів вищої освіти освітнього ступеня магістр за спеціальністю 112 Статистика

Розробник: Акбаш Катерина Сергіївна, доцент кафедри математики та методики її навчання, кандидат фізико-математичних наук

Робочу програму обговорено та схвалено на засіданні кафедри математики та методики її навчання

Протокол від «4» серпня 2022 року № 1

В.о. завідувача кафедри _____ д.пед.н., доцент Ботузова Ю.В.



Робоча програма навчальної дисципліни **Аналіз даних та комп'ютерна статистика** для студентів спеціальності 112 Статистика за другим (магістерським) рівнем вищої освіти. – ЦДПУ імені В. Винниченка, 2022.

© Акбаш К.С., 2022 рік

© ЦДПУ імені В.Винниченка, 2022 рік

1. ПОЯСНЮВАЛЬНА ЗАПИСКА

1.1. Опис навчальної дисципліни

Найменування показників	Галузь знань, спеціальність, рівень вищої освіти	Характеристика навчальної дисципліни	
		денна форма навчання	заочна форма навчання
Кількість кредитів – 5	Галузь знань <u>11 Математика та статистика</u> (шифр і назва)	Нормативна	
Індивідуальне навчально-дослідне завдання (назва)	Спеціальність: <u>112 Статистика</u> (шифр і назва)	Рік підготовки	
		1-й	-й
Загальна кількість годин – 150	Освітня програма: <u>Фінансова, страхова та комп'ютерна статистика</u> (шифр і назва)	Семестр	
		1-й	-й
Тижневих годин для денної форми навчання: аудиторних – 4 самостійної роботи студента – 4	Рівень вищої освіти: <u>другий (магістерський)</u>	Лекції	
		26 год.	год.
		Практичні, семінарські	
		16 год.	год.
		Лабораторні	
		32 год.	год.
		Самостійна робота	
		76 год.	год.
		Індивідуальні завдання:	
		год.	
Вид контролю:			
Екзамен			

1.2. Мета та завдання навчальної дисципліни

Метою викладання дисципліни є формування теоретичних і практичних компетенцій в області цілісного представлення, розуміння місця і ролі, а також застосування методів статистичного аналізу даних; навчання студентів сучасних програмних засобів, в яких реалізовані модулі, які здійснюють розв'язування задач аналізу даних.

Завданнями вивчення навчальної дисципліни є:

- Дати студентам теоретичні основи по спектру найбільш розповсюджених статистичних методів аналізу даних і умов їх експерименту;
- Привити навиків і умінь практичного застосування комп'ютерних технологій при аналізі і прогнозуванні соціально-економічних показників;
- Вивчення концепції та технології сучасного аналізу даних на комп'ютері;

- Вивчення сучасних візуальних методів аналізу даних і використання їх для статистичних висновків і формування гіпотез про структуру даних;
- Вироблення вмінь самостійного розв'язування задач по вибору методів аналізу в практичних ситуаціях.

У результаті вивчення навчальної дисципліни у студента мають бути сформовані такі *компетентності*:

загальні	фахові
ЗК3. Здатність проведення теоретичних та прикладних досліджень на відповідному рівні. ЗК4. Здатність вчитися і оволодівати сучасними знаннями. ЗК6. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.	ФК1. Здатність аналізувати предметні області, формувати, аналізувати та моделювати процеси їх функціонування. ФК5 Спроможність розробляти ймовірно-статистичні моделі ситуацій з реального світу та переносити математичні знання у нематематичні контексти. ФК7. Здатність самостійно розробляти проекти шляхом творчого застосування існуючих та генерування нових математичних та статистичних підходів. ФК14. Здатність систематизувати професійні знання, працювати з експертними, онлайн- та оффлайн-джерелами інформації для інтеграції даних і знань, з метою поглиблення набутих та здобуття нових фахових знань.

1.3. Очікувані програмні результати навчання:

ПРН1. Знання методології аналізу та моделювання прикладної предметної області, виявлення інформаційних потреб і збору вихідних даних для статистичного дослідження процесів її функціонування.

ПРН2. Знати та розуміти фундаментальні і прикладні аспекти наук у сфері математики та статистики.

ПРН5. Володіння статистичними методами аналізу, прогнозування та оцінки параметрів моделей, способами інтерпретації числових даних та принципами функціонування процесів реального світу.

ПРН8. Відтворювати знання фундаментальних розділів математики в обсязі, необхідному для володіння математичним апаратом відповідної галузі знань і використання ймовірно-статичних методів у обраній професії.

ПРН12. Застосовувати методології обґрунтованого вибору парадигм і спеціалізоване програмне забезпечення для вирішення прикладних статистичних завдань.

ПРН15. Здійснювати науково-дослідну роботу в сферах фінансової, страхової та комп'ютерної статистики.

ПРН20. Здатність працювати з експертними, онлайн- та оффлайн-джерелами інформації для інтеграції даних і знань, реалізовувати раціональні способи пошуку та використання науково-технічної інформації для розв'язання прикладних професійних завдань.

ПРН22. Здатність формулювати та вдосконалювати дослідницьку задачу, збирати необхідну інформацію та формулювати висновки, які можна захищати в науковому контексті.

2. ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Змістовий модуль 1. Методи аналізу законів розподілів ймовірностей випадкових величин

Тема 1. Критерії нормальності розподілу

Загальні критерії згоди, модифіковані для перевірки нормальності розподілу (модифікований критерій χ^2 , Критерії типу Колмогорова-Смирнова). Спеціальні критерії нормальності (Критерій Шапіро-Уїлка, Критерій асиметрії та ексцесу). Порівняння потужностей різних критеріїв нормальності.

Тема 2. Критерії експоненційності та рівномірності розподілу

Критерій Шапіро-Уїлка, Критерії типу Колмогорова-Смирнова. Критерій Манн-Фертінга-Шуера для розподілу Вейбула.

Тема 3. Криві розподілу Джонсона

Сімейство розподілів S_L . Сімейство розподілів S_B . Сімейство розподілів S_U . Застосування кривих Джонсона при обробці експериментальних даних.

Тема 4. Криві розподілу Пірсона

Криві Пірсона типів I–VII. Застосування кривих Пірсона при обробці експериментальних даних.

Змістовий модуль 2. Критерії тренду і випадковості

Тема 5. Критерії викидів у випадку нормального розподілу

Критерій Шовене. Критерій Грубса. Критерій Титьєна-Мура для виявлення декількох викидів.

Тема 6. Критерії викидів у випадку експоненційного розподілу та розподілу Вейбула

Критерій Смоляка-Титаренко. Критерій Кімбера для декількох викидів. Критерій викидів для експоненційного розподілу.

Змістовий модуль 3. Багатовимірний статистичний аналіз

Тема 7. Кластерний аналіз і його реалізація в статистичних пакетах

Ієрархічний кластерний аналіз з двома та більше змінними. Ієрархічний кластерний аналіз з попереднім факторним аналізом.

Тема 8. Дискримінантний аналіз і його реалізація в статистичних пакетах

Лінійний дискримінантний кластерний аналіз. Дискримінантний аналіз при нормальному законі розподілу показників

Тема 9. Факторний аналіз і його реалізація в статистичних пакетах

Основні проблеми факторного аналізу. Метод головних компонент і його алгоритм. Проблема обертання. Проблема оцінки факторів та задача класифікації.

Змістовий модуль 4. Аналіз часових рядів

Тема 10. Стаціонарні та звідні до них часові ряди

Білий шум. $MA(q)$ процес. Процес авторегресії. Перетворення $MA(q)$ процесів. $ARMA$ процес. $ARIMA$ процеси. Аналіз та прогнозування часових рядів у статистичному пакеті SPSS

Тема 11. Новітні моделі та методи аналізу часових рядів

Лінійні моделі. Нелінійні моделі. Нейронні мережі. Векторні часові ряди.

2. СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Назви розділів і тем	Кількість годин					
	денна форма					
	усього	у тому числі				
л		п	лаб.	інд	СРС	
1	2	3	4	5	6	7
Змістовий модуль 1. Методи аналізу законів розподілів ймовірностей випадкових величин						
Тема 1. Критерії нормальності розподілу	9	2		2		5
Тема 2. Критерії експоненційності та рівномірності розподілу	9	2		2		5
Тема 3. Криві розподілу Джонсона	11	2	2	2		5
Тема 4. Криві розподілу Пірсона	11	2	2	2		5
<i>Разом за Змістовним модулем 1</i>	<i>40</i>	<i>8</i>	<i>4</i>	<i>8</i>		<i>20</i>
Розділ 2. Двоетапні задачі стохастичного програмування						
Тема 5. Критерії викидів у випадку нормального розподілу	11	2	2	2		5
Тема 6. Критерії викидів у випадку експоненційного розподілу та розподілу Вейбула	11	2	2	2		5
<i>Разом за Змістовним модулем 2</i>	<i>22</i>	<i>4</i>	<i>4</i>	<i>4</i>		<i>10</i>
Змістовий модуль 3. Багатовимірний статистичний аналіз						

Тема 7. Кластерний аналіз і його реалізація в статистичних пакетах	16	2	2	4		8
Тема 8. Дискримінантний аналіз і його реалізація в статистичних пакетах	12	2		4		6
Тема 9. Факторний аналіз і його реалізація в статистичних пакетах	16	2	2	4		8
<i>Разом за Змістовним модулем 3</i>	<i>44</i>	<i>6</i>	<i>4</i>	<i>12</i>		<i>22</i>
Змістовий модуль 4. Аналіз часових рядів						
Тема 10. Стаціонарні та звідні до них часові ряди	22	4	2	4		12
Тема 11. Новітні моделі та методи аналізу часових рядів	22	4	2	4		12
<i>Разом за Змістовним модулем 4</i>	<i>44</i>	<i>8</i>	<i>4</i>	<i>8</i>		<i>24</i>
Усього годин	150	26	16	32		76

4. ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

4.1. Теми лекційних занять

4.1.1 денна форма навчання

№ з/п	Назва теми	Кількість годин
1	Критерії нормальності розподілу	2
2	Критерії експоненційності та рівномірності розподілу	2
3	Криві розподілу Джонсона	2
4	Криві розподілу Пірсона	2
5	Критерії викидів у випадку нормального розподілу	2
6	Критерії викидів у випадку експоненційного розподілу та розподілу Вейбула	2
7	Кластерний аналіз і його реалізація в статистичних пакетах	2
8	Дискримінантний аналіз і його реалізація в статистичних пакетах	2
9	Факторний аналіз і його реалізація в статистичних пакетах	2
10	Стаціонарні та звідні до них часові ряди	4
11	Новітні моделі та методи аналізу часових рядів	4
	Разом	26

4.2. Теми семінарських (практичних) занять

4.2.1 денна форма навчання

№ з/п	Назва теми	Кількість годин
1	Підбір кривих Джонсона за експериментальними даними	2
2	Підбір кривих Пірсона за експериментальними даними	2
3	Критерії викидів у випадку нормального розподілу	2
4	Критерії викидів у випадку експоненційного розподілу та розподілу Вейбула	2
5	Методи кластерного аналізу	2
5	Методи факторного аналізу	2
7	Первинний аналіз часового ряду та вибір методів обробки	2
8	Нелінійні моделі часових рядів	2
	Разом	16

4.3. Теми лабораторних занять

4.3.1 денна форма навчання

№ з/п	Назва теми	Кількість годин
1	Застосування критеріїв нормальності розподілу на реальних даних. Можливості статистичного пакету STATGRAPHICS.	2
2	Застосування критеріїв експоненційності та рівномірності розподілу на реальних даних. Можливості статистичного пакету STATGRAPHICS.	2

3	Підбір кривих Джонсона за експериментальними даними (сімейства S_L, S_B, S_U).	2
4	Підбір кривих Пірсона за експериментальними даними (сімейства I-VII).	2
5	Аналіз викидів у випадку нормального розподілу за експериментальними даними. Можливості статистичного пакету SPSS	2
6	Аналіз викидів у випадку експоненційного розподілу та розподілу Вейбула за експериментальними даними	2
7	Кластерний аналіз і його реалізація в статистичних пакетах	4
8	Дискримінантний аналіз і його реалізація в статистичних пакетах	4
9	Факторний аналіз і його реалізація в статистичних пакетах	4
10	Побудова та аналіз ARIMA моделей в SPSS. Прогнозування на основі ARIMA моделей	4
11	Нелінійні моделі та методи аналізу часових рядів	4
	Разом	32

4.4. Завдання для самостійної роботи

4.4.1 денна форма навчання

№ з/п	Назва теми	Кількість годин
1	Загальні критерії згоди (критерії, що базуються на порівнянні щільності розподілу і емпіричної гистограми; критерії, що базуються на порівнянні теоретичної і емпіричної функції розподілу ймовірностей)	5
2	Критерії, що базуються на трансформації експоненційного розподілу у рівномірний	5
3	Криві розподілу Джонсона	5
4	Криві розподілу Пірсона	5
5	Критерії викидів у випадку нормального розподілу	5
6	Критерії викидів у випадку експоненційного розподілу та розподілу Вейбула	5
7	Кластерний аналіз методом k-середніх	8
8	Порівняння дискримінантної функції і моделі логістичної регресії	6
9	Методи обертання у задачі факторного аналізу	8
10	Стационарні та звідні до них часові ряди	12
11	Новітні моделі та методи аналізу часових рядів	12
	Разом	76

4.5. Індивідуальне навчально-дослідне завдання

Не передбачено.

4.6. Методи навчання

Під час вивчення дисципліни «Аналіз даних та комп'ютерна статистика» передбачено комплексне використання різноманітних методів організації і здійснення навчально-пізнавальної діяльності студентів та методів стимулювання і мотивації їх навчання, що сприяє розвитку творчих засад особистості майбутнього фахівця, з урахуванням індивідуальних особливостей учасників освітнього процесу.

З метою формування професійних компетентностей широко впроваджуються інноваційні методи навчання. Це – комп'ютерна підтримка освітнього процесу, впровадження інтерактивних методів навчання.

Бали студенти накопичують на основі виконаних завдань у практичних і лабораторних роботах. Одне комплексне завдання розраховане на кожну тему. Оцінювання відбувається на основі зданих коротких аналітичних звітів по кожному завданню курсу.

4.7. Засоби діагностики результатів навчання здобувачів освіти.

Порядок та критерії виставлення балів

Навчальна дисципліна «Аналіз даних та комп'ютерна статистика» оцінюється за модульно–рейтинговою системою. Дисципліна викладається один семестр. Складається з чотирьох змістовних модулів: до першого входить 4 теми, до другого 2 теми, до третього 3 теми, до четвертого дві теми.

Контрольні заходи здійснюються з дотриманням вимог об'єктивності, індивідуального підходу, системності, всебічності.

Використовуються такі методи контролю (усний, письмовий), які мають сприяти підвищенню мотивації студентів до навчально-пізнавальної діяльності, зокрема, поточний контроль та підсумковий контроль.

Завданням поточного контролю є перевірка розуміння та засвоєння певної частини учбового матеріалу, рівня сформованості навичок, умінь самостійно опрацьовувати навчальний матеріал, здатності осмислити зміст теми.

Об'єктами поточного контролю знань студента є систематичність та активність роботи на заняттях; виконання завдань для самостійної роботи. Оцінюванню можуть підлягати: рівень знань, продемонстрований у комплексних аналітичних звітах; активність при обговоренні питань практичного заняття.

У разі невиконання завдань поточного контролю студент має право скласти їх індивідуально до останнього практичного заняття за дозволом завідувача кафедри. Порядок такого контролю регламентований викладачем.

Завданням підсумкового контролю є перевірка розуміння студентом програмного матеріалу в цілому, логіки та взаємозв'язків між окремими розділами, здатності творчого використання накопичених знань, уміння сформулювати своє ставлення до певної проблеми навчальної дисципліни тощо.

4.8. Перелік програмових питань для самоконтролю:

1. Критерії нормальності розподілу: модифікований критерій Хі-квадрат
2. Критерії нормальності розподілу: критерії типу Колмогорова-Смирнова
3. Критерії перевірки експоненційності розподілу: критерій Шапіро-Уїлка
4. Критерії перевірки експоненційності розподілу: критерії типу Колмогорова-

Смирнова

5. Критерії згоди для рівномірного розподілу
6. Критерії згоди для розподілу Пуасона
7. Критерії згоди для Гамма-розподілу
8. Підбір кривих розподілу ймовірностей по експериментальним даним: криві розподілу Джонсона типу S_B
9. Підбір кривих розподілу ймовірностей по експериментальним даним: криві розподілу Джонсона типу S_L
10. Підбір кривих розподілу ймовірностей по експериментальним даним: криві розподілу Джонсона типу S_U
11. Оцінка параметрів нормального розподілу Підбір кривих розподілу ймовірностей по експериментальним даним: криві розподілу Пірсона типу I
12. Оцінка параметрів нормального розподілу Підбір кривих розподілу ймовірностей по експериментальним даним: криві розподілу Пірсона типу II
13. Оцінка параметрів нормального розподілу Підбір кривих розподілу ймовірностей по експериментальним даним: криві розподілу Пірсона типу III
14. Оцінка параметрів нормального розподілу Підбір кривих розподілу ймовірностей по експериментальним даним: криві розподілу Пірсона типу IV
15. Оцінка параметрів нормального розподілу Підбір кривих розподілу ймовірностей по експериментальним даним: криві розподілу Пірсона типу V.
16. Оцінка параметрів нормального розподілу Підбір кривих розподілу ймовірностей по експериментальним даним: криві розподілу Пірсона типу VI
17. Оцінка параметрів нормального розподілу Підбір кривих розподілу ймовірностей по експериментальним даним: криві розподілу Пірсона типу VII
18. Критерії перевірки викидів у випадку нормального розподілу. Критерій Грубса
19. Критерії перевірки викидів у випадку нормального розподілу. Критерій Титьєна Мура для виявлення декількох викидів
20. Критерії перевірки викидів у випадку експоненційного розподілу
21. Ієрархічний кластерний аналіз з двома та більше змінними
22. Ієрархічний кластерний аналіз з попереднім факторним аналізом
23. Ієрархічний кластерний аналіз з попереднім факторним аналізом
24. Дискримінантний аналіз при нормальному законі розподілу показників
25. Факторний аналіз. Метод головних компонент і його алгоритм
26. Білий шум
27. $MA(q)$ процес
28. Процес авторегресії
29. ARMA процес
30. ARIMA процеси
31. Нелінійні моделі часових рядів

4.9. Схема нарахування балів, які отримують студенти

Екзамен (I семестр)

Поточне тестування та самостійна робота		Разом	Екзамен	Сума
Розділ 2			Макс. - 40	
T1-T6	T7-T11	60	40	100
5*6=30	6*6=30			

T1, T2 ... T11 – теми курсу.

Шкала оцінювання: національна та ЄКТС

Сума балів за всі види навчальної діяльності	Оцінка за національною шкалою	
	для екзамену, курсового проєкту (роботи), практики	для заліку
90-100	відмінно	зараховано
82-89	добре	
74-81		
64-73	задовільно	
60-63		
35-59	незадовільно	незараховано
1-34	незадовільно	незараховано

5. НАВЧАЛЬНО-МЕТОДИЧНЕ ЗАБЕЗПЕЧЕННЯ

5.1. Рекомендована література

Основна

1. Бахрушин В.Є. Методи аналізу даних : навчальний посібник для студентів / В.Є. Бахрушин. – Запоріжжя : КПУ, 2011. – 268 с. ISBN 978-966-414-103-8
2. Горват А.А., Молнар О.О., Мінькович В.В. Методи обробки експериментальних даних з використанням MS Excel: Навчальний посібник. Ужгород: Видавництво УжНУ "Говерла", 2019. – 160 с.
3. Грицюк П.М., Остапчук О.П. Аналіз даних: Навчальний посібник.– Рівне: НУВГП, 2008. – 218 с.
4. Ланде Д.В., Субач І.Ю., Бояринова Ю.Є. Основи теорії і практики інтелектуального аналізу даних у сфері кібербезпеки: навчальний посібник. — К.: ІСЗІ КПІ ім. Ігоря Сікорського», 2018. — 300 с.
5. Лук'яненко І. Г., Жук В. М. Аналіз часових рядів. Частина перша : Побудова ARIMA, ARCH/GARCH моделей з використанням пакета E.Views 6.0. Практичний посібник для роботи в комп'ютерному класі / І. Г. Лук'яненко, В. М. Жук. – К. : НаУКМА; Аграр Медіа Груп, 2013. – 187 с.
6. Лупан І.В., Авраменко О.В., Акбаш К.С. Комп'ютерні статистичні пакети: навчально-методичний посібник. – 2-е вид. Кіровоград: «КОД» 2015.
7. Майборода Р. Є. Комп'ютерна статистика : підручник / Р. Є. Майборода. – К. : ВПЦ "Київський університет", 2019. – 589 с. ISBN 978-966-933-010-9

8. Оленко А.Я. Комп'ютерна статистика: Навчальний посібник. – К.: Видавничо-поліграфічний центр «Київський університет», 2007. – 174 с.

9. Шкільний Є.П., Гончарова Л.Д., Миротворська Н.К. Методи обробки та аналізу гідрометеорологічної інформації (збірник задач і вправ): навчальний посібник. – К.: Міносвіти і науки України, 2000. – 419 с.

6. ПОЛІТИКА ЩОДО АКАДЕМІЧНОЇ ДОБРОЧЕСНОСТІ

Політика щодо академічної доброчесності формується на основі дотримання принципів академічної доброчесності відповідно до Законів України «Про освіту», «Про вищу освіту», «Про наукову і науково-технічну діяльність», «Про авторське право і суміжні права», «Про видавничу справу», з урахуванням норм Положення «Про академічну свободу та академічну доброчесність в Центральноту українському державному педагогічному університеті імені Володимира Винниченка» (затверджене вченою радою, протокол №2 від 30.09.2019; №10 від 07.02.2022).