

Міністерство освіти і науки України
Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»

Кафедра Комп'ютерних наук та інформаційних технологій (№ 302)

ЗАТВЕРДЖУЮ

Гарант освітньої програми

 О.В. Малєєва
(ініціали та прізвище)

« 29 » _____ 08 _____ 2023 р.

**РОБОЧА ПРОГРАМА ОБОВ'ЯЗКОВОЇ
НАВЧАЛЬНОЇ ДИСЦИПЛІНИ**

Дата-аналіз в інформаційних системах

(назва навчальної дисципліни)

Галузь знань: 12 «Інформаційні технології»
(шифр і найменування галузі знань)

Спеціальність: 126 - Інформаційні системи та технології
(код та найменування спеціальності)

Освітня програма: Розподілені інформаційні системи
(найменування освітньої програми)

Форма навчання: денна

Рівень вищої освіти: перший (бакалаврський)

Харків 2023 рік

Розробник: Малєєва О.В., професор, д.т.н.
(прізвище та ініціали, посада, науковий ступінь та вчене звання)


(підпис)

Робочу програму розглянуто на засіданні кафедри комп'ютерних наук та інформаційних технологій

Протокол № 659/09 від « 29 » _серпня_ 2023 р.

Завідувач кафедри д.т.н., проф.
(науковий ступінь і вчене звання)


(підпис)

О.Є. Федорович
(ініціали та прізвище)

1. Опис навчальної дисципліни

Найменування показників	Галузь знань, спеціальність, освітня програма, рівень вищої освіти	Характеристика навчальної дисципліни (денна форма навчання) 4- й семестр – обов’язкова 5- й семестр – обов’язкова
Кількість кредитів – семестр 4 – 7,5 семестр 5 – 2	Галузь знань 12 «Інформаційні технології» <small>(шифр та найменування)</small> Спеціальність: 126 - Інформаційні системи та технології <small>(код та найменування спеціальності)</small> Освітня програма: Розподілені інформаційні системи <small>(найменування освітньої програми)</small> Рівень вищої освіти: перший (бакалаврський)	Навчальний рік
Кількість модулів – 3		2023/2024
Кількість змістових модулів – 2		Семестр
Індивідуальне завдання РР “ <u>Дата-аналіз варіаційних рядів</u> ” <small>(назва)</small>		4-й, 5-й
Загальна кількість годин: семестр 4 – 225, семестр 5 – 76		Лекції*
Кількість тижневих годин для денної форми навчання: - аудиторних семестр 4 – 5 години, семестр 5 – 1 година; - самостійної роботи студента: семестр 4 – 9 годин, семестр 5 – 3 години.		4- й семестр – 32 години, 5- й семестр – 0 годин
		Практичні, семінарські*
		4- й семестр – 16 годин, 5- й семестр – 16 годин
	Лабораторні*	
	4- й семестр – 32 години, 5- й семестр – 0 годин	
	Самостійна робота	
	4- й семестр – 145 годин, 5- й семестр – 44 години	
	Вид контролю	
	4- й семестр – модульний контроль, іспит 5- й семестр – діф. залік	

Співвідношення кількості годин аудиторних занять до самостійної роботи становить: 80/145

* Аудиторне навантаження може бути зменшене або збільшене на одну годину залежно від розкладу занять.

2. Мета та завдання навчальної дисципліни

Мета вивчення: надання студентам основних положень теорії ймовірності та математичної статистики для проведення дата-аналізу стохастичних процесів та їх оцінювання.

Завдання: вивчення сучасних підходів теорії ймовірності для стохастичного дата-аналізу інформаційних систем.

Компетентності, які набуваються:

Загальні:

КЗ 1. Здатність до абстрактного мислення, аналізу та синтезу.

КЗ 2. Здатність застосовувати знання у практичних ситуаціях.

КЗ 4. Здатність спілкуватися іноземною мовою.

КЗ 5. Здатність вчитися і оволодівати сучасними знаннями.

КЗ 6. Здатність до пошуку, оброблення та узагальнення інформації з різних джерел.

Спеціальні (фахові):

КС 2. Здатність застосовувати стандарти в області інформаційних систем та технологій при розробці функціональних профілів, побудові та інтеграції систем, продуктів, сервісів і елементів інфраструктури організації.

КС 3. Здатність до проектування, розробки, налагодження та вдосконалення системного, комунікаційного та програмно-апаратного забезпечення інформаційних систем та технологій, Інтернету речей (IoT), комп'ютерно-інтегрованих систем та системної мережної структури, управління ними.

КС 4. Здатність проектувати, розробляти та використовувати засоби реалізації інформаційних систем, технологій та інфокомунікацій (методичні, інформаційні, алгоритмічні, технічні, програмні та інші).

КС 6. Здатність використовувати сучасні інформаційні системи та технології (виробничі, підтримки прийняття рішень, інтелектуального аналізу даних та інші), методики й техніки кібербезпеки під час виконання функціональних завдань та обов'язків.

КС 10. Здатність вибору, проектування, розгортання, інтегрування, управління, адміністрування та супроводжування інформаційних систем, технологій та інфокомунікацій, сервісів та інфраструктури організації.

КС 12. Здатність управляти та користуватися сучасними інформаційно-комунікаційними системами та технологіями (у тому числі такими, що базуються на використанні Інтернет).

КС 16. Здатність використовувати інтелектуальні інформаційні технології для прийняття управлінських рішень в складних системах та процесах (аерокосмічна галузь, атомна енергетика, критичні технологічні об'єкти промисловості тощо).

Очікувані результати навчання:

ПР 3. Використовувати базові знання інформатики й сучасних інформаційних систем та технологій, навички програмування, технології безпечної роботи в комп'ютерних мережах, методи створення баз даних та інтернет-ресурсів, технології розроблення алгоритмів і комп'ютерних програм мовами високого рівня із застосуванням об'єктно-орієнтованого програмування для розв'язання задач проектування і використання інформаційних систем та технологій.

ПР 6. Демонструвати знання сучасного рівня технологій інформаційних систем, практичні навички програмування та використання прикладних і спеціалізованих комп'ютерних систем та середовищ з метою їх запровадження у професійній діяльності.

Пререквізити – ОК 1, ОК 3, ОК 11 - Вища математика, ОК 3 - Моделі та методи дискретної математики, ОК 7 - Створення візуальних інтерфейсів, ОК 8 - Структуризація інформації в управлінні, ОК 13 - Мобільні та хмарні технології, ОК 14 - Веб-технології в розподілених інформаційних системах.

Кореквізити – ОК 22 - Методи дослідження та оптимізації бізнес-рішень, ОК 23 - Системне уявлення та інтеграція інформаційних систем, ОК 24 - Дата-аналіз в інформаційних системах (КР).

3. Програма навчальної дисципліни

Модуль 1.

Змістовний модуль 1. Основи теорії ймовірностей. Ряди розподілу випадкових величин.

Тема 1. Вступ до навчальної дисципліни «Дата-аналіз в інформаційних системах»

Історія розвитку теорії ймовірностей та математичної статистики. Предмет та методи дисципліни. Зв'язок з іншими дисциплінами. Стандарти в області інформаційних систем.

Тема 2. Основи теорії ймовірностей

Вступ до навчальної дисципліни. Випадкові події. Аксиоми теорії ймовірностей. Визначення імовірності. Геометрична ймовірність. Задачі з обчислення геометричної ймовірності. Умовна ймовірність. Незалежність подій. Формула повної імовірності. Формула Байєса. Схема Бернуллі, гранична теорема Пуассона. Локальна та інтегральна теореми Муавра – Лапласа.

Тема 3. Статистична сукупність. Ряди розподілу.

Предмет та метод дата-аналізу. Системне, комунікаційне та інформаційне забезпечення інформаційних систем. Інтелектуальний аналізу даних під час виконання функціональних завдань та обов'язків. Загальна, математична та спеціальні види статистики. Сукупність та її елементи. Варіаційні ряди. Графічне зображення варіаційних рядів. Функція і щільність розподілу

безперервних величин. Види кривих розподілу. Статистичні показники. Абсолютні величини. Відносні величини. Інформаційні технології дата-аналізу. Основні можливості програмного пакету STATISTICA. Інтерфейс користувача. Робота з файлами даних в пакеті STATISTICA.

Тема 4. Середні величини.

Середня арифметична. Загальна і часткові середні. Властивості середньої арифметичної. Середня гармонійна і геометрична. Загальний вид середніх. Медіана і мода. Квартилі і децилі. Середні в рядах розподілу випадкових величин.

Модульний контроль

Модуль 2.

Змістовний модуль 2. Статистичні гіпотези та критерії.

Тема 5. Вимір варіації

Показники варіації. Правило додавання дисперсій. Дисперсія в рядах розподілу випадкових величин. Моменти розподілу. Характеристики асиметрії й ексцесу.

Тема 6. Статистичні гіпотези та критерії.

Дискретні закони розподілу. Безперервні закони розподілу. Види розподілу: біноміальний, рівномірний, бета-розподіл. Задачі перевірки гіпотез. Види статичних критеріїв, потужність критерію. Типи задач з перевірки гіпотез. t-критерій. Критерії згоди. Число ступенів свободи. Критерій Пірсона, Романовського та Колмогорова-Смірнова.

Тема 7. Вимір зв'язку. Кореляційний, регресійний і дисперсійний аналіз.

Рівняння регресії. Алгоритмічні засоби реалізації інформаційних систем (метод найменших квадратів). Лінійний коефіцієнт кореляції. Коефіцієнт детермінації. Нелінійні рівняння регресії. Множинна регресія. Побудова рівнянь залежностей методом відхилень. Розрахунок коефіцієнта й індексу кореляції, коефіцієнта стійкості зв'язку. Дисперсійний аналіз. Ф-критерій.

Модульний контроль

Змістовний модуль 3. Ряди динаміки та теорія вибірки.

Тема 8. Ряди динаміки.

Інтелектуальні інформаційні технології для прийняття управлінських рішень в складних системах та процесах. Характеристики рядів динаміки. Методи вирівнювання динамічних рядів. Вирівнювання динамічних рядів за прямою, за показовою функцією, за параболою, за рядом Фур'є. Автокореляція, кореляція в рядах динаміки.

Тема 9. Основи теорії вибірки.

Інтегрування інформаційних систем. Помилки спостереження. Генеральна сукупність і вибірка. Помилки вибірки при повторному і

безповторному доборі. Три задачі теорії вибірки: визначення довірчих інтервалів, довірчій імовірності і чисельності вибірки. Способи добору: випадковий добір, районований добір, серійний добір. Мала вибірка.

Модульний контроль

4. Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин				
	Усього	У тому числі			
		л	п	лаб.	с. р.
Модуль 1					
Змістовий модуль 1. Основи теорії ймовірностей. Ряди розподілу випадкових величин.					
Тема 1. Вступ до навчальної дисципліни «Дата-аналіз в інформаційних системах»	1	1	-	-	-
Тема 2. Основи теорії ймовірностей.	33	6	6	-	21
Тема 3. Статистична сукупність. Ряди розподілу.	26	3	1	4	18
Тема 4. Середні величини.	17	2	2	4	9
Модульний контроль	2	2	-	-	-
Разом за змістовним модулем 1	79	14	9	8	48
Модуль 2					
Змістовий модуль 2. Статистичні гіпотези та критерії.					
Тема 5. Вимір варіації.	13	2	2	4	5
Тема 6. Статистичні гіпотези та критерії.	25	3	2	8	12
Тема 7. Вимір зв'язку. Кореляційний, регресійний і дисперсійний аналіз.	30	3	1	8	18
Модульний контроль	2	2	-	-	-
Разом за змістовним модулем 2	70	10	5	20	35
Змістовий модуль 3. Ряди динаміки та теорія вибірки					
Тема 8. Ряди динаміки	29	3	1	4	21
Тема 9. Основи теорії вибірки	13	3	1	-	9
Модульний контроль	2	2			
Разом за змістовним модулем 3	44	8	2	4	30
Індивідуальне завдання		-	-	-	32
Усього за 4-й семестр	225	32	16	32	145
Курсова робота з дисципліни	60	-	16	-	44
Усього годин за 5-й семестр	60	-	16	-	44
Усього годин з дисципліни	285	32	32	32	189

5. Теми семінарських занять

№ з/п	Назва теми	Кількість годин
1	Не передбачено навчальним планом	

6. Теми практичних занять

№ з/п	Назва теми	Кількість годин
1	Визначення ймовірності. Залежні і незалежні події. Додавання і множення ймовірностей.	2
2	Умовна ймовірність. Формула повної імовірності.	2
3	Формули Бернуллі і Пуассона. Локальна і інтегральна формули Муавра – Лапласа.	2
4	Побудова дискретних і інтервальних варіаційних рядів. Графічне зображення варіаційних рядів.	1
5	Обчислення середньої арифметичної зваженої в інтервальних рядах. Розрахунок загальної і часткової середніх. Розрахунок середньої гармонійної і геометричної. Визначення медіани і моди, кватилей і децилей у дискретних та інтервальних рядах.	2
6	Розрахунок показників варіації. Розрахунок загальної дисперсії на основі міжгрупової і внутрігрупової. Характеристики асиметрії та ексцесу.	2
7	Оцінка відповідності варіаційних рядів дискретним та безперервним розподілам за критеріями згоди (Пірсона, Колмогорова-Смірнова, Романовського). Дисперсійний аналіз.	2
8	Розрахунок характеристик рядів динаміки. Аналітичне вирівнювання динамічних рядів. Визначення автокореляції, кореляції в рядах динаміки.	2
9	Рішення задач теорії вибірки (визначення довірчих інтервалів, довірчій імовірності і чисельності вибірки).	1
	Разом	16

7. Теми лабораторних занять

№ з/п	Назва теми	Кількість годин
1	Основні прийоми роботи в пакеті Statistica	4
2	Розвідницький аналіз даних: обчислення основних статистик.	4
3	Оцінка виду розподілу за статистичними критеріями.	4
4	Дослідження статистичної значущості розходжень в групах(ч.1).Т-критерій	4
5	Дослідження статистичної значущості розходжень в групах(ч. 2). Дисперсійний аналіз.	4
6	Дослідження статистичної залежності даних: аналіз таблиця спряженості, кореляційний аналіз.	4
7	Проведення регресійного аналізу: множинна	4

	регресія, нелінійна регресія	
8	Часові ряди	4
	Разом	32

8. Самостійна робота

№ з/п	Назва теми	Кількість годин
4-й семестр		
Змістовий модуль 1		
1	Тема 2. Задачі з обчислення геометричної ймовірності. Відхилення відносної частоти від постійної імовірності в незалежних випробуваннях. Закон великих чисел.	21
2	Тема 3. Робота з файлами даних в пакеті STATISTICA. Графічні можливості програмного пакету STATISTICA.	18
3	Тема 4. Метод моментів.	9
Змістовий модуль 2		
4	Тема 5. Метод доданків обчислення середньої та дисперсії. Метод найбільшої правдоподібності.	5
5	Тема 7. Види розподілу: біноміальний, рівномірний, бета-розподіл.	12
6	Тема 3. Визначення виду залежності методом ломаної. Рангова кореляція.	18
Змістовий модуль 3		
7	Тема 8. Корелювання послідовних різниць. Методи прогнозування. Метод експоненційного вирівнювання. Помилки прогнозування.	21
8	Тема 9. Помилки вибірки при районованому та серійному доборі. Індексний метод.	9
9	Індивідуальне завдання	32
Разом за 4-й семестр		145
5-й семестр		
1	Аналіз існуючих систем, публікацій за предметною областю	10
2	Математичні методи дослідження	10
3	Вибір інструментальних засобів	5
4	Реалізація дослідження	15
5	Оформлення розрахунково-пояснювальної записки за зазначеними вимогами.	4
Разом за 5-й семестр		44
Усього годин з дисципліни		189

9. Індивідуальні завдання

РР на тему «Дата-аналіз варіаційних рядів»

10. Методи навчання

Проведення лекцій, практичних занять, лабораторних робіт, індивідуальні консультації з питань нового матеріалу, самостійна робота студентів за матеріалами, опублікованими кафедрою (методичні посібники) та іншими матеріалами, в тому числі електронними.

11. Методи контролю

Виконання самостійних письмових робіт, захист лабораторних робіт, написання тестів з теоретичного матеріалу, іспит (4 семестр), діф. залік (5 семестр)

12. Критерії оцінювання та розподіл балів, які отримують здобувачі

12.1. Розподіл балів, які отримують студенти (кількісні критерії оцінювання)

Складові навчальної роботи	Бали за одне заняття	Кількість занять	Сумарна кількість балів
Змістовний модуль 1			
Робота на практичних заняттях	0...2	4	0...8
Виконання і захист лабораторних робіт	0...5	2	0...10
Модульний контроль (самостійні письмові роботи)	0...6	2	0...12
Модульний контроль (тест)	0...8	1	0...8
Змістовний модуль 2			
Робота на практичних заняттях	0...2	3	0...6
Виконання і захист лабораторних робіт	0...5	5	0...23
Модульний контроль (самостійні письмові роботи)	0...6	1	0...6
Модульний контроль (тест)	0...7	1	0...7
Змістовний модуль 3			

Робота на практичних заняттях	0...2	1	0...2
Виконання і захист лабораторних робіт	0...5	1	0...5
Модульний контроль (самостійні письмові роботи)	0...6	1	0...6
Модульний контроль (тест)	0...7	1	0...7
Усього за семестр			0...100

Семестровий контроль (іспит) проводиться у разі відмови студента від балів поточного тестування й за наявності допуску до іспиту. Під час складання семестрового іспиту студент має можливість отримати максимум 100 балів.

Білет для іспиту складається з теоретичного тесту та задачі. За тест (20 запитань) студент отримує максимально 70 балів. За повне та правильне вирішення задачі – 30 балів.

12.2 Критерії оцінювання роботи студента протягом семестру

Задовільно (60-74). Мати мінімум знань та умінь. Відпрацювати та захистити всі лабораторні роботи. Знати основи теорії ймовірностей. Знати способи графічного зображення варіаційних рядів. Знати способи обчислення середніх величин, дисперсії; критерій Пірсона, метод найменших квадратів. Знати деякі методи вирівнювання динамічних рядів. Знати основні задачі теорії вибірки.

Добре (75-89). Твердо знати мінімум знань, виконати усі завдання. Показати вміння виконувати та захищати всі лабораторні роботи в обумовлений викладачем строк з обґрунтуванням рішень та заходів, які запропоновано у роботах. Знати формули повної імовірності, Байєса, схему Бернуллі. Вміти будувати функцію та щільність розподілу. Знати різні види середніх величин. Вміти обчислювати показники варіації. Знати основні закони розподілу та критерії згоди. Вміти будувати рівняння регресії та оцінювати тісноту зв'язку. Вміти визначати автокореляцію у часових рядах. Знати основи теорії вибірки.

Відмінно (90-100). Повно знати основний та додатковий матеріал. Знати усі теми. Орієнтуватися у підручниках та посібниках. Досконально знати методи теорії ймовірностей. Знати загальний вид середніх та вміти обчислювати структурні середні. Знати правило додавання дисперсій та моменти розподілу. Вміти застосовувати декілька критеріїв для перевірки гіпотез. Знати основи кореляційного, регресійного та дисперсійного аналізу. Знати аналітичні способи вирівнювання часових рядів. Вміти визначати помилки вибірки при повторному і безповторному відборі.

Шкала оцінювання: бальна і традиційна

Сума балів	Оцінка за традиційною шкалою	
	Іспит, диференційований залік	Залік
90 – 100	Відмінно	Зараховано
75 – 89	Добре	
60 – 74	Задовільно	
0 – 59	Незадовільно	Не зараховано

13. Методичне забезпечення

1. Імовірісно-статистичні методи в інформаційно-керуючих системах / О.Є.Федорович, О.В. Малєєва, Н.В. Нечіпорук. - Навч. посібник. - Харків: Нац. аерокосм. ун-т «Харків. авіац. ін-т», 2005. - 202 с.
2. Математична статистика в інформаційних системах/О.В. Малєєва, О.В. Максименко. - Навч. посібник з практичних занять. - Харків: Нац. аерокосм. ун-т «Харків. авіац. ін-т», 2006. – 62 с.
3. Малєєва, О. В. Статистичний аналіз даних (у програмному пакеті STATISTICA 10.0): навч. посіб. до лаб. практикуму / О. В. Малєєва, А. Ю. Юркевич. – Харків : Нац. аерокосм. ун-т ім. М. Є. Жуковського «Харків. авіац. ін-т», 2018. – 48 с.
4. Комплекс навчально-методичного забезпечення дисципліни; презентаційні матеріали; інтерактивні лабораторні роботи (<https://mentor.khai.edu/course/view.php?id=1302>)

14. Рекомендована література

Базова

1. Руденко В. М. Математична статистика. Навч. посіб. – К.: Центр учбової літератури, 2012. – 304 с.
2. Кармелюк Г. Теорія ймовірності та математична статистика Посібник з розв'язування задач. – К.: Центр навчальної літератури, 2017, 576 с.
3. Алілуйко А. П. Практикум з теорії ймовірностей та математичної статистики / Алілуйко А. М., Дзюбановська Н. В., Єрмоєнко В. О., Мартинюк О. М., Шинкарик М. І. — Тернопіль, 2018. — 352 с.

Допоміжна

1. Бобик, О. І., Гладунський В.Н. Довідник з теорії ймовірностей і математичної статистики : Навч. посібник для студентів ВНЗ– Львів : ЛБІ НБУ, 2005.– 143 с.
2. Теорія ймовірностей та математична статистика : [наук. зб.] / Київ. нац. ун-т ім. Тараса Шевченка. – - Виходить щоквартально. - ISSN 0868-6904. - ISSN 1547-7363

3. Статистика: навч. посіб. / Р. В. Фещур, В. П. Кічор, А. Ф. Барвінський, М. Р. Тимошук . – 4-те вид., оновл. і доповн. – Л. : Бух. центр "Ажур", 2010. – 256 с.
4. Чернобай О. Б. Про деякі методичні особливості викладання курсу теорії ймовірностей : дис. – Ірпінь: Університет ДФС України., 2020.

15. Інформаційні ресурси

1. Теорія ймовірностей та математична статистика у прикладах і задачах : навч. посіб. для студ. вищ. навч. закл. / О.Б. Жильцов ; за ред. Г.О. Михаліна. — К. : Київ. ун-т ім. Б. Грінченка, 2015. — 336 с. http://elibrary.kubg.edu.ua/id/eprint/13578/1/O_Zhyltsov_KUBG_TY_UN.pdf
2. Теорія ймовірностей і математична статистика : навчальний посібник / В. М. Астахов, Г. С. Буланов, В. О. Паламарчук. – Краматорськ : ДДМА, 2009. – 64 с. <http://www.dgma.donetsk.ua/metod/vm/tims.pdf>
3. Сайт науково-технічної бібліотеки університету [Електронний ресурс]. – Режим доступу: <http://library.khai.edu>.