

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КРЕМЕНЧУЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ МИХАЙЛА ОСТРОГРАДСЬКОГО
Навчально-науковий інститут електричної інженерії
та інформаційних технологій
Кафедра комп'ютерної інженерії та електроніки

«ЗАТВЕРДЖУЮ»

Проректор з науково-педагогічної та
методичної роботи



Віктор КОСТІН

09

2024 року

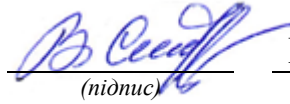
РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

«Імовірісно-статистичні методи інформаційних технологій»

першого (бакалаврського) освітнього рівня
спеціальності 123 «Комп'ютерна інженерія»
освітньо-професійної програми «Комп'ютерна інженерія»

Робоча програма навчальної дисципліни «Імовірно-статистичні методи інформаційних технологій» розроблена на основі освітньо-професійної програми «Комп'ютерна інженерія» підготовки здобувачів вищої освіти освітнього ступеня «Бакалавр» за спеціальністю 123 «Комп'ютерна інженерія» та відповідних нормативних документів

Робочу програму розробив:
доцент кафедри КІЕ, к. т. н.



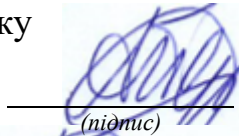
(підпис)

Валерій СИДОРЕНКО
(Власне ім'я ПРИЗВИЩЕ)

Робочу програму обговорено та схвалено на засіданні випускової кафедри освітньо-професійної програми «Комп'ютерна інженерія» спеціальності 123 «Комп'ютерна інженерія»,

Протокол № 1 від «19» вересня 2024 року

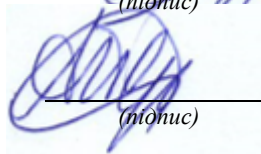
Гарант освітньої програми



(підпис)

Андрій ПЕРЕКРЕСТ
(Власне ім'я ПРИЗВИЩЕ)

Завідувач кафедри



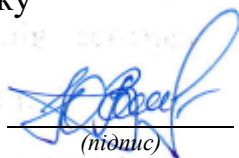
(підпис)

Андрій ПЕРЕКРЕСТ
(Власне ім'я ПРИЗВИЩЕ)

Робочу програму обговорено та схвалено на засіданні науково-методичної ради інституту електричної інженерії та інформаційних технологій,

Протокол № 1 від «24» вересня 2024 року

Голова науково-методичної ради



(підпис)

Юрій ЗАЧЕПА
(Власне ім'я ПРИЗВИЩЕ)

1. Опис навчальної дисципліни

Найменування показників	Галузь знань, напрям підготовки, освітньо-кваліфікаційний рівень	Характеристика навчальної дисципліни	
		денна форма	
Кількість кредитів – 6	Галузь знань 12 «Інформаційні технології»	Обов'язкова	
Модулів – 1	Спеціальність (професійне спрямування): 123 «Комп'ютерна інженерія» Освітньо-професійна програма «Комп'ютерна інженерія»	Рік підготовки:	
Змістових модулів – 2		1-й	
Розрахунково-графічна робота		Семестр	
Загальна кількість годин – 180		2-й	
		Лекції	
Тижневих годин для денної форми навчання: аудиторних – 3 самостійної роботи студента – 6	Освітній ступінь: Бакалавр	30 год.	
		Практичні	
		14	
		Лабораторні	
		16 год.	
		Самостійна робота	
		120 год.	
		Індивідуальні завдання	
–			
–			
Вид контролю: Іспит			

Співвідношення кількості годин аудиторних занять до самостійної і індивідуальної роботи становить:

для денної форми навчання – $60/120=0,5$.

2. Мета та завдання навчальної дисципліни

Мета: є набуття студентами професійних компетенцій в області ймовірнісно-статистичних методів і підготовка студентів до ефективного їх використання в навчальному процесі, подальшій інженерній та науковій діяльності.

Завдання: набуття знань закономірностей випадкових явищ і процесів, вміння використовувати ймовірнісно-статистичні методи при аналізі, моделюванні та проектуванні складових інформаційних технологій.

У результаті вивчення навчальної дисципліни студент повинен

отримати досвід з компетентностей:

ІК. Здатність розв'язувати складні спеціалізовані задачі та практичні проблеми під час професійної діяльності в комп'ютерній галузі або навчання, що передбачає застосування теорії та методів комп'ютерної інженерії і характеризується комплексністю та невизначеністю умов.

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

ЗК 2. Здатність вчитися і оволодівати сучасними знаннями.

ЗК 3. Здатність застосовувати знання у практичних ситуаціях.

ЗК 7. Вміння виявляти, ставити та вирішувати проблеми.

ФК 15. Здатність аргументувати вибір методів розв'язування спеціалізованих задач, критично оцінювати отримані результати, обґрунтовувати та захищати прийняті рішення.

набути навички та уміння:

ПРН 2. Мати навички проведення експериментів, збирання даних та моделювання в комп'ютерних системах.

ПРН 6. Вміти застосовувати знання для ідентифікації, формулювання і розв'язування технічних задач спеціальності, використовуючи методи, що є найбільш придатними для досягнення поставлених цілей.

ПРН 8. Вміти системно мислити та застосовувати творчі здібності до формування нових ідей.

ПРН 11. Вміти здійснювати пошук інформації в різних джерелах для розв'язання задач комп'ютерної інженерії.

ПРН 17. Спілкуватись усно та письмово з професійних питань українською мовою та однією з іноземних мов (англійською, німецькою, італійською, французькою, іспанською).

ПРН 18. Використовувати інформаційні технології для ефективного спілкування на професійному та соціальному рівнях.

ПРН 20. Усвідомлювати необхідність навчання впродовж усього життя з метою поглиблення набутих та здобуття нових фахових знань, удосконалення креативного мислення.

ПРН 21. Якісно виконувати роботу та досягати поставленої мети з дотриманням вимог професійної етики.

3. Програма навчальної дисципліни

Змістовий модуль 1 Теорія ймовірностей та ймовірнісні процеси

Тема 1 Випадкові події та їх аналіз

Тема 1.1 Основи комбінаторики.

Основні формули комбінаторики. Теорема про добуток шансів. Урни та кульки. Урнова схема: вибір без повернення та з урахуванням порядку. Урнова схема: вибір без повернення та без урахування порядку. Урнова схема: вибір з поверненням та з урахуванням порядку. Урнова схема: вибір з поверненням та без урахування порядку.

Тема 1.2 Класична ймовірнісна схема.

Основні поняття елементарної теорії ймовірностей. Предмет теорії ймовірностей. Статистичне визначення ймовірності. Простір елементарних подій. Операції над подіями (алгебра подій). Ймовірність на дискретному просторі елементарних подій. Класичне визначення ймовірності. Гіпергеометричний розподіл.

Тема 1.3 Геометрична ймовірність.

Задача про зустріч. Задача Бюффона. Парадокс Бертрана.

Тема 1.4 Аксиоматика теорії ймовірностей.

σ -алгебра подій. Аксиоми σ -алгебри. Поняття міри. Ймовірність як нормована міра. Аксиоми ймовірностей. Властивості ймовірностей, що витікають з аксіом.

Тема 1.5 Умовна ймовірність, незалежність.

Умовна ймовірність. Незалежність. Формула повної ймовірності. Формула Байеса. Роль байєсівської підхід в машинному навчанні.

Тема 1.6 Схема Бернуллі.

Розподіл числа успіхів у n випробуваннях. Найбільш ймовірне число успіхів. Номер першого успішного випробування. Наближення гіпергеометричного розподілу біноміальним. Наближені формули Лапласа. Наближена формула Пуассона.

Тема 2 Випадкові величини

Тема 2.1 Дискретні випадкові величини. Закони розподілу дискретних випадкових величин. Біноміальний розподіл. Розподіл Пуассона. Геометричний розподіл. Гіпергеометричний розподіл. Числові характеристики дискретних випадкових величин. Математичне сподівання дискретної випадкової величини. Дисперсія дискретної випадкової величини. Начальні та центральні моменти k -го порядку. Числові характеристики деяких дискретних законів.

Тема 2.2 Функція розподілу ймовірностей випадкової величини.

Неперервні випадкові величини. Щільність розподілу. Числові характеристики. Представлення функції розподілу та щільності розподілу дискретної випадкової величини за допомогою функції Хевісайда та δ -функції Дірака.

Тема 2.3 Числові характеристики неперервних випадкових величин.

Числові характеристики неперервних випадкових величин: математичне сподівання, дисперсія, початкові та центральні моменти k -го порядку. Закони розподілу ймовірностей неперервної випадкової величини: рівномірний, експоненціальний, нормальний. Властивості нормального розподілу. Правило «трьох сигм». Розподіл Вейбулла. Розподіл Парето.

Тема 3 Системи і функції випадкових величин.

Тема 3.1 Багатовимірні випадкові величини.

Функція розподілу багатовимірної випадкової величини. Властивості функції сумісного розподілу. Типи багатовимірних розподілів. Властивості щільності сумісного розподілу. Залежність та незалежність випадкових величин.

Система двох випадкових величин. Числові характеристики системи двох випадкових величин. Кореляційний момент. Коефіцієнт кореляції. Властивості коваріації. Властивості коефіцієнта кореляції.

Тема 3.2 Корельованість та залежність випадкових величин.

Корельованість та залежність випадкових величин. Поняття регресії. Рівняння лінійної регресії для двох випадкових величин. Метод найменших квадратів (МНК). Лінійна кореляція. Нормальна кореляція.

Тема 3.3 Перетворення випадкових величин.

Закони розподілення функцій випадкових аргументів. Перетворення однієї випадкової величини. Закон розподілення монотонної функції одного випадкового аргументу. Закон розподілення лінійної функції від аргументу, що підпорядковується нормальному закону. Закон розподілення немонотонної функції одного випадкового аргументу.

Функції від двох випадкових величин. Приклади використання формули згортки.

Тема 3.4 Закон великих чисел.

Попередні зауваження. Нерівність Чебишова. Закон великих чисел у формі Чебишова. ЗВЧ у формі Бернуллі. ЦГТ.

Тема 4 Випадкові процеси

Тема 4.1 Найпростіший потік подій.

Випадкові процеси. Процес Пуассона. Вінеровський процес. Марківський процес.

Тема 4.2 Елементи теорії СМО. Елементи теорії СМО. Ланцюги Маркова

Модель черги на базі теорії СМО. Загальні відомості про СМО. n – канална СМО без черги. Граф станів СМО « n -клієнтів - сервер». Диференціальні рівняння і розрахунок ймовірності станів СМО. n – канална СМО з чергою. Задача обслуговування найпростішого потоку подій СМО з очікуванням та експоненціальним часом обслуговування.

Початкові відомості про ланцюг Маркова. Однорідний ланцюг Маркова. Перехідні ймовірності. Матриця переходу. Рівняння Маркова.

Змістовий модуль 2 Математична статистика та обробка результатів вимірювання в середовищі R

Тема 5 Математична статистика і обробка результатів вимірювань.

Тема 5.1 Основи вибіркового методу.

Поняття вибірки. Статистичний розподіл. Інтервальний статистичний розподіл. Міркування щодо вибору кількості інтервалів інтервальної таблиці. Гістограма і полігон. Вибіркові математичне сподівання, дисперсія, моменти вищих порядків, асиметрія, ексцес, медіана. Емпірична функція розподілу. Теорема Гливленко. Теорема Колмогорова, довірчі межі для емпіричної функції розподілу.

Тема 5.2 Точкові статистичні оцінки.

Точкові статистичні оцінки. Метод моментів та метод максимальної правдоподібності. Незміщені, обґрунтовані та ефективні оцінки. Нерівність Крамера-Рао. Точкові оцінки параметрів розподілів: нормального, рівномірного, експоненціального, Коші, логнормального, Пуассона, Бернуллі, біноміального.

Тема 5.3 Точні вибіркові розподілення.

Точні вибіркові розподіли, їхня сутність. Теореми про точне представлення для χ^2 -розподілу, розподілу Стюдента, розподілу Фішера.

Тема 5.4 Інтервальні статистичні оцінки

Інтервальні оцінки (ІО). ІО для параметрів нормального розподілу. ІО параметрів розподілів: рівномірного, експоненціального, Пуассона, Бернуллі. Загальний підхід до інтервального оцінювання на основі ЦГТ.

Тема 6 Перевірка статистичних гіпотез.

Тема 6.1 Статистичні гіпотези. Основні поняття.

Нульова (основна) та альтернативна (конкуруюча) Проста і складна статистичні гіпотези. Статистичний критерій, критичні точки. Поняття критичної області. Лівостороння, правостороння та двостороння критичні області. Похибка першого та другого роду. Потужність статистичного критерію.

Тема 6.2 Перевірка статистичних гіпотез щодо закону розподілення.

Перевірка статистичної гіпотези щодо виду математичної моделі закону розподілу випадкової величини. Параметричні та непараметричні критерії. Критерії χ^2 -Пірсона, Романовського, Смірнова-Колмогорова, ω^2 .

Тема 6.3 Перевірка статистичних гіпотез щодо параметрів розподілень випадкових величин.

Перевірка статистичних гіпотез щодо рівності математичних сподівань двох вибірок. Пов'язані та непов'язані вибірки.

Перевірка статистичної гіпотези щодо однорідності дисперсій; критерії Фішера, Бартлетта, Кохрена. Проблеми робастності статистичних критеріїв.

Тема 7 Прикладні методи математичної статистики

Тема 7.1 Основи дисперсійного аналізу.

Однофакторний дисперсійний аналіз за Фішером. Загальна, факторна та залишкова суми квадратів відхилень. Модель даних та оцінка її параметрів. Зв'язок між загальною, факторною та залишковою сумами. Загальна, факторна та залишкова суми квадратів відхилень. Порівняння декількох середніх методом дисперсійного аналізу.

Оцінка впливу фактора. Методи множинних порівнянь Тьюкі та Шеффе.

Дисперсійний аналіз з повтореннями.

Тема 7.2 Основи кореляційного аналізу

Елементи теорії кореляції. Функціональна, статистична та кореляційна залежність. Умовні середні. Кореляційна таблиця. Точкова та інтервальна оцінки коефіцієнта кореляції Пірсона. Перевірка значимості оцінки коефіцієнта кореляції.

Непараметрична кореляція. Номінальні та порядкові шкали. Коефіцієнти кореляції Спірмена та Кендала.

Тема 7.3 Основи регресійного аналізу

Елементи регресійного аналізу. Передумови лінійної регресії. Вибіркове рівняння лінійної регресії. МНК-оцінки рівняння лінійної регресії. Перевірка значимості коефіцієнтів моделі, критерій Стьюдента. Перевірка адекватності, критерій Фішера та коефіцієнт детермінації.

Внутрішньо лінійні регресійні моделі. Лінеаризація нелінійних моделей.

4. Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин				
	денна форма				
	усього	у тому числі			
л		п	лаб	с. р.	
1	2	3	4	5	6
Змістовий модуль 1 Теорія ймовірностей та ймовірнісні процеси					
Тема 1 Випадкові події та їх аналіз					
Тема 1.1 Основи комбінаторики		1	1	-	-
Тема 1.2 Класична ймовірнісна схема		1	1	-	2
Тема 1.3 Геометрична ймовірність		1	1	-	2
Тема 1.4 Аксиоматика теорії ймовірностей		1	1	-	2
Тема 1.5 Умовна ймовірність, незалежність		1	1	-	3
Тема 1.6 Схема Бернуллі		1	1	-	3
Тема 2 Випадкові величини					
Тема 2.1 Дискретні випадкові величини		0,5	0,5	-	5
Тема 2.2 Функція розподілу ймовірностей випадкової величини.		1	1	-	5
Тема 2.3 Числові характеристики неперервних випадкових величин.		0,5	0,5	-	5
Тема 3 Системи і функції випадкових величин					
Тема 3.1 Багатовимірні випадкові величини		1	-	-	5
Тема 3.2 Корельованість та залежність випадкових величин		1	1	-	5
Тема 3.3 Перетворення випадкових величин		1	1	-	5
Тема 3.4 Закон великих чисел		1	1	-	-
Змістовий модуль 4. Випадкові процеси					
Тема 14. Найпростіший потік подій.		4	1	-	7
Тема 15. Елементи теорії СМО. Елементи теорії СМО. Ланцюги Маркова		4	1	-	15
Разом за змістовим модулем 1		20	13	-	64
Змістовий модуль 2 Математична статистика та обробка результатів вимірювання в середовищі R					
Тема 5 Математична статистика і обробка результатів вимірювань					
Тема 5.1 Основи вибіркового методу		1	1	2	2
Тема 5.2 Точкові статистичні оцінки		1	-	1	4
Тема 5.2 Точні вибіркові розподілення		1	-	-	-
Тема 5.3 Інтервальні статистичні оцінки		1	-	1	4
Тема 6 Перевірка статистичних гіпотез					

Тема 6.1 Статистичні гіпотези. Основні поняття.		1	–	–	–
Тема 6.2 Перевірка статистичних гіпотез щодо закону розподілення.		1	–	2	3
Тема 6.3 Перевірка статистичних гіпотез щодо параметрів розподілень випадкових величин		–	–	2	3
Семестровий контроль за змістовим модулем 1				–	2
Тема 7 Прикладні методи математичної статистики					
Тема 7.1 Основи дисперсійного аналізу		2	–	2	4
Тема 7.2 Основи кореляційного аналізу.		1	–	2	3
Тема 7.3 Основи регресійного аналізу.		1	–	4	5
Семестровий контроль за змістовим модулем 2				-	4
ІНДЗ (РГР)		–	–	–	22
Семестровий контроль (іспит)		–	–	–	–
Разом за змістовим модулем 2		10	1	16	56
Разом годин		30	14	16	120

5. Теми практичних занять

№ з/п	Назва теми	Кількість годин
		дфн
1	Елементи комбінаторики. Класичне визначення ймовірності. Застосування комбінаторики для розрахунку ймовірностей	2
2	Геометрична ймовірність. Алгебра подій. Теорема множення та додавання ймовірностей. Формула повної ймовірності та формула Байєса	3
3	Схема Бернуллі	1
4	Закони розподілу та числові характеристики дискретних випадкових величин. Закони розподілу та числові характеристики неперервних випадкових величин. Нормальний закон	2
5	Закони розподілу функцій випадкових величин. Композиція законів розподілу. Розподіл екстремальних значень	3
6	Найпростіший потік подій. Елементи теорії СМО. Ланцюги Маркова	2
7	Основи вибіркового методу. Точкові статистичні оцінки. Точні вибіркові розподілення. Інтервальні статистичні оцінки	1
	Разом	14

6. Теми лабораторних занять

№ з/п	Назва теми	Кількість годин
		дфн
1	Лаб. роб. 1. Налаштування Git+GitHub середовища	2
2	Лаб. роб. 2. Структура проекту зі статистичної обробки даних у стилі literate programming. Налаштування R+RStudio+Quarto середовища	2
3	Лаб. роб. 3. Основи вибіркового методу	2
4	Лаб. роб. 4. Перевірка статистичних гіпотез щодо закону розподілу. Перевірка на нормальність	4
5	Лаб. роб. 5. Дисперсійний аналіз (ANOVA)	2
6	Лаб. роб. 6. Випадок з двома вибірками. Побудова моделей регресії	4
	Разом	16

7. Тематика РГР

№ з/п	Назва теми	Опис
1	Аналіз розподілу часу відповіді на запити до бази даних	<p>Отримати набір даних, що містить час відповіді (у мілісекундах) на певну кількість запитів до бази даних. Необхідно:</p> <ul style="list-style-type: none"> • Побудувати гістограму та оцінити параметри розподілу (наприклад, експоненціальний, нормальний). • Перевірити гіпотезу про відповідність вибірки певному розподілу (наприклад, за допомогою критерію χ^2-квадрат). • Обчислити довірчий інтервал для середнього часу відповіді
2	Статистичний аналіз кількості помилок у програмному коді	<p>Отримати дані про кількість помилок, знайдених під час тестування певного програмного модуля (наприклад, кількість помилок на 1000 рядків коду). Необхідно:</p> <ul style="list-style-type: none"> • Оцінити середню кількість помилок та її дисперсію. • Побудувати розподіл кількості помилок та порівняти його з теоретичним розподілом (наприклад, розподілом Пуассона). • Перевірити гіпотезу про те, чи перевищує середня кількість помилок заданий поріг
3	Аналіз залежності між обсягом вхідних даних та обсягом оперативної пам'яті, необхідної для обробки цих даних	<p>Розглянути програму, яка обробляє дані (наприклад, зчитує дані з файлу, виконує певні обчислення). Він змінює обсяг вхідних даних та вимірює обсяг оперативної пам'яті (RAM), який використовується програмою під час обробки. Необхідно:</p> <ul style="list-style-type: none"> • Виміряти обсяг оперативної пам'яті (RAM), що використовується програмою для різних обсягів вхідних даних. • Побудувати діаграму розсіювання (scatter plot) залежності обсягу RAM від обсягу вхідних даних. • Обчислити коефіцієнт кореляції між обсягом RAM та обсягом вхідних даних. • Побудувати лінійну регресійну модель для прогнозування обсягу RAM на основі обсягу вхідних даних. • Зробити висновки про те, чи існує залежність між обсягом вхідних даних та обсягом оперативної пам'яті, та чи можна використовувати розмір вхідних даних для прогнозування використання RAM
4	Аналіз часу безвідмовної роботи (MTBF) серверів	<p>Отримати дані про час безвідмовної роботи (у годинах) певної кількості серверів. Завдання:</p> <ul style="list-style-type: none"> • Оцінити середній час безвідмовної роботи (MTBF). • Побудувати гістограму розподілу часу безвідмовної роботи та перевірити гіпотезу про його відповідність експоненційному розподілу. • Обчислити ймовірність того, що сервер пропрацює безвідмовно заданий час (наприклад, 1000 годин)
5	Статистичний аналіз розміру файлів у файловій системі	<p>Зібрати дані про розміри файлів (у байтах) у певній директорії. Завдання:</p> <ul style="list-style-type: none"> • Побудувати гістограму розподілу розмірів файлів. • Оцінити середній розмір файлу та його дисперсію. • Визначити, яка частка файлів має розмір, що перевищує заданий поріг (наприклад, 1 МБ)

6	Аналіз кількості запитів до API протягом певного часу	Студент отримує дані про кількість запитів, що надходять до певного API протягом кожної хвилини (години, дня). Завдання: <ul style="list-style-type: none"> • Побудувати графік часового ряду кількості запитів. • Оцінити середню кількість запитів та її дисперсію. • Виявити періоди з підвищеним та зниженим навантаженням на API
7	Оцінка ймовірності виникнення колізій хеш-функції	Дослідити хеш-функцію з відомою кількістю можливих значень хешу. Він генерує випадкові вхідні дані та обчислює їхні хеші. Завдання: <ul style="list-style-type: none"> • Експериментально оцінити ймовірність виникнення колізії (два різних вхідних значення мають однаковий хеш). • Порівняти отриману оцінку з теоретичною оцінкою ймовірності колізії (наприклад, використовуючи парадокс днів народження)
8	Аналіз часу завантаження веб-сторінок	Збирає дані про час завантаження певної веб-сторінки з різних місць розташування або в різний час доби. Завдання: <ul style="list-style-type: none"> • Оцінити середній час завантаження та його дисперсію. • Побудувати довірчий інтервал для середнього часу завантаження. • Перевірити гіпотезу про те, чи перевищує середній час завантаження заданий поріг (наприклад, 3 секунди)
9	Статистичний аналіз кількості спам-повідомлень, отриманих користувачем	Отримати дані про кількість спам-повідомлень, отриманих користувачем електронної пошти за певний період часу (наприклад, за день, тиждень, місяць). Завдання: <ul style="list-style-type: none"> • Оцінити середню кількість спам-повідомлень. • Побудувати розподіл кількості спам-повідомлень та порівняти його з розподілом Пуассона. • Обчислити ймовірність того, що користувач отримає більше заданої кількості спам-повідомлень за певний період
10	Аналіз кількості відмов (errors) при доступі до веб-сервісу	Зібрати дані про кількість HTTP-помилки (наприклад, 404, 500) при доступі до певного веб-сервісу за певний період часу. Завдання: <ul style="list-style-type: none"> • Оцінити середню кількість помилок. • Виявити періоди з підвищеною кількістю помилок. • Проаналізувати причини виникнення помилок (якщо можливо)
11	Оцінка ефективності алгоритму стиснення даних	Дослідити алгоритм стиснення даних (наприклад, Lempel-Ziv). Він стискає різні файли та вимірює коефіцієнт стиснення (відношення розміру стиснутого файлу до розміру оригінального файлу). Завдання: <ul style="list-style-type: none"> • Оцінити середній коефіцієнт стиснення для різних типів файлів. • Визначити, які типи файлів найкраще піддаються стисненню. • Порівняти ефективність різних алгоритмів стиснення на основі статистичного аналізу коефіцієнтів стиснення
12	Аналіз продуктивності мережі (ping)	Використати утиліту ping для вимірювання часу відгуку (у мілісекундах) при з'єднанні з певним сервером. Він виконує ping кілька разів та збирає дані. Завдання: <p>Оцінити середній час відгуку та його дисперсію.</p> <p>Обчислити довірчий інтервал для середнього часу відгуку.</p> <p>Визначити, чи є значні відмінності у часі відгуку в різний час доби</p>
13	Аналіз впливу різних факторів на кількість завантажень мобільного додатку	Зібрати дані про кількість завантажень мобільного додатку, а також про різні фактори, які можуть впливати на цю кількість (наприклад, рейтинг додатку в магазині, кількість позитивних відгуків, витрати на рекламу). Необхідно: <p>Визначити, які фактори мають найбільший вплив на кількість завантажень за допомогою кореляційного аналізу</p> <p>Побудувати</p>

		<p>множинну регресійну модель для прогнозування кількості завантажень на основі цих факторів.</p> <p>Оцінити, на скільки збільшиться кількість завантажень при збільшенні рейтингу додатку або збільшенні витрат на рекламу</p>
14	Порівняльний аналіз продуктивності різних алгоритмів машинного навчання на одному наборі даних	<p>Обрати набір даних для задачі класифікації або регресії та декілька різних алгоритмів машинного навчання. Необхідно:</p> <p>Навчити кожну модель та оцінити її продуктивність за допомогою певних метрик.</p> <p>Застосувати дисперсійний аналіз (ANOVA) для визначення, чи є статистично значуща різниця між продуктивністю різних алгоритмів.</p> <p>Якщо різниця є значущою, провести попарне порівняння алгоритмів, щоб визначити, які з них працюють краще за інші</p>
15	Дослідження впливу типу SSD-накопичувача на швидкість читання/запису даних	<p>Протестувати декілька різних типів SSD-накопичувачів (наприклад, SATA, NVMe) та вимірює швидкість читання/запису даних для кожного з них. Необхідно:</p> <p>Застосувати дисперсійний аналіз (ANOVA) для визначення, чи є статистично значуща різниця між швидкістю читання/запису для різних типів накопичувачів.</p> <p>Якщо різниця є значущою, провести попарне порівняння типів накопичувачів, щоб визначити, які з них працюють швидше за інші</p>
16	Аналіз впливу різних параметрів мережі на затримку (latency) передачі даних	<p>проводить вимірювання затримки передачі даних (ping) в мережі, змінюючи різні параметри (наприклад, розмір пакету, відстань до сервера). Необхідно:</p> <p>Визначити, які параметри мають найбільший вплив на затримку за допомогою кореляційного аналізу.</p> <p>Побудувати регресійну модель для прогнозування затримки на основі цих параметрів.</p> <p>Провести дисперсійний аналіз для порівняння затримки при різних значеннях певного параметра (наприклад, при різних розмірах пакету)</p>
17	Аналіз залежності між розміром коду програми (в рядках) та часом її виконання	<p>Студент пише декілька програм, що виконують однакову задачу (наприклад, сортування масиву, обчислення факторіалу) різними способами, щоб отримати різний розмір коду (в рядках). Далі, для кожної програми він заміряє час її виконання. Необхідно:</p> <ul style="list-style-type: none"> • Виміряти розмір коду (в рядках) кожної програми. • Виміряти час виконання кожної програми на одному й тому ж наборі вхідних даних. • Побудувати діаграму розсіювання (scatter plot) залежності часу виконання від розміру коду. • Обчислити коефіцієнт кореляції між розміром коду та часом виконання. • Побудувати лінійну регресійну модель для прогнозування часу виконання на основі розміру коду. • Зробити висновки про те, чи існує залежність між розміром коду та часом виконання, та чи можна використовувати розмір коду для прогнозування часу виконання
18	Аналіз залежності між кількістю коментарів до посту в соціальній	<p>Зібрати дані про кількість коментарів та лайків для певних постів у соціальній мережі (наприклад, Facebook, Instagram, Twitter). Дані можна зібрати вручну або за допомогою API соціальної мережі. Необхідно:</p> <ul style="list-style-type: none"> • Обчислити коефіцієнт кореляції між кількістю коментарів та кількістю лайків.

	мережі та кількістю лайків	<ul style="list-style-type: none"> • Побудувати діаграму розсіювання (scatter plot) залежності кількості лайків від кількості коментарів. • Побудувати лінійну регресійну модель для прогнозування кількості лайків на основі кількості коментарів. • Зробити висновки про те, чи існує залежність між кількістю коментарів та кількістю лайків, та чи можна використовувати кількість коментарів для прогнозування популярності посту
19	Аналіз впливу різних факторів на час відгуку веб-сервера	<p>Виміряти час відгуку веб-сервера (наприклад, за допомогою утиліти ping або інструментів розробника веб-браузера), змінюючи різні фактори, такі як:</p> <ul style="list-style-type: none"> • Розмір HTTP-запиту (наприклад, розмір зображення, яке завантажується). • Кількість одночасних запитів. • Відстань до сервера (наприклад, вимірюючи час відгуку до різних серверів). • Побудувати регресійну модель, яка описує залежність часу відгуку від цих факторів. • Визначити, які фактори мають найбільший вплив на час відгуку. • Зробити висновки про оптимальні параметри для досягнення мінімального часу відгуку
20	Аналіз залежності між кількістю переглядів відео на YouTube та тривалістю відео	<p>Зібрати дані про кількість переглядів та тривалість відео на YouTube. Дані можна зібрати вручну або за допомогою API YouTube. Необхідно:</p> <ul style="list-style-type: none"> • Обчислити коефіцієнт кореляції між кількістю переглядів та тривалістю відео. • Побудувати діаграму розсіювання (scatter plot) залежності кількості переглядів від тривалості відео. • Побудувати лінійну регресійну модель для прогнозування кількості переглядів на основі тривалості відео. • Зробити висновки про те, чи існує залежність між тривалістю відео та його популярністю

8. Самостійна робота

№ з/п	Назва теми	Кількість годин
		ДФН
1	Класична ймовірнісна схема	2
2	Геометрична ймовірність	2
3	Аксиоматика теорії ймовірностей	2
4	Умовна ймовірність, незалежність	3
5	Схема Бернуллі	3
6	Дискретні випадкові величини	5
7	Функція розподілу ймовірностей випадкової величини	5
8	Числові характеристики неперервних випадкових величин	5
9	Багатовимірні випадкові величини	5
9	Корельованість та залежність випадкових величин	5
10	Перетворення випадкових величин	5
11	Найпростіший потік подій	7
12	Елементи теорії СМО. Елементи теорії СМО. Ланцюги Маркова	15
13	Основи вибіркового методу	2
14	Точкові статистичні оцінки	4
15	Інтервальні статистичні оцінки	4
16	Перевірка статистичних гіпотез щодо закону розподілення	3
17	Перевірка статистичних гіпотез щодо параметрів розподілень випадкових величин	3
18	Семестровий контроль по змістовому модулю 1	2
19	Основи дисперсійного аналізу	4
20	Основи кореляційного аналізу	3
21	Основи регресійного аналізу	5
22	Семестровий контроль по змістовому модулю 2	4
23	Робота над індивідуальним завданням (РГР)	22
	Разом	120

9. Методи навчання

Пояснювально-ілюстративні, репродуктивні (опитування, тестування, розв'язування задач, виконання вправ за зразком).

Лекції, практичні заняття, лабораторні роботи, консультації, самостійна робота.

Лекції викладаються зокрема з використанням мультимедійних засобів.

Самостійне опрацювання навчального матеріалу виконується з використанням конспекту лекцій, відеоматеріалів, основної та додаткової навчальної літератури, інформаційних ресурсів.

10. Методи контролю

Облік відвідування, опитування, захист лабораторних робіт, комплекти тестових завдань для проведення поточного та підсумкового контролю.

11. Розподіл балів, що отримують студенти

Критерії оцінювання		
Вид роботи	Зміст	Бали
Лекції	Робота на лекціях	10
Практичні	Робота на практичних заняттях, РГР	20
Завдання з ЛР (виконання, захист)	Лабораторна робота № 1	5
	Лабораторна робота № 2	5
	Лабораторна робота № 3	5
	Лабораторна робота № 4	5
	Лабораторна робота № 5	5
	Лабораторна робота № 6	5
Контроль	Тест зі змістового модуля 1	10
	Тест зі змістового модуля 2	10
	Іспит	20
	Разом балів	100

Шкала оцінювання: національна та ECTS

Сума балів за 100-бальною шкалою	Оцінка в ECTS	Значення оцінки ECTS	Критерії оцінювання	Рівень компетентості	Оцінка за національною шкалою
					іспит, диференційований залік
90-100	A	відмінно	Студент виявляє особливі творчі здібності, вміє самостійно здобувати знання, без допомоги викладача знаходить та опрацьовує необхідну інформацію, вміє використовувати набуті знання і вміння для прийняття рішень у нестандартних ситуаціях, переконливо аргументує відповіді, самостійно розкриває власні обдарування і нахили	Високий (творчий)	відмінно
82-89	B	дуже добре	Студент вільно володіє вивченим обсягом матеріалу, застосовує його на практиці, вільно розв'язує вправи і задачі у стандартних ситуаціях, самостійно виправляє допущені помилки, кількість яких незначна	Достатній (конструктивно-варіативний)	добре
74-81	C	добре	Студент вміє зіставляти, узагальнювати, систематизувати інформацію під керівництвом викладача; в цілому самостійно застосовувати її на практиці; контролювати власну діяльність; виправляти помилки, серед яких є суттєві, добирати аргументи для підтвердження думок		
64-73	D	задовільно	Студент відтворює значну частину теоретичного матеріалу, виявляє знання і розуміння основних положень; з допомогою викладача може аналізувати навчальний матеріал, виправляти помилки, серед яких є значна кількість суттєвих	Середній (репродуктивний)	задовільно

60-63	E	достатньо	Студент володіє навчальним матеріалом на рівні, вищому за початковий, значну частину його відтворює на репродуктивному рівні		
35-59	FX	незадовільно з можливістю повторного складання семестрового контролю	Студент володіє матеріалом на рівні окремих фрагментів, що становлять незначну частину навчального матеріалу	Низький (рецептивно-продуктивний)	незадовільно
1-34	F	незадовільно з обов'язковим повторним вивченням залікового кредиту	Студент володіє матеріалом на рівні елементарного розпізнання і відтворення окремих фактів, елементів, об'єктів		

10. Методичне забезпечення

1. Методичні вказівки щодо виконання практичних робіт та самостійної роботи з навчальної дисципліни «Теорія ймовірностей та математична статистика» для студентів денної форми навчання зі спеціальностей 122 –

«Комп'ютерні науки» та 123 – «Комп'ютерна інженерія», Частина I. 2023. – 43 стор.

2. Методичні вказівки щодо виконання лабораторних та самостійної робіт студентів з навчальної дисципліни «Імовірісно-статистичні методи інформаційних технологій» для студентів усіх форм навчання для спеціальностей 122 – «Комп'ютерні науки» та 123 – «Комп'ютерна інженерія». Частина II, 2023. – 42 стор.

3. Методичні вказівки щодо виконання розрахунково-графічної роботи з навчальної дисципліни «Імовірісно-статистичні методи інформаційних технологій» для студентів денної форми навчання для спеціальностей 122– «Комп'ютерні науки» та 123–«Комп'ютерна інженерія», 2023. – 16 стор.

11. Рекомендована література

Основна

1. Dembo, A. (2019, April 23). Probability Theory: STAT310/MATH230. Department of Mathematics, Stanford University. Email: amir@math.stanford.edu. URL: <https://web.stanford.edu/class/stats310a/lnotes.pdf>

2. McMullen, C. (2011). *Probability Theory: Course Notes*. Harvard University. Retrieved March 29, 2021. URL: <https://people.math.harvard.edu/~ctm/papers/home/text/class/harvard/154/course/course.pdf>

3. F. M. Dekking, C. Kraaikamp, H. P. Lopuhaä, L. E. A Modern Introduction to Probability and Statistics: Understanding Why and How Meester. Springer Science & Business Media, 2005 - 486 стор.

4. David F. Anderson, Timo Seppäläinen, Benedek Valkó. Introduction to Probability. Cambridge University Press, 2 лист. 2017 р.

5. Mark Ward, Ellen Gundlach, W. H. Freeman. Introduction to Probability. 12 черв. 2015 р. – 704 стор.

6. Курс теорії ймовірностей : підручник / Б. В. Гнеденко – К.: Видавничо-поліграфічний центр «Київський університет», 2010. 464 с. URL: <https://probability.knu.ua/userfiles/yamnenko/Gnedenko.pdf>

7. Математична статистика : навч. посіб. [Електронне видання] / [С. М. Григулич, В. П. Лісов-ська, О. І. Макаренко та ін.]. К. : КНЕУ, 2015. 203 с. URL: <https://core.ac.uk/reader/43282100>

8. Васильків І.М. Основи теорії ймовірностей і математичної статистики: навч. посібник. Львів : ЛНУ імені Івана Франка, 2020. 184 с. URL: https://new.mmf.lnu.edu.ua/wp-content/uploads/2020/04/VasyI-kiv-I.M.-TIMS_CHASTYNA_1.pdf

9. Медведєв М. Г., Пащенко І. О. Теорія ймовірностей та математична статистика. Підручник. К.: Вид-во "Ліра-К". 2008. 536 с. https://pdf.lib.vntu.edu.ua/books/2015/Medvedev_2008_536.pdf

Допоміжна

10. NR Draper, H Smith. Applied regression analysis. John Wiley & Sons. New York 1981.
11. Stroock, D. W. (2011). *Probability Theory: An Analytic View* (2nd ed.). Cambridge University Press. ISBN-10: 0521132509, ISBN-13: 978-0521132503.
12. MIT OpenCourseWare. (2014). *Theory of Probability* (Spring 2014). Instructor: Prof. Scott Sheffield. Department: Mathematics. Course URL: <https://ocw.mit.edu/courses/18-175-theory-of-probability-spring-2014/>
13. Найко Д.А. Шевчук О. Ф. Теорія ймовірностей та математична статистика: навч. посіб. / Д.А. Найко, О.Ф. Шевчук – Вінниця: ВНАУ, 2020. – 382 с. <http://repository.vsau.org/getfile.php/24513.pdf>
14. Сидоренко В. М., Садовнича С. А., Долударєва Є. В. Оптимізація структури тестових завдань навчальних онлайн-курсів на основі ймовірнісної моделі / Інженерні та освітні технології. 2022. Т. 10. № 2. С. 27–36. doi: <https://doi.org/10.30929/2307-9770.2022.10.02.03>
15. Artem Artemenko, Oleksii Chorny, Valeriy Sydorenko, Serhii Serhiienko, Yurii Zacheпа, Vitaliy Kuznetsov, & Alisa Kuznetsova. (2021). ESTIMATING THE PROBABILITY OF THE EMERGENCY OPERATION OF THE QUARRY ELECTRIC LOCOMOTIVE TRACTION ELECTRIC DRIVE. World Science, (2(63)). https://doi.org/10.31435/rsglobal_ws/28022021/7448
16. Lecture Notes in Networks and Systems, vol 707. Springer, Cham. pp. 36-44. https://doi.org/10.1007/978-3-031-34721-4_5
17. Сидоренко В. М., Кирилах Н. Г. Дидактико-методичні аспекти викладання теорії ймовірностей та математичної статистики студентам ІТ напрямку. Інженерні та освітні технології. 2023. Т. 11. № 3. С. 17–23. Doi: <https://doi.org/10.32782/2307-9770.2023.11.03.02>
18. Guchenko M., Sydorenko V., Belska V., Liutenko M., Fesenko N. DComFra Project Learning Module M20 Advanced Spreadsheets in Mathematical Modeling Tasks of Electrical and Computer Engineers Education Proceedings of the 20th IEEE International Conference on Modern Electrical and Energy Systems, MEES 2021

12. Інформаційні ресурси

1. The Comprehensive R Archive Network. (n.d.). Retrieved from <https://cran.r-project.org/>
2. Posit. (n.d.). Put data into production with Posit Connect. Retrieved December 22, 2023, from <https://posit.co/>
3. Shiny. (n.d.). Easy web apps for data science without the compromises. Retrieved December 22, 2023, from <https://posit.co/>
4. The Comprehensive R Archive Network. (n.d.). Retrieved from <https://cran.r-project.org/>
5. Welcome to Quarto. (n.d.). Retrieved from <https://quarto.org/>
6. Wickham, H., Navarro, D., & Pedersen, T. L. (2023). *ggplot2: Elegant Graphics for Data Analysis (3e)*. Retrieved from <https://ggplot2-book.org/>
7. Wickham, H., Çetinkaya-Rundel, M., & Grolemund, G. (n.d.). *R for Data Science (2e)*. Retrieved from <https://r4ds.hadley.nz/>
8. Сидоренко, В. М. (2022, April 1). *Data Science на R. Лабораторний практикум (draft version)*. Retrieved from <https://vgamaley.github.io/DS-book-lab/>