

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА

МЕХАНІКО-МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ

Кафедра теорії ймовірностей, статистики та актуарної математики



РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Дескриптивна статистика багатовимірних даних

для студентів

галузь знань	11 «Математика та статистика»
спеціальність	112 «Статистика»
освітній рівень	перший (бакалавр)
освітня програма	«Статистика»
вид дисципліни	вибіркова

Форма навчання	денна
Навчальний рік	2020/2021
Семестр	5
Кількість кредитів ECTS	3
Мова викладання, навчання та оцінювання	українська
Форма заключного контролю	іспит

Викладачі: професор Майборода Ростислав Євгенович, доктор.ф.-м.н., професор, професор кафедри теорії ймовірностей та актуарної математики

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

КИЇВ – 2020

Розробник: професор Майборода Ростислав Євгенович, доктор.ф.-м.н., професор, професор кафедри теорії ймовірностей та актуарної математики.

ЗАТВЕДЖЕНО
Зав. кафедри
кафедри теорії ймовірностей,
статистики та актуарної математики
Мішура Мішура Ю.С.

Протокол № 1 від 28.08.2020 р.

Схвалено науково-методичною комісією механіко-математичного факультету

Протокол від "31" 08 2020 року № 1

Голова науково-методичної комісії Олійник професор, д.ф.-м.н. Олійник А.С.
(підпис)

1. Мета дисципліни – оволодіння сучасними методами, теоретичними положеннями та основними застосуваннями описової статистики та візуалізації статистичних даних, зокрема дослідницького факторного аналізу, кореляційного аналізу, кластерного аналізу, ознайомлення з прикладами практичного застосування статистичних алгоритмів.

2. Попередні вимоги до опанування або вибору навчальної дисципліни:

1. *Знати:* основи математичного аналізу та лінійної алгебри.
2. *Вміти:* розв'язувати задачі елементарної математики.
3. *Володіти елементарними навичками:* роботи на персональному комп'ютері.

3. Анотація навчальної дисципліни:

Навчальна дисципліна «Дескриптивна статистика багатовимірних даних» є складовою освітньої програми підготовки фахівців за освітнім рівнем «бакалавр» галузі знань 11 Математика та статистика зі спеціальності 112 Статистика освітньої програми «Статистика».

Дана дисципліна входить до вибіркового блоку «Комп'ютерна статистика та аналіз даних». Дисципліна «Дескриптивна статистика багатовимірних даних» включає в себе теорію статистик середнього положення, розкиду, кореляції, їх алгебраїчні та статистичні властивості. Візуалізацію одновимірних даних на основі цих статистик та за допомогою гістограм. Аналіз статистичної інформації на географічній карті. Аналіз та візуалізацію кореляцій та простих регресійних залежностей, дослідницький факторний аналіз і зниження вимірності, кластерний аналіз. У курсі передбачено вивчення основ мови статистичного програмування R та технік застосування R для дескриптивного аналізу даних.

Викладається у 5 семестрі в обсязі 90 год. (3 кредити ECTS¹) зокрема: лекції – всього 28 год, практичні 14 год., самостійна робота – 46 год. У курсі передбачено 2 змістових модулі та виконання 6 індивідуальних самостійних завдань. Завершується дисципліна іспитом у п'ятому семестрі.

4. Завдання (навчальні цілі):

Формування здатності розв'язувати складні спеціалізовані математичні та статистичні задачі, що характеризуються комплексністю і невизначеністю умов і передбачає застосування теоретико-ймовірнісних і статистичних методів; набуття знань, умінь та навичок (компетентностей) на рівні новітніх досягнень у математиці та статистиці, відповідно до освітнього рівня «Бакалавр». Зокрема, професійне оволодіння компетентностями:

- 1) Здатність до абстрактного мислення, аналізу та синтезу.
- 2) Здатність застосовувати знання у практичних ситуаціях.
- 3) Знання й розуміння предметної області та професійної діяльності.
- 4) Здатність спілкуватися українською мовою як усно, так і письмово.
- 5) Здатність вчитися і оволодівати сучасними знаннями.
- 6) Здатність до пошуку, обробки та аналізу інформації з різних джерел.
- 7) Здатність приймати обґрунтовані рішення.
- 8) Здатність працювати автономно.
- 9) Визначеність і наполегливість щодо поставлених завдань і взятих обов'язків.
- 10) Здатність оцінювати та забезпечувати якість виконуваних робіт.
- 11) Здатність діяти на основі етичних міркувань (мотивів).
- 12) Здатність здійснювати логічні математичні міркування із чітким зазначенням припущень та висновків.
- 13) Здатність до математичного формулювання задач та вибору методів їх розв'язання.
- 14) Здатність до кількісно-статистичного мислення.

¹ кредитів ECTS – кредит кратний 30 годинам.

- 15) Здатність до ймовірнісного мислення, що передбачає сприйняття стохастичної природи явищ.
- 16) Здатність робити якісні висновки з кількісних даних.
- 17) Здатність проводити дослідження ймовірнісно-статистичних моделей та інтерпретувати одержані результати.
- 18) Здатність подавати статистичні процедури та результати їхнього застосування у формі, придатній для цільової аудиторії, до якої звертаються, як усно, так і письмово
- 19) Здатність до аналізу основ і властивостей статистичних алгоритмів та розуміння переваг тих чи інших підходів, у тому числі до оцінки їх обґрунтованості й ефективності.

5. Результати навчання за дисципліною:

Результат навчання (РН) (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 (Формуються розробником)			
РН 1.1	Знати основні статистики середнього положення та розкиду, їхні властивості та способи їх обчислення у системі R.	<i>Лекція, практичне заняття</i>	<i>Іспит, активна робота на лекції, усні відповіді, результати виконання самостійних індивідуальних завдань</i>	5%
РН 1.2	Знати методи візуалізації розподілу одновимірних даних.			5%
РН 1.3	Знати теорію коефіцієнтів кореляції Пірсона, Кендалла і Спірмена, засоби їх візуального відображення.			10%
РН 1.4	Знати методи аналізу простої лінійної регресії, їх реалізацію в R.			10%
РН 1.5	Знати теорію та застосування методу головних компонент.			10%
РН 2.1	Вміти реалізувати в R техніки зниження вимірності та візуалізації на основі методу головних компонент.	<i>Практичне заняття, самостійна робота</i>	<i>Розв'язання задач на практичних заняттях іспит, виконання завдань, винесених на самостійну роботу</i>	15%
РН 2.2	Вміти проводити ієрархічний кластер-аналіз в R.			15%
РН 2.3	Вміти відображати статистичні дані на географічній карті.			25%
РН 3.1	Здатність грамотно будувати комунікацію, виходячи з мети і ситуації спілкування	<i>Практичне заняття</i>	<i>активна робота практичних заняттях, усні відповіді</i>	2.5%
РН 3.2	Вироблення навиків командної роботи	<i>Практичне заняття</i>	<i>активна робота практичних заняттях, усні відповіді</i>	2.5%

6. Співвідношення результатів навчання дисципліни із програмними результатами навчання

Результати навчання дисципліни	PH 1.1	PH 1.2	PH 1.3	PH 1.4	PH 1.5	PH 2.1	PH 2.2	PH 2.3	PH 3.1	PH 3.2
Програмні результати навчання										
PH-1 - Здійснювати професійну письмову й усну комунікацію українською мовою та, принаймні, однією з іноземних мов	+	+	+	+	+	+	+	+	+	+
PH-11 - Вміти аналізувати та прогнозувати лінійні статистичні моделі та моделі регресії, оцінювати їхні параметри	+	+	+	+	+	+	+	+	+	+
PH-12 - Вміти збирати та обробляти дані, застосовувати статистичні процедури для аналізу даних за допомогою обчислювальної техніки та програмних засобів	+	+		+	+	+	+		+	+
PH-15 - Володіти математичними та статистичними методами аналізу, прогнозування та оцінки параметрів математичних моделей, статистичними методами інтерпретації та обробки числових даних	+		+			+	+			
PH-16 - Вміти використовувати в практичній діяльності спеціалізоване статистичне програмне забезпечення	+	+	+	+	+	+	+	+	+	+
PH-17 - Знати методи моделювання природничих та/або соціальних процесів			+	+	+					
PH-18 - Вміти застосовувати ймовірнісно-статистичні моделі та методи для розв'язання прикладних проблем і задач	+	+	+	+	+	+	+	+		

Відповідно до матриці ОП мають в таблиці Співвідношення результатів навчання дисципліни із програмними результатами навчання мають бути PH-1, PH-6, PH-8, PH-9, PH-13, PH-15, PH-16, PH-17, PH-20.

7. Схема формування оцінки.

7.1. Форми оцінювання студентів:

- оцінювання впродовж навчального періоду:

1. Активна робота на лекції, усні відповіді: PH1.1, PH1.2, PH1.3, PH1.4, PH1.5, PH2.1, PH2.2, PH2.3, PH3.1, PH3.2 – 6 балів/5 балів;
 2. Виконання індивідуальних самостійних завдань 1-2: PH1.1, PH1.2, PH3.1 – 18 балів/10 балів;
 3. Виконання індивідуального самостійного завдання 3: PH2.3, PH3.1, PH3.2 – 9 балів/5 балів;
 4. Виконання індивідуального самостійного завдання 4: PH1.3, PH1.4 – 9 балів/5 балів;
 5. Виконання самостійних індивідуальних завдань 5-6: PH2.1, PH2.2, PH3.1, PH3.2 – 18 балів/10 балів;
- Разом 60/35

- підсумкове оцінювання: іспит.

- максимальна кількість балів, які можуть бути отримані: 40 балів;

- результати навчання, які будуть оцінюватись: РН1.1, РН1.2, РН1.3, РН1.4, РН1.5, РН2.1, РН2.2, РН2.3
- форма проведення і види завдань: письмова робота з усним обговоренням.

7.2. Організація оцінювання:

Критично-розрахунковий мінімум балів за навчання впродовж семестру становить **20** балів, рекомендований мінімум, розрахований з урахуванням специфіки дисципліни становить **35** балів. Студенти, які протягом семестру набрали сумарно меншу кількість балів ніж рекомендований мінімум **35** балів для підвищення балів отримують можливість написати додаткову контрольну роботу та доскласти домашні завдання. Мінімальна кількість балів, які додаються до семестрових – 24 бали, тобто, якщо оцінка студента на іспиті є нижчою від мінімального порогового рівня (24 бали), то бали за іспит не додаються до семестрової оцінки (вважаються рівними нулю), а підсумкова оцінка із дисципліни є незадовільною;

У випадку відсутності студента з поважних причин відпрацювання та перездачі форм контролю здійснюються у відповідності до „Положення про організацію освітнього процесу в Київському національному університеті імені Тараса Шевченка” (2018), <http://www.univ.kiev.ua/pdfs/official/Organization-of-the-educational-process.pdf>.

Форма іспиту в кожному семестрі – письмово-усна. Екзаменаційний білет іспиту складається із 4 завдань, перші два з яких є теоретичними, одне – розрахункове і одне – по обговоренню результатів виконання індивідуальних самостійних завдань. Кожне завдання оцінюється від 0 до 10 балів. Додатково від 0 до 5 балів студент отримує за додаткові запитання на усному опитуванні. Всього за іспит можна отримати від 0 до 40 балів.

Терміни проведення форм оцінювання:

1. Виконання індивідуальних самостійних завдань 1-3: на 1-му - 6-му тижні навчального періоду.

2. Виконання індивідуальних самостійних завдань 4-6: на 7-му - 12-му тижні навчального періоду

7.3. Шкала відповідності оцінок

Відмінно / Excellent	90-100
Добре / Good	75-89
Задовільно / Satisfactory	60-74
Незадовільно / Fail	0-59

8. СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ. ТЕМАТИЧНИЙ ПЛАН ЛЕКЦІЙ І ПРАКТИЧНИХ ЗАНЯТЬ

теми	Назва теми	Кількість годин				
		Лекції	Практичні заняття	Самост. робота	Контр. модульна робота	Інші форми контролю

Змістовий модуль 1 „ Статистика одновимірних даних ”						
1	Дескриптивні статистики одновимірних даних	6	2	6		
2	Алгебраїчні, статистичні та оптимальні властивості дескриптивних статистик	6	2	8		
3	Візуалізація розподілів одновимірних даних	4	2	8		
Змістовий модуль 2 „ Статистика багатовимірних даних ”						
1	Регресійний та кореляційний аналіз	4	4	8		
2	Зниження вимірності даних на основі методу головних компонент	4	2	8		
3	Алгоритми кластерного аналізу	4	2	8		
Всього годин за I семестр		28	14	46		

Загальний обсяг 90 годин, у тому числі:
лекції – 28 годин,
практичні заняття – 14 годин,
консультації – 2 годин,
самостійна робота – 46 годин.

9. Рекомендовані джерела

Основні:

1. Майборода Р.Є. "Комп'ютерна статистика". – Київ, ВПЦ "Київський університет", 2019. - 589с.
2. Майборода Р.Є. "Методичні рекомендації по курсу "Дескриптивна статистика" для студентів III курсу".- 2020 , 34 с. Режим доступу: <http://probability.univ.kiev.ua/userfiles/mre/descrtasks.pdf>
3. Майборода Р.Є. Регресія: Лінійні моделі./ К.:ВПЦ «Київський університет», 2007.- 296 с.

Додаткові:

1. Гихман И.И., Скороход А.В., Ядренко М.И. Теория вероятностей и математическая статистика. – Киев, Выща школа, 1988. – 439 с.
2. Venables W.N., Ripley B.D. Modern Applied Statistics with S/ Springer, 2002.- 562 p.
3. Härdle W., Simar L. Applied Multivariate Statistical Analysis-Springer 2007.- 455p.