

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ СЕМЕНА КУЗНЕЦЯ**

ЗАТВЕРДЖЕНО

на засіданні кафедри
статистики і економічного прогнозування
Протокол № 2 від 02.09.2024 р.

ПОРОДЖЕНО

Проректор з навчально-методичної роботи

 Карина НЕМАШКАЛО

АНАЛІЗ ДАНИХ ТА БАГАТОВИМІРНА СТАТИСТИКА

робоча програма навчальної дисципліни (РПНД)

Галузь знань	всі
Спеціальність	всі
Освітній рівень	третій (освітньо-науковий)
Освітньо-наукова програма	всі

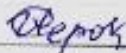
Статус дисципліни	вибіркова
Мова викладання, навчання та оцінювання	українська

Розробники:

д.е.н., професор
к.е.н., доцент

підписано КЕП

Олена РАЄВНЄВА



Ірина СЕРОВА

Завідувач кафедри
статистики і економічного
прогнозування

підписано КЕП

Олена РАЄВНЄВА

**Харків
2024**

ВСТУП

Соціально-економічні процеси характеризуються великою кількістю показників і являють собою багаторівневу систему зі складними взаємозв'язками між елементами. Для здійснення досліджень великих масивів інформації доцільно використовувати багатовимірні статистичні методи, що дозволяють аналізувати структуру взаємозв'язків між різноманітними показниками.

Моделюючи й аналізуючи взаємозв'язки між параметрами економічних систем, багатовимірні статистичні методи дозволяють оцінити стан і перспективи розвитку складних соціально-економічних процесів, вибрати обґрунтовані тактичні і стратегічні управлінські рішення, оцінити надійність і точність висновків, зроблених на підставі обмежених статистичних даних.

Навчальна дисципліна «Аналіз даних та багатовимірна статистика» досліджує різноманітні сучасні методи статистичного аналізу, що найчастіше використовуються в дослідницькій практиці, й формує практичні навички застосування цих методів на підставі пакетів прикладних програм.

Метою викладання навчальної дисципліни «Аналіз даних та багатовимірна статистика» є набуття теоретичних знань і практичних навичок роботи зі статистичною інформацією як основою наукових досліджень.

Завдання навчальної дисципліни полягають у вивченні сутності інструментів аналізу даних та багатовимірної статистики як основи роботи з наукою інформацією, в опануванні статистичними методами аналізу часових рядів, узагальнення та групування даних, структурного, кластерного, дискримінантного та факторного аналізів.

Об'єктом навчальної дисципліни є соціально-економічні явища та процеси, що відбуваються на макро-, мезо- та мікрорівнях.

Предметом вивчення навчальної дисципліни є методи попередньої обробки даних, статистичні та економіко-математичні методи групування, узагальнення та аналізу інформації наукового дослідження.

Результати навчання та компетентності, які формує навчальна дисципліна визначено в табл. 1.

Таблиця 1

Результати навчання та компетентності, які формує навчальна дисципліна

Результати навчання	Компетентності, якими повинен оволодіти здобувач вищої освіти
Використовувати параметричні та непараметричні критерії для рядів розподілу величин, визначати та усувати аномальні рівні часових рядів	Здатність формувати інформаційний простір дослідження різноманітних процесів підприємницької діяльності та працювати аномальними даними
Проводити дослідження закономірностей в рядах динаміки, визначати стійкість динамічного ряду	Здатність формувати систему показників оцінки та аналізу різних аспектів підприємницької діяльності в умовах повного та обмеженого доступу до інформації, флуктуаційних викидів та аномальних даних

Будувати одномірні та багатомірні групування та проводити їх аналіз	Здатність використовувати статистичні та економіко-математичні методи узагальнення та групування статистичної інформації
Визначати аналітичні показники простої та багатовимірної структури та структурних зрушень, аналізувати концентрацію, диференціацію та подібність структур різноманітних об'єктів, застосовувати на практиці кластерний, дискримінантний та факторний аналізи	Здатність здійснювати оцінювання та аналіз багатовимірних процесів підприємницької діяльності, формувати науково-інформаційне обґрунтування управлінських рішень

ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Зміст навчальної дисципліни

Змістовий модуль 1. Прості статистичні методи узагальнення та аналізу даних

Тема 1. Попередній аналіз статистичної інформації наукового дослідження.

1.1. Основні категорії аналізу часових рядів.

Поняття часового ряду та основна мета його аналізу. Методи та моделі вивчення складових часового ряду. Стаціонарні та нестаціонарні часові ряди. Способи отримання пропущених значень часового ряду. Вимоги до інформації, представленої в часовому ряду. Етапи проведення аналізу однорідності сукупності. Причини появи в сукупності аномальних спостережень. Методи виявлення та усунення аномальних спостережень.

1.2. Методи виявлення закономірностей динаміки досліджуваних явищ.

Характеристика інтенсивності динаміки на підставі системи базисних та ланцюгових показників. Цілі виявлення закономірності динаміки.

1.3. Методи вивчення та вимірювання стійкості рівнів ряду і тренду.

Показники стійкості та коливань динамічних рядів. Процентний розмах, середня процентна зміна, коефіцієнт стійкості, коефіцієнт Спірмена. Експоненційний та параболічний тренди, їхня характеристика.

Тема 2. Компонентний аналіз часового ряду.

2.1. Компоненти часового ряду.

Схема статистичного аналізу одновимірних часових рядів. Причини коливання рівнів часового ряду. Поняття компонентного аналізу: основна тенденція, тренд, закон розвитку. Проблеми, що виникають при вивченні часових рядів. Види трендів та їх графічне зображення. Методи розкладання часових рядів на компоненти.

2.2. Методи визначення наявності випадкової компоненти у часовому ряду.

Перевірка гіпотези про наявність тенденції у вихідному ряду. Моделі аналізу часових рядів: трендові моделі, моделі декомпозиції, моделі згладжування, авторегресійні й лагові моделі. Методи визначення наявності тренду в часовому ряду. Метод F-критерію Фішера, метод Форстера-Стюарта,

метод середніх. Послідовність кроків критерію серій, заснованого на медіані вибірки.

Тема 3. Статистичні методи узагальнення та групування даних.

3.1. Методи одномірного групування.

Складові статистичного зведення. Види зведення за різноманітними класифікаційними ознаками. Види групувань залежно від поставлених завдань. Прості та складні групування. Типи одновимірних групувань та основи їх побудови.

3.2. Методи багатомірного групування.

Види причинно-наслідкових зв'язків. Функціональні та стохастичні зв'язки. Побудова та аналіз комбінаційної матриці. Поняття та типи багатовимірних групувань. Метод багатовимірної середньої. Типи заходів подібності. Аналітичне групування, дисперсійний аналіз та їх практичне застосування в наукових дослідженнях.

Змістовий модуль 2. Методи структурного аналізу та кластеризації даних в економічних дослідженнях

Тема 4. Методи аналізу структури соціально-економічних даних.

4.1. Статистичний аналіз простої та багатовимірної структури економічних процесів.

Процеси та їх види. Поняття структури та основні напрямки її статистичного вивчення. Характеристика та приклади одномірної, багатомірної та ієрархічної структури. Розрахунок структурних показників.

4.2. Методи оцінки нерівномірності розвитку процесів. Аналіз концентрації, диференціації та подібності розподілу.

Спеціальні показники ступеня рівномірності розподілу. Крива Лоренца та її інтерпретація. Оцінювання відмінностей двох розподілів у просторі та часі. Коефіцієнти локалізації, диференціації, подібності структур. Показники структурних відмінностей.

4.3. Статистичний аналіз інтенсивності структурних зрушень.

Структурні зрушення в економіці та їх статистичне вимірювання. Індивідуальні показники абсолютних та відносних структурних зрушень. Індекси структурних зрушень. Відносні показники структурних зрушень.

Тема 5. Економіко-математичні методи дослідження багатомірних процесів.

5.1. Кластерний аналіз.

Сутність кластерного аналізу. Стандартизація та нормування. Поняття відстані. Методи кластерного аналізу.

5.2. Сутність дискримінантного аналізу та сфери його застосування.

Основні поняття дискримінантного аналізу. Дискримінантні функції. Критерії оцінки якості класифікації. Використання дискримінантного аналізу в економіці.

5.3. Поняття факторного аналізу та його мета.

Основні поняття факторного аналізу. Методи факторного аналізу. Метод головних компонент.

Перелік лабораторних занять / завдань за навчальною дисципліною наведено в табл. 2

Таблиця 2

Перелік лабораторних занять / завдань

Назва теми та / або завдання	Зміст
Тема 1. Завдання 1-3.	Завдання 1. Перевірка динамічних рядів на нормальність розподілу величин. Показники асиметрії та ексцесу. Завдання 2. Розрахунок параметричних та непараметричних критеріїв для рядів розподілу величин. Завдання 3. Визначення та усунення аномальних рівнів часового ряду. Дослідження закономірностей у часовому ряді.
Тема 2. Завдання 1-2.	Завдання 1. Згладжування часових рядів. Завдання 2. Побудова трендових моделей часових рядів.
Тема 3. Завдання 1-3.	Завдання 1. Побудова групувань за однією та двома ознаками. Завдання 2. Методика аналітичних групувань. Завдання 3. Розрахунок багатовимірної середньої.
Тема 4. Завдання 1-3.	Завдання 1. Способи побудови простої та багатовимірної структури. Завдання 2. Розрахунок індивідуальних показників структури та індексів структурних зрушень. Розрахунок абсолютних та відносних коефіцієнтів структурних зрушень. Завдання 3. Визначення показників структурних відмінностей.
Тема 5. Завдання 1-3.	Завдання 1. Використання кластерного аналізу для дослідження економічних процесів. Завдання 2. Вирішення завдання класифікації методом дискримінантного аналізу. Завдання 3. Побудова моделі факторного аналізу.

Перелік самостійної роботи за навчальною дисципліною наведено в табл. 3.

Таблиця 3

Перелік самостійної роботи

Назва теми та / або завдання	Зміст
Тема 1-5.	Виконання лабораторних завдань
Тема 1-5.	Виконання домашніх завдань у вигляді case study за темами дисципліни
Тема 1-5.	Підготовка до лабораторних занять та тестового контролю
Тема 1-5.	Підготовка до написання контрольної роботи

Кількість годин лекційних, лабораторних занять та годин самостійної роботи наведено в робочому плані (технологічній карті) з навчальної дисципліни.

МЕТОДИ НАВЧАННЯ

У процесі викладання навчальної дисципліни для набуття визначених результатів навчання, активізації освітнього процесу передбачено застосування таких методів навчання, як:

Словесні (лекція (Тема 1-5), проблемна лекція (Тема 2, 4), міні-лекція (Тема 1, 3, 5)).

Наочні (демонстрація (Тема 1-5)).

Практичні (лабораторна робота (Тема 1-5), есе (Тема 1-5), case study (Тема 1-5), тестовий контроль (тема 1-5)).

ФОРМИ ТА МЕТОДИ ОЦІНЮВАННЯ

Університет використовує 100 бальну накопичувальну систему оцінювання результатів навчання здобувачів вищої освіти.

Поточний контроль здійснюється під час проведення лекційних та лабораторних занять і має на меті перевірку рівня підготовки здобувачів ступеня доктора філософії до виконання конкретної роботи і оцінюється сумою набраних балів:

для дисциплін з формою семестрового контролю залік: максимальна сума – 100 балів; мінімальна сума – 60 балів.

Підсумковий контроль включає семестровий контроль та атестацію здобувача ступеня доктора філософії.

Семестровий контроль проводиться у формі диференційованого заліку.

Підсумкова оцінка за навчальною дисципліною визначається сумуванням всіх балів, отриманих під час поточного контролю.

Під час викладання навчальної дисципліни використовуються наступні контрольні заходи:

Поточний контроль: лабораторні заняття (20 балів), захист звіту з лабораторної роботи (30 балів), тестовий контроль (15 балів), домашнє завдання у вигляді case study (15 балів), письмова контрольна робота (20 балів).

Семестровий контроль: Диференційований залік.

Критерії поточного оцінювання знань здобувача ступеня доктора філософії (спеціальності):

Лабораторне заняття (максимальна кількість балів – 20 балів). Кожне лабораторне заняття оцінюється в 2 бала:

1 – виконання поставленого завдання;

1 – активність у вирішенні завдань.

Захист звіту з лабораторної роботи (максимальна кількість балів – 30). Кожний звіт з лабораторної роботи оцінюється в 6 балів:

6 – робота виконана і захищена в строк без помилок у розрахунках, надані

обґрунтовані висновки і правильні відповіді за темою роботи;

5 - робота виконана і захищена без помилок у розрахунках, але не в строк; надані обґрунтовані висновки і правильні відповіді за темою роботи;

4 – робота виконана та захищена з незначними помилками у розрахунках та висновках;

3 та менше – робота виконана з помилками у розрахунках та висновках, захист роботи відсутній;

0 – робота не виконана і незахищена.

Тестовий контроль (максимальна кількість балів – 15). Кожне тестове завдання оцінюється в 3 бали:

30% правильних відповідей – 1 бал;

60% правильних відповідей – 2 бали;

100% правильних відповідей - 3 бали.

Домашнє завдання (максимальна кількість балів – 15). Кожне домашнє завдання за темою оцінюється в 3 бали:

2 – правильність застосування аналітичного інструментарію при виконанні завдання, ілюстрація на конкретних практичних прикладах вміння вирішувати завдання наукового дослідження із застосуванням відповідних статистичних методів;

0,5 – правильна відповідь на запитання за домашнім завданням;

0,5 – своєчасність виконання домашнього завдання.

Письмова контрольна робота (максимальна кількість балів – 20):

19-20 – усі завдання вирішені вірно, зроблено обґрунтовані висновки щодо проаналізованої ситуації;

17-18 – усі завдання вирішені вірно, але здобувач припустився деяких неточностей у формулюванні економічних висновків;

15-16 – усі завдання вирішені вірно, але здобувач зробив не повні економічні висновки;

12-14 – усі завдання вирішені, але здобувач припустився незначних помилок у вирішенні завдань та формулюванні висновків;

9-11 – усі завдання вирішені, але здобувач припустився значних помилок у їх вирішенні та формулюванні висновків;

6-8 – завдання вирішені наполовину зі значною кількістю помилок, висновки відсутні;

0 – здобувач не з'явився на письмову контрольну роботу.

Більш детальну інформацію щодо системи оцінювання наведено в робочому плані (технологічній карті) з навчальної дисципліни.

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

Основна

1. Болюбаш Н. М. Інтелектуальний аналіз даних : навч. посіб. / Н. М. Болюбаш. – Миколаїв : Вид-во ЧНУ ім. Петра Могили, 2023. 320 с

2. Горкавий В. К. Статистика: Підручник. Третє вид., переробл. і доповн. Київ: Алерта, 2020. 644 с.

3. Malyarets, Lyudmyla, Babenko, Vitalina, Iastremska, Olena, Barannik, Igor. Assessment of the Level of Environmental and Economic Sustainability of Subjects of the Public Sector of the Country's Economy, *Lecture Notes in Networks and Systems*, 2023, 753, 3-11. Режим доступу: <http://repository.hneu.edu.ua/handle/123456789/30626>
4. Негрей М., Гнот Т. Аналітика з R: навчальний посібник / Негрей М., Гнот Т. Київ: ФОП Ямчинський О. В., 2020. 236 с.
5. Статистика [Електронний ресурс] : навч. посіб. / О. В. Раєвнева, І. В. Аксьонова, О. І. Бровко [та ін.]; за заг. ред. О. В. Раєвневої. Самостійне електрон. текстове мережеве вид. (389 с./3,59 МБ). Харків: ХНЕУ ім. С. Кузнеця, 2019. Режим доступу: <http://repository.hneu.edu.ua/handle/123456789/24523>
6. Статистичне моделювання та прогнозування: навчальний посібник/ Під ред. О. В. Раєвневої. Х.: ВД «Інжек», 2014. 578 с.
7. Статистичний аналіз даних : навчальний посібник / Т. М. Паянок, Т. М. Задорожня. Ірпінь : Університет державної фіскальної служби України, 2020. 312 с.
8. Willard, C.A. (2020) Statistical Methods: An Introduction to Basic Statistical Concepts and Analysis. 2nd Edition, Routledge, New York. Режим доступу: <https://doi.org/10.4324/9780429261039>
9. Prasad S. Advanced Statistical Methods. Springer Nature Singapore Pte Ltd, 2024. 218 p. Режим доступу: <https://doi.org/10.1007/978-981-99-7257-9>

Додаткова

10. Bruce P., Bruce A., Gedeck P. Practical Statistics for Data Scientists. 50+ Essential Concepts Using R and Python. 2nd Edition: O'Reilly Media, 2020. – 368 p.
11. Засоби підготовки та аналізу даних. Лабораторний практикум [Електронний ресурс] : навчальний посібник для студентів спеціальності 113 «Прикладна математика» / КПІ ім. Ігоря Сікорського ; уклад.: А. Ю. Шелестов, Н. М. Куссуль. – Електронні текстові дані (1 файл: 1,56 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2021. 31 с. Режим доступу: <https://ela.kpi.ua/handle/123456789/43491>
12. Rand R. Wilcox A Guide to Robust Statistical Methods: Springer Cham 2023. – 326 p. <https://link.springer.com/book/10.1007/978-3-031-41713-9>
13. Rayevnyeva, O.; Ponomarenko, V.; Matusova, S.; Stryzhychenko, K.; Filip, S.; Brovko, O. Models of the Impact of Socio-Economic Shocks on Higher Education Development. *Adm. Sci.* 2024, 14, 278. <https://doi.org/10.3390/admsci14110278> Scopus
14. O. Rayevnyeva, K. Stryzhychenko and O. Brovko, "The Investigation of the Long-Term and Short-Term Cointegration Effects between GDP and Migration," 2024 14th International Conference on Advanced Computer Information Technologies (ACIT), Ceske Budejovice, Czech Republic, 2024, pp. 116-121, doi: 10.1109/ACIT62333.2024.10712544 Scopus
15. Rayevnyeva O, Stryzhychenko K, Matúšová S. Impact of Migration Processes on GDP. *Engineering Proceedings*. 2023; 39(1):86. <https://doi.org/10.3390/engproc2023039086> Scopus

16. Scitovski R., Sabo K., Martínez-Álvarez F., Šime Ungar. Cluster Analysis and Applications: Springer Cham, 2021. – 271 p.
17. Економетрика з R : навчальний посібник / А.В. Скрипник, Д.М. Жерліцин, Ю.О. Нам'ясенко. Київ: ФОП Ямчинський О.В., 2020. 248 с.

Інформаційні ресурси

18. Загальнодоступна інформаційна база даних Національної комісії з цінних паперів та фондового ринку про ринок цінних паперів [Електронний ресурс] / Режим доступу: <https://stockmarket.gov.ua/>.
19. Інформаційна база даних емітентів. Режим доступу: www.Smida.gov.ua.
20. Основи статистики та аналізу даних: відкритий посібник з відкритих даних. Режим доступу: <https://socialdata.org.ua/manual/manual4/>.
21. Сайт Головного управління статистики в Харківській області. Режим доступу: <https://kh.ukrstat.gov.ua/>.
22. Сайт Державної служби статистики України. Режим доступу: www.ukrstat.gov.ua.
23. Офіційний сайт департаменту статистики Організації Об'єднаних Націй [Електронний ресурс]. Режим доступу: <http://unstats.un.org/unsd/default.htm>.