

Харківський національний університет імені В.Н. Каразіна
Факультет комп'ютерних наук
Кафедра теоретичної та прикладної системотехніки

УХВАЛЕНО
Вченою радою факультету
комп'ютерних наук, протокол № 4
від «03» грудня 2019 р.
Голова Вченої ради _____



Назва курсу	Діагностика комп'ютерних систем
Викладач (-і)	Проф. Угрюмов М.Л., доцент Бакуменко Н.С.
Профайл викладача (-ів)	http://www-csd.univer.kharkov.ua/about-us/sub-faculty/tase-department/personalnij-sklad/
Контактний тел.	
E-mail:	tps@karazin.ua
Сторінка курсу в системі дистанційного навчання	-
Консультації	Очні консультації: 6 годин, кожен парний вівторок з 15.20 до 16.20, ауд. 320

1. **Коротка анотація до курсу** – курс «Діагностика комп'ютерних систем» спрямований на опанування методами аналізу даних при управлінні складними комп'ютерними системами, оцінки їх ефективності, обробці даних та ін., для рішення яких розробляється математичне забезпечення комп'ютерних систем, а також використовуються спеціалізовані пакети прикладних програм, а також навичок по їх використанню..

2. **Мета та цілі курсу** – засвоєння студентами основ моделювання чисельних даних та методології статистичного аналізу, оцінювання взаємозв'язку величин, зниження розмірності даних та ін., вироблення навичок по адаптації стандартних алгоритмів до нових чисельних рішень складних прикладних задач, а також придбання знань про пакети прикладних програм спеціального призначення..

3. **Формат курсу** - очний

4. **Результати навчання** –

знати:

- основні цілі та вихідні передумови застосування статистичних методів при управлінні складними комп'ютеризованими системами;
- методи попередньої обробки даних;
- основні поняття вибіркового методу;
- методи перевірки статистичних гіпотез;
- характеристики статистичного зв'язку кількісних даних;
- моделі дисперсійного аналізу статистичних даних;
- методи кореляційного аналізу статистичних даних;
- методи регресійного аналізу;
- методи оцінювання інформативності (значущості) змінних при невизначеності даних;
- методи прогнозування часових рядів;

вміти:

- застосовувати знання у практичних ситуаціях
- використовувати сучасні методи і мови програмування для розроблення алгоритмічного та програмного забезпечення;
- ідентифікувати, класифікувати та описувати роботу програмно-технічних засобів, комп'ютерних та кіберфізичних систем, мереж та їхніх компонентів шляхом використання аналітичних методів і методів моделювання;
- аргументувати вибір методів розв'язування спеціалізованих задач, критично оцінювати отримані результати, обґрунтовувати та захищати прийняті рішення;
- виконувати експериментальні дослідження за професійною тематикою
- застосовувати знання для ідентифікації, формулювання і розв'язування технічних задач спеціальності, використовуючи методи, що є найбільш придатними для досягнення поставлених цілей;
- вибирати адекватні методи статистичного аналізу даних у відповідності з метою дослідження та характером статистичних даних
- знаходити чисельні характеристики статистичних розподілів вибірок даних;

- перевіряти основні гіпотези щодо параметрів розподілення даних;
- видаляти аномальні спостереження у скалярних та векторних даних;
- робити одно факторний та двофакторний дисперсійний аналіз даних;
- знаходити рівняння лінійної регресії;
- перевіряти значущість коефіцієнтів лінійної регресії;
- знаходити довірчі інтервали для коефіцієнтів лінійної регресії;
- будувати лінійну множинну регресію;
- знаходити рівняння нелінійної регресії;
- визначати статистичні оцінки для параметрів нелінійної регресії;
- оцінювати інформативність змінних з врахуванням точності вимірювання змінних стану і наявності парної кореляції між ними;
- оцінювати довірчі інтервали для математичного очікування нелінійних залежностей методом Монте-Карло;
- представляти змістовну інтерпретацію результатів статистичного аналізу;
- працювати з сучасними програмними системами статистичного аналізу даних (Statistica, MATLAB Statistic Toolbox, SPSS);
- придбати навички:
 - проведення експериментів, збирання даних та моделювання в комп'ютерних системах
 - формулювання змістовної та математичної постановок задач, здійснювання формалізації представлення даних, структуризації поставлених задач;
- засвоєння основних методів і прийомів аналізу та обробки різних видів інформації; розробки моделей та методів статистичного аналізу даних;
- проведення верифікації математичних методів, оцінки їх якості на основі перевірки існуючих статистичних гіпотез ;
- вирішення задач чисельного характеру з застосуванням спеціалізованих пакетів;

мати уявлення:

- про класичні і сучасні методи статистичного аналізу даних;
- про межу можливих застосувань методів статистичного аналізу даних.

5. Обсяг курсу

Вид заняття	Загальна кількість годин
лекції	24
практичні	24
самостійна робота	42

6. Ознаки курсу:

Рік викладання	семестр	спеціальність	Курс (рік навчання)	Нормативний/ вибірковий
2019	2	Комп'ютерна інженерія	4	вибірковий (В)

7. **Пререквізити** – перед вивченням курсу студент повинен знати вищу математику, основи теорії ймовірностей та математичної статистики, основи алгоритмізації та програмування, основи роботи з базами даних.

8. **Технічне та програмне забезпечення /обладнання** - Вказати, якщо є специфічні вимоги які студент повинен врахувати

9. Політики курсу

Академічна доброчесність - це сукупність етичних принципів та визначених законом правил, якими мають керуватися учасники освітнього процесу під час навчання, викладання та провадження наукової (творчої) діяльності з метою забезпечення довіри до результатів навчання та/або наукових (творчих) досягнень. Дотримання академічної доброчесності здобувачами освіти передбачає самостійне виконання навчальних завдань, завдань поточного та підсумкового контролю результатів навчання; посилаючись на джерела інформації у разі використання ідей, розробок, тверджень, відомостей.

10. Схема курсу

Тиж. / акад.год.	Тема, план, короткі тези	Форма діяльності (заняття)* *лекція, ПЗ,ЛР, СРС) / Формат** **аудиторна, СРС	Матеріали	Завдання, год
Розділ 1. Оцінювання невизначеностей при вимірюванні в статистичному дослідженні				
Тиж. 1 9 год.	Тема 1. Діагностування систем на основі даних моніторингу. Лекція 1. Діагностування як процес розпізнавання стану систем. Етапи процесу діагностування.	Лекція / аудиторна	Презентація лекції	2 год.

	Показники якості діагностування.			
	СРС. Інформаційно-аналітичне забезпечення процесів діагностування систем на основі даних моніторингу.	СРС / СРС	Конспект лекції, Інформаційна технологія діагностування складних технічних систем в умовах невизначеності вхідних даних [Текст]: монографія / В. Є. Стрілець, Е. М. Угрюмова та ін - Х.: Нац. аерокосм. ун-т ім. Н.Є. Жуковського «Харк. авіац. ін-т», 2015. –104 с. (с.5-57)	опрацювати лекцію, ознайомитись з літературою, 5 год.
	ПЗ 1. Знайомство з середовищем R Studio. Первинна обробка даних	Практичне заняття	Гнатюк В. Вступ до R на прикладах. Харків : ХНЕУ, 2010. – 101 с., Завдання на ПЗ 1	опрацювати лекцію, ознайомитись з літературою, 2 год
Тиж. 2 9 год.	Тема 2. Статистичне оцінювання параметрів розподілів. Лекція 2.Статистичні оцінки математичного очікування, факторної дисперсії по вибірці. Метод максимальної правдоподібності (М-оцінювання). Формулювання та перевірка гіпотез про рівність	Лекція / аудиторна	Презентація лекції	2 год

	центрів розподілів, рівність дисперсій.			
	CPC. Типи невизначеності: епістеміческа, алеаторна (параметрична). Форми обліку результатів спостереження. Типи та види робастного оцінювання.	<i>CPC</i>	Конспект лекції, Analyst Robust statistics. How Not to Reject Outliers. December 1989. Vol. 114. Режим доступу - https://www.aafco.org/proficiency-testing-program/CSPStatsDocs/Reports/HowNotToRejectOutliersAMC1989.pdf	опрацювати лекцію, ознайомитись з літературою, 5 год
	ПЗ 2. Перевірка гіпотези про вид розподілу ознаки	Практичне заняття.	Конспект лекції, завдання на ПЗ 2	опрацювати лекцію, ознайомитись з літературою, 2 год.
Тиж. 3 8 год.	<i>Тема 2.</i> Статистичне оцінювання параметрів розподілів. <i>Лекція 3.</i> Поняття довірчого інтервалу. Довірча ймовірність. Побудова довірчого інтервалу для математичного очікування при відомій, невідомій дисперсії. Побудова довірчого інтервалу для дисперсії.	<i>Лекція / аудиторна</i>	Презентація лекції	2 год.

	СРС. Довірчі та предикативні інтервальні прогнози	СРС	Конспект лекції, В.Є. Бахрушин Методи аналізу даних : навчальний посібник для студентів. Запоріжжя : КПУ, 2011. – 268 с. (с. 43- 73)	опрацювати лекцію, ознайомитись з літературою, 4 год.
	ПЗ 3. Статистичне оцінювання параметрів розподілів	Практичне заняття	Конспект лекції, завдання на ПЗ 3	опрацювати лекцію, ознайомитись з літературою, 2 год.
Тиж. 4 8 год.	Тема 3. Дисперсійний аналіз даних. Лекція 3. Постановка задачі однофакторного дисперсійного аналізу. Основні положення дисперсійного аналізу. Постановка задачі двофакторного дисперсійного аналізу. Метод головних компонент.	Лекція / аудиторна	Презентація лекції	2 год.
	СРС. Факторний аналіз.	СРС	Конспект лекції, В.Є. Бахрушин Методи аналізу даних : навчальний посібник для студентів. Запоріжжя : КПУ, 2011. – 268 с. (с. 43- 73)	Опрацювати лекцію, ознайомитись з літературою, 4 год.
	ПЗ 4. Дисперсійний аналіз даних (однофакторний)	Практичне заняття.	Конспект лекції, завдання на ПЗ 4	опрацювати лекцію, ознайомитись з літературою,

				2 год.
Тиж. 5 8 год	<i>Тема 4. Кореляційний аналіз даних. Коефіцієнт детермінації як універсальна характеристика ступеню тісноти статистичного зв'язку.</i>	<i>Лекція / аудиторна</i>	Презентація лекції	2 год.
	СРС. Дослідження лінійної залежності за допомогою парного коефіцієнта кореляції.	<i>СРС</i>	Конспект лекції	опрацювати лекцію, ознайомитись з літературою, 4 год.
	ПЗ 5. Дисперсійний аналіз даних (двофакторний)	Практичне заняття.	Конспект лекції, завдання на ПЗ 5	опрацювати лекцію, ознайомитись з літературою, 2 год.
Тиж. 6 8 год	Лекція 6. Кореляційне відношення. Множинні та часткові коефіцієнти кореляції. Рангова кореляція	<i>Лекція / аудиторна</i>	Презентація лекції	2 год.
	СРС. Конкордація ознак.	<i>СРС</i>	Конспект лекції	опрацювати лекцію, ознайомитись з літературою, 2 год
	ПЗ 6. Кореляційний аналіз даних	Практичне заняття.	Конспект лекції, завдання на ПЗ 6	опрацювати лекцію, ознайомитись з літературою, 2 год.
Розділ 2. Формальні математичні моделі систем і процесів				
Тиж. 7 8 год	<i>Тема 5. Множинний регресійний аналіз. Лекція 7. Регресійний аналіз. Модель лінійної регресії. Обмеження регресійної моделі - мультиколінеарність, гомоскедастичність.</i>	<i>Лекція / аудиторна</i>	Презентація лекції	2 год

	СРС. Основні методи побудови регресійних моделей.	СРС	О. О. Сергеев-Горчинський, Г. В. Іщенко Інтелектуальний аналіз даних: Комп'ютерний практикум. КПІ ім. Ігоря Сікорського. – 2018. – 73 с. (с. 55-61)	опрацювати лекцію, ознайомитись із літературою, 4 год
	ПЗ 7. Рангова кореляція.	Практичне заняття	Конспект лекції, завдання на ПЗ 7	опрацювати лекцію, ознайомитись з літературою, 2 год.
Тиж. 8 8 год	Тема 5. Множинний регресійний аналіз. Лекція 8. Модель логістичної регресії. Інтерпретація коефіцієнтів логістичної регресії. Некоректно поставлені завдання. Алгоритми, що регуляризують (робастні алгоритми): адаптивні, інваріантні.	Лекція / аудиторна	Презентація лекції	2 год.
	СРС. Методи регуляризації в задачах ідентифікації, апроксимації даних та прогнозування часових рядів.	СРС	Конспект лекції	опрацювати лекцію, ознайомитись із літературою, 4 год.
	ПЗ 8. Множинний регресійний аналіз (лінійна форма регресійної залежності).	Практичне заняття.	Конспект лекції	опрацювати лекцію, ознайомитись з літературою, 2 год.
Тиж. 9 8 год	Тема 6. Методи оцінювання інформативності (значущості) змінних при невизначеності даних. Лекція 9. Методи оцінюванні диференціальної інформативності з врахуванням точності	Лекція / аудиторна	Презентація лекції	2 год

	вимірювання змінних стану и наявність парної кореляції между ними. Методи розпізнавання образів: детерміністські (дискримінантного аналізу, багатовимірною шкалювання і логічні), ймовірносно-статистичні (методи Байеса, послідовного аналізу і оцінювання на основі теорії інформації). Стохастичний аналіз інформативності: індекси Соболя. Taguchi S / N Ratio.			
	СРС. Ймовірносно-статистичні (методи Байеса, послідовного аналізу і оцінювання на основі теорії інформації) методи оцінювання інформативності.	<i>СРС</i>	Конспект лекції, Черняк О.І., Захарченко П.В Інтелектуальний аналіз даних: К., 2014. — 599 с.	опрацювати лекцію, ознайомитись із літературою, 4 год
	ПЗ 9. Множинний регресійний аналіз (поліноміальна форма регресійної залежності).	Практичне заняття	Конспект лекції, завдання на ПЗ 9	опрацювати лекцію, ознайомитись з літературою, 2 год.
Тиж. 10 7 год	Лекція 10. Оцінювання інформативності на основі методів структурно-параметричного аналізу і синтезу регресійних моделей.	<i>Лекція / аудиторна</i>	Презентація лекції	опрацювати лекцію, переглянути презентацію, 2 год
	СРС. Методи послідовного аналізу варіантів, вагові з адаптацією, локально-стохастичні на основі самоорганізації.	<i>СРС</i>	Конспект лекції, Черняк О.І., Захарченко П.В Інтелектуальний аналіз даних: К., 2014. — 599 с.	опрацювати лекцію, ознайомитись із літературою, 4 год
	ПЗ 10. Оцінювання інформативності (значущості) змінних при невизначеності даних.	Практичне заняття	Конспект лекції, завдання на ПЗ 10	опрацювати лекцію, ознайомитись з літературою, 2 год.

Тиж. 11 8 год	Тема 7. Прогнозування часових рядів. Стаціонарність, автоковаріації і автокореляції. Основні описові статистики для часових рядів. Використання трендовій лінійної регресії з детермінованими чинниками для моделювання часового ряду. Прогнози по регресії з детермінованими чинниками.	<i>Лекція / аудиторна</i>	Презентація лекції	опрацювати лекцію, переглянути презентацію, 2 год
	СРС. Стабільне та ефективне оцінювання параметрів трендових регресійних моделей.	<i>СРС</i>	Конспект лекції	опрацювати лекцію, 4 год
	ПЗ 11. Прогнозування часових рядів.	Практичне заняття	Конспект лекції, завдання на ПЗ 11	опрацювати лекцію, ознайомитись з літературою, 2 год.
Тиж. 12 8 год	Лекція 12. Згладжування часового ряду. Авторегресійні трендові моделі. Розладнання часових рядів. Критерії тренду.	<i>Лекція Аудит.</i>	Презентація лекції	опрацювати лекцію, переглянути презентацію, 2 год
	СРС. Багатовимірні часові ряди. Багатовимірні трендові регресійні моделі. Ранг коінтеграції (розмірність простору коінтегрованих часових рядів).	<i>СРС</i>	Конспект лекції	опрацювати лекцію, 4 год
	Контрольна робота	<i>Контр. робота / аудиторна</i>	Варіанти завдань КР	Повторити матеріал розділів 1-2

***якщо література подається в скороченому вигляді, то розшифрування подаєте вкінці

11. Система оцінювання та вимоги

Загальна система оцінювання курсу.

Поточний контроль, самостійна робота, індивідуальні завдання					Залікова робота	Сума
Розділ 1	Розділ 2	Контрольна	Індивідуальн	Разом		

							робота, передбачена навчальним планом	є завдання			
T1	T2	T3	T4	T5	T6	T7					
5	10	10	5	10	5	5	10		60	40	100

T1, T2 ... – теми розділів.

Загальна система оцінювання курсу	<i>участь в роботі впродовж семестру/залік - 60/40</i> Розподіл балів, що присвоюються студентам з навчальної дисципліни, є сумою балів за виконання всіх видів завдань та самостійну роботу плюс бали, отримані під час екзамену. Впродовж семестру студент за виконання всіх завдань може отримати до 60 балів і до 40 балів під час складання заліку.
Вимоги до контрольної роботи	Завдання з контрольної роботи повинні бути виконані в повному обсязі. При оцінюванні роботи враховується повнота виконання, здатність студента формулювати закони та закономірності, структурувати судження, обґрунтовувати виконані операції, алгоритми, правила визначення понять, робити висновки, описи, тощо. Максимальна оцінка за контрольну роботу – 10 балів. Студент може скласти залік лише за наявності виконаних практичних завдань.
Умови допуску до підсумкового контролю	Умовою допуску студента до підсумкового семестрового контролю (екзамену) є наявність усіх виконаних практичних робіт та контрольної роботи. Згідно рішення кафедри теоретичної та прикладної системотехніки факультету комп'ютерних наук до екзамену не допускаються студенти, які мають заборгованість контрольним роботам. Наприкінці курсу всі бали підсумовуються і студент має можливість (в разі успішного виконання всіх завдань) отримати максимальну підсумкову оцінку 100 балів.

Шкала оцінювання

Сума балів за всі види навчальної діяльності протягом семестру	Оцінка	
	для екзамену	для заліку
90 – 100	відмінно	зараховано
70-89	добре	
50-69	задовільно	

12. Рекомендована література

Основна література

1. Олійник А. О. Інтелектуальний аналіз даних : навчальний посібник /А. О. Олійник, С. О. Субботін, О. О. Олійник. – Запоріжжя : ЗНТУ, 2012. – 278 с
2. Лапач С.М., Чубенко О.В., Бабіч П.М. Статистика в науці та бізнесі. - К.: МОРІОН, 2002. – 640 с.
3. Field, Andy (2009). Discovering Statistics Using SPSS. – 855 p.
4. Field, Andy (2012). Discovering Statistics Using R. – 995 p.
5. Аналіз даних / Н.С. Бакуменко, О.С. Радівоненко. - Навч. посібник з лабораторного практикуму.-Харків: Нац. аерокосм. ун-т «Харків. авіац. ін-т», 2007.– 87 с.

Додаткова література

1. Методологія наукових досліджень в медицині: навч. посібник / В. Д. Бабаджан, Н. С. Бакуменко, О. І. Кадикова та ін.; за ред. П. Г. Кравчуна, В. Д. Бабаджана, В. В. М'ясоєдова. – Харків : ХНМУ, 2020. – 260 с.
2. Розв'язання економічних завдань з використанням статистичних пакетів обробки даних: навч. посібник з лаб. практикуму/М.С. Мазорчук, Н.С. Бакуменко. - Харків: Нац. аерокосм. ун-т «Харків. авіац. ін-т», 2008. – 95 с.
3. Аналіз даних результатів статистичних досліджень у середовищі R [Текст]: навч. посібник / М. С. Мазорчук, В. С. Добряк, Н.С. Бакуменко. - Х.: Нац. аерокосм. ун-т ім. Н.Є. Жуковського «Харк. авіац. ін-т», 2018. – 36 с.
4. Статистичний аналіз даних у соціології [Текст]: навч. посібник / М. С. Мазорчук, В. С. Добряк. - Х.: Нац. аерокосм. ун-т ім. Н.Є. Жуковського «Харк. авіац. ін-т», 2012. – 36 с.