

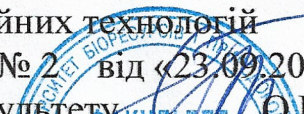
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

Факультет інформаційних технологій
Кафедра економічної кібернетики

«ЗАТВЕРДЖУЮ»

Проректор з науково-педагогічної
роботи та розвитку

 М. Кваша
« 2021 р.

РОЗГЛЯНУТО І СХВАЛЕНО
на засіданні Вченої ради факультету
інформаційних технологій
Протокол № 2 від «23.09.2021»
Декан факультету  О.Г. Глазунова

на засіданні кафедри економічної
кібернетики
Протокол № 2 від «16.09.2021»
Завідувач кафедри  Д.М. Жерліцин

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ
МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА ПЛАНУВАННЯ
ЕКСПЕРИМЕНТУ

РІВЕНЬ ВИЩОЇ ОСВІТИ: ТРЕТІЙ (ОСВІТНЬО-НАУКОВИЙ)

ГАЛУЗЬ ЗНАНЬ : 12 «Інформаційні технології»,

РОЗРОБНИК: доктор економічних наук, професор А. В. Скрипник

Київ 2021

1. Опис навчальної дисципліни

Математичне моделювання та планування експерименту

Галузь знань, спеціальність, освітній ступінь		
Галузь знань	«Інформаційні технології»,	
Освітньо-науковий рівень	Третій	
Освітній ступінь	PhD (доктор філософії)	
Спеціальність		
Освітньо-наукова програма		
Характеристика навчальної дисципліни		
Вид	дисципліна циклу спеціальної (фахової) підготовки	
Загальна кількість годин	90	
Кількість кредитів ECTS	3	
Кількість змістових модулів	Не передбачено	
Курсовий проект (робота)	Не передбачено	
Форма контролю	залік	
Показник навчальної дисципліни для денної та заочної форми навчання		
	денна форма навчання	заочна форма навчання
Рік підготовки (курс)	1	1
Семестр	1	1
Лекційні заняття	10 год	10 год
Практичні, семінарські заняття	20 год	20 год
Лабораторні заняття		
Самостійна робота	60 год	60 год
Індивідуальні завдання		
Кількість тижневих аудиторних годин для денної форми навчання	4	4

2. Мета та завдання навчальної дисципліни

Метою навчальної дисципліни «Математичне моделювання та планування експерименту» є формування у здобувачів **вищої освіти ступеня доктора філософії** комплексу здатностей до дослідження й розкриття

механізму процесів, виявлення сутності, взаємозв'язків і причин виникнення процесів/явищ для одержання певних наукових/прикладних результатів і розв'язання поставлених завдань шляхом використання методів регресійного, кореляційного та ін. видів аналізу даних для перевірки висунутих гіпотез і розробки різних видів моделей.

Предметом дисципліни «Математичне моделювання та планування експерименту» є вивчення ключових положень із статистики, теорії ймовірностей, економетрики.

Основними **компетентностями**, яких повинен набути здобувач після вивчення дисципліни є:

здатність працювати з великими обсягами інформації з використанням алгоритмів Big Data, Data Mining;

– здатність до оцінки обсягів інформації інформації, яка необхідна для отримання обґрунтованих висновків за темою дослідження;

– здатність здійснювати інтервальні оцінки випадкових величин та їх характеристик;

– здатність використати кількісні методи для оцінок ризиків інформаційної та мережевої безпеки;

– здатність використовувати сучасні інформаційні методи (нейромережи, панельну регресію, логіт та пробіт моделі) для детального аналізу стану інформаційного простору;

– набути компетентності, що дозволять користуватися на кількісному рівні поняттями оптимізації діяльності інформаційних систем при припустимому ступені ризику що досягається за рахунок диверсіфікації;

– демонструвати здатність до пошуку, обробки та аналізу великого масиву інформації з різних джерел;

– бути спроможним досліджувати процеси або явища шляхом створення стохастичних моделей.

Основними організаційними формами вивчення дисципліни «Математичне моделювання та планування експерименту» є: лекції, практичні заняття, самостійна робота та контрольні заходи – опитування, тести, залік.

Самостійна робота є основним засобом засвоєння аспірантами навчального матеріалу в час, вільний від аудиторних занять. Вона передбачає опрацювання монографій та інших наукових праць, підручників, навчальних і методичних посібників тощо, активне використання різнобічних вітчизняних і світових інформаційних джерел.

Форма підсумкового контролю – оцінка за результатами поточного та підсумкового контролю, яка здійснюється в сертифікованому електронному курсі. Залік є підсумковою формою контролю засвоєння знань і формування вмінь, передбачених внаслідок вивчення дисципліни.

3. Структура навчальної дисципліни

- повного терміну денної (заочної) форми навчання

Назви змістових модулів і тем	Кількість годин											
	денна форма						Заочна форма					
	усього	у тому числі					усього	у тому числі				
		л	п	л а б	Інд	с.р .		л	п	л і н	і н .	с.р .
Тема 1. Організація і планування експерименту при розробці інформаційних та інформаційно-управляючих систем з використанням алгоритмів Big Data, Data Mining.	14	2	4			8	14	2	4			8
Тема 2. Розподіли випадкових величин, що використовуються	14	2	4			8	14	2	4			8

для оцінки інформаційної безпеки.											
Тема 3. Оцінка обсягів інформації для отримання рівня значимості безпеки інформаційних мереж	26	2	4			20	26	2	4		20
Тема 4. Основні поняття інформаційної безпеки: «інформаційний ризик», «інформаційна небезпека», інформаційна загроза», та кількісні методи їх оцінювання.	20	2	4			14	20	2	4		14
Тема 5. Стохастичне моделювання інформаційної небезпеки: метод головних компонент, нейромереж, панельної регресії.	16	2	4			10	16	2	4		10
Усього годин	90	10	20			60	90	10	20		60

4. Теми практичних занять

№ п/п	Назва теми	Кількість годин
1	2	3
1.	Тема 1. Розподіли випадкових величин, що використовуються для оцінки інформаційної безпеки	4
2.	Тема 2. Інформаційна ентропія Шеннона та її практичне застосування.	4
3.	Тема 3. Оцінка обсягів інформації для отримання надійних оцінок інформаційної безпеки	4
4.	Тема 4. Логіт і пробіт моделювання для оцінки ймовірності реалізації хакерських атак	4
5.	Тема 5. Оцінка ступеня небезпеки інформаційних мереж методами нейромереж та панельної регресії.	4
Разом		20

5. Самостійна робота

№ з/п	Назва теми	Кількість годин	
		денна форма	заочна форма
1.	Статистика успішних хакерських атак за останнє десятиріччя	8	8
2.	Стан безпеки українських інформаційних мереж на державному рівні	10	10
3.	Стан безпеки приватних мереж українських компаній	12	12
4.	Технології дослідницької роботи в інформаційних технологіях	8	8
5.	Особливості обробки великих масивів даних	12	12

6.	Сучасні програмні засоби аналізу та обробки інформації, що використовуються в інформаційних технологіях	10	10
	Разом	60	60

6. Анотація тем навчальної дисципліни

Тема 1. Організація і планування експерименту при розробці інформаційних та інформаційно-управляючих систем з використанням алгоритмів Big Data, Data Mining. Методики отримання та обробки великих масивів інформації. Генератор випадкових чисел. Інформаційна ентропія Шеннона та її практичне застосування, перехресна ентропія, дивергенція, фундаментальне попереднє скорочення інформації для передачі з відомим розподілом ймовірностей.

Питання для самопідготовки за темою 1

1. Оцініть ентропію українського інформаційного простору?
2. Що таке перехресна ентропія?
3. Як пов'язані обсяг інформації та її ентропії?
4. Як зростають обсяги та ентропія глобальної системи інформації?

Тема 2. Розподіли випадкових величин, що використовуються для оцінки інформаційної небезпеки. Розподіли Фішера, Гаусса, χ^2 , Пуассона.

Використання критерію χ^2 для оцінок приналежності емпіричного розподілу до деякого теоретичного розподілу. Прикладне використання розподілів в наукових дослідженнях з інформаційних технологій.

Питання для самопідготовки за темою 2

1. Як створюється розподіл χ^2 ?
2. Що вважати великим масивом і як забезпечити достатню точність оцінки математичного очікування?
3. Чому страхові компанії не використовують поняття математичного очікування?

4. В яких випадках можна використати нормальний розподіл без перевірки?

5. Наведіть приклади використання рівномірного розподілу в своїх дослідженнях ?

8. Похибки використання правила трьох сігм для великих масивів інформації.

9. Якщо розподіл не є відомим, як оцінити ймовірність великих відхилень від математичного очікування?

Тема 3. Оцінка обсягів інформації для отримання рівня значимості безпеки інформаційних мереж. Поняття точності оцінки різних характеристик вибірки. Оцінка частки p з заданою точністю, оцінка інших характеристик p з заданою точністю.

Питання для самопідготовки за темою 3

1. На підставі яких розподілів будуються оцінки?
2. Чим відрізняються підходи в оцінках точності вибірки в соціології та інформаційних технологіях?
3. Який потрібен обсяг вибірки при похибці 1 %, на рівні значимості 0,05 для оцінки частки?

Тема 4. . Основні поняття інформаційної безпеки: «інформаційний ризик», «інформаційна небезпека», інформаційна загроза», та кількісні методи їх оцінювання. Логіт і пробіт моделювання для оцінки ймовірності реалізації хакерських атак в сучасному інформаційному суспільстві

Питання для самопідготовки за темою 4

1. Як визначається загальний ризик інформаційної безпеки.
2. Кількісні оцінки окремих типів інформаційної небезпеки.
3. Якщо випадки реалізації небезпеки відбуваються рідко, то який розподіл можна використати для їх реалізації?

4. В якому випадку слід використати біноміальний розподіл?

5. Яку характеристику слід використати замість математичного очікування для кількості випадків реалізації інформаційної небезпеки.

6. Чім відрізняються логіт і пробіт моделі оцінки ймовірності

Тема 5. Стохастичне моделювання інформаційної небезпеки: метод головних компонент, нейромереж, панельної регресії. Основні переваги та недоліки методу головних компонент, особливості використання нейромереж для прогнозування кількості випадків реалізації інформаційної небезпеки. Випадки аналізу рознесених у просторі і часі подій за допомогою панельної регресії.

Питання для самопідготовки за темою 5

1. Яка інформація потрібна для реалізації методу головних компонент?
2. В чому перевага використання нейромереж?
3. Які головні переваги панельної регресії і як їх можна застосувати для аналізу кібербезпеки? Математичне очікування та дисперсія бінарного процесу.
4. Який розподіл використовується при оцінці обсягу вибірки?

7. Індивідуальна робота

Індивідуальні завдання для проміжного контролю знань здобувачів **вищої освіти ступеня доктора філософії** з дисципліни «Математичне моделювання та планування експерименту» виконують відповідно до навчального плану. Аспірантам пропонується спланувати експеримент, що може буде використано для проведення дослідження, або проаналізувати вже існуючу кількісну інформацію що має відношення до тематики дослідження. Результати таких досліджень доповідаються на практичних заняттях.

Метою виконання індивідуального завдання є закріплення і систематизація отриманих знань у процесі самостійної підготовки у міжсесійний період.

7. Контрольні питання, комплекти тестів для визначення рівня засвоєння знань здобувачами зі спеціальності «Інформаційні технології»

1. Оцініть ентропію українського інформаційного простору?
2. Що таке перехресна ентропія?
3. Як пов'язані обсяг інформації та її ентропія?
4. Як зростають обсяги та ентропія глобальної системи інформації?
5. Як створюється розподіл χ^2 ?
6. Що вважати великим масивом і як забезпечити достатню точність оцінки математичного очікування?
7. Чому страхові компанії не використовують поняття математичного очікування?
8. В яких випадках можна використати нормальний розподіл без перевірки?
9. Наведіть приклади використання рівномірного розподілу ?
10. Похибки правила трьох сігм для великих масивів інформації?
11. Якщо розподіл не є відомим, як оцінити ймовірність великих відхилень від математичного очікування?
12. На підставі яких розподілів будуються оцінки?
13. Чим відрізняються підходи в оцінках точності вибірки в соціології та інформаційних технологіях?
14. Який потрібен обсяг вибірці при похибці 1 %, на рівні значимості 0,05 для оцінки частки?
15. Як визначається загальний ризик інформаційної безпеки.
16. Як робляться кількісні оцінки окремих типів інформаційної небезпеки.

17. Якщо випадки реалізації небезпеки відбуваються рідко, то який розподіл можна використати для їх реалізації?
18. В якому випадку слід використати біноміальний розподіл?
19. Яку характеристику слід використати замість математичного очікування для кількості випадків реалізації інформаційної небезпеки.
20. Чим відрізняються логіт і пробіт моделі оцінки ймовірності?
21. Яка інформація потрібна для реалізації методу головних компонент?
22. В чому перевага використання нейромереж?
23. Які головні переваги панельної регресії і як їх можна застосувати для аналізу кібербезпеки? Математичне очікування та дисперсія бінарного процесу.
24. Який розподіл використовується при оцінці обсягу вибірки?

Загальні питання для здобувачів третього рівня вищої освіти з усіх галузе знань

Приклади питання з яких складається тест для здобувачів освітньо-наукової програми Інформаційні технології

Позначення

***n* - кількість спостережень, *k* - кількість входів**

Питання

Відповіді

Як виглядає панельна лінійна регресія з двома входами?

$$y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 x_2 + \varepsilon$$

Чим відрізняються регресійні коефіцієнти β_1 і $\hat{\beta}_1$?

Величина β_1 не є відомою, тоді як $\hat{\beta}_1$ отримуємо після процедури МНК

Який розподіл має $\frac{\beta_j - \hat{\beta}_j}{\sigma_{\beta_j}}$?

Розподіл Стюдента з $n - (k + 1)$ ступенями свободи

Які параметри впливають на довірчий інтервал прогнозу?	Стандартна похибка, горизонт прогнозування
Яка інформація потрібна для логіт та пробіт моделей інформаційної безпеки.	Часовий ряд реалізації випадків інформаційної безпеки
На який інтервал можна прогнозувати рівень інформаційної безпеки	На середній інтервал зміни інформаційних технологій
Як з окремих показників інформаційної безпеки отримати Інтегральний показник	Шляхом множення ймовірності реалізації окремих станів інформаційної безпеки
Існує два часових ряди. Як поділити їх на причину (x) та наслідок (y)?	За допомогою взаємно-кореляційної функції. Ендогенна зміна повинна мати часовий лаг відносно екзогенної.
Що означає P-value в спеціалізованих програмах обробки інформації?	Рівень значимості
В чому переваги та недоліки зростання інформаційної ентропії Шенону	Зростання ентропії свідчить о наявності багато джерел, однак збільшує час пошуку інформації
Два з трьох понять інформаційної безпеки мають рівень значимості 0,1, який рівень значимості повинен мати третій параметр щоб загальний рівень значимості був менш за 10^{-6}	10^{-4}

10. Рекомендована література

Основна навчальна література

1. Єріна А. М. Статистичне моделювання та прогнозування: Навч. посібник. К: КНЕУ, 2014. 340 с.
2. Заєць С. В., Томіленко В. М. Статистика. Ірпінь: Видавництво Національного університету ДПС України, 2015. 512 с. (текст парал.: укр., англ. мовами).
3. Інформаційна та кібербезпека: соціотехнічний аспект: підручник / [В. Л. Бурячок, В. Б. Толубко, В. О. Хорошко, С. В. Толюпа]; за заг. ред. д-ра техн. наук, професора В. Б. Толубка.— К.: ДУТ, 2015.— 288 с.
4. 2. Корченко А.Г. Анализ и оценивание рисков информационной безопасности / А.Г. Корченко, А.Е. Архипов, С.В. Казмирчук // Монография. – К.: ООО «Лазурит-Полиграф», 2013. – 275 с.
5. 3. Корченко А.Г. Системы анализа и оценивания рисков информационной безопасности / А.Г. Корченко, А.Е. Архипов, С.В. Казмирчук. – К. : Palmarium Academic Publishing, 2013. – 316 с.
6. Скрипник А. В. Методичні вказівки для дисципліни «Математичні моделі та планування експерименту» для аспірантів очної, заочної форми навчання. Київ, 2018. 193с.
7. Скрипник А. В., Галаєва Л. В., Коваль Т. В., Шульга Н. Г. Навчальний посібник «Математична статистика». К.: ВЦ «Компринт», 2018. 380с.
8. Скрипник А., Негрей Н. Економетрика : навчальний посібник. Київ, ЦП «Компринт», 2016. 320 с.
9. Про схвалення Концепції розвитку цифрової економіки та суспільства України на 2018-2020 роки та затвердження плану заходів щодо її реалізації від 17 січня 2018 р. № 67-р <https://zakon.rada.gov.ua/laws/show/67-2018-%D1%80>

10. Green W. Econometric analysis// NY University, Person Education Limitation. 2012, 1233 p.

11. Wooldridge M. Introductory Econometrics a Modern Approach//2013 South-Western/ Canada. 869 p.

Додаткові джерела

1. Скрипник А.В., Герасимчук Н.. Економічні і фінансові ризики / Житомир: Видавництво ЖДУ, 2013. 371 с.

2. Barrow M. Statistics for economics, accounting and business studies. 2016 Harlow, 384 p.

3. Floyd J. Statistics for Economist: Beginning. Toronto 2020, 218 p.

4. Gadium A., Pannell D. Risk attitudes and risk perceptions of crops producers in Western Australia Babcock, B.A., R.W. Fraser, J.N. Lekakis, Amsterdam:, 2020. 113–134 с.

5. Hoag D.L. Applied Risk Management in Agriculture / D.L. Hoag, Washington DC: CRS Press, 2010. 380 с.

6. UNCTAD. The Least Developed Countries Report 2020. CHAPTER 4: Transition to the digital economy: technological capabilities as drivers of productivity [Електронний ресурс]. – Режим доступу: https://unctad.org/system/files/official-document/ldcr2020_ch4_en.pdf.

7. World Bank (2016). World Development Report 2016: Digital Dividends. Washington, DC. [Електронний ресурс]. – Режим доступу: <https://www.worldbank.org/en/publication/wdr2016>

8. World Bank (2018a). World Development Report 2019: The Changing Nature of Work. Washington DC. [Електронний ресурс]. – Режим доступу: <https://www.worldbank.org/en/publication/wdr2019>