



ПРИКЛАДНА СТАТИСТИКА

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Перший (бакалаврський)</i>
Галузь знань	<i>12 Інформаційні технології</i>
Спеціальність	<i>122 КОМП'ЮТЕРНІ НАУКИ, 124 СИСТЕМНИЙ АНАЛІЗ</i>
Освітня програма	<i>Системний аналіз і управління, Системи і методи штучного інтелекту</i>
Статус дисципліни (код)	<i>Вибіркова</i>
Форма навчання	<i>очна(денна)/дистанційна/змішана</i>
Рік підготовки, семестр	<i>3 курс, осінній семестр</i>
Обсяг дисципліни	<i>4 кредити ЄКТС</i>
Семестровий контроль/ контрольні заходи	<i>Залік</i>
Розклад занять	<i>http://roz.kpi.ua/</i>
Мова викладання	<i>Українська</i>
Інформація про керівника курсу / викладачів	Лектор: доктор філософії, старший викладач <i>Левенчук Людмила Борисівна</i> levenchuk.liudmyla@iit.kpi.ua ORCID: http://orcid.org/0000-0002-8600-0890 Практичні: доктор філософії, старший викладач <i>Левенчук Людмила Борисівна</i>
Розміщення курсу	<i>Googleclassroom</i>

Програма навчальної дисципліни

1. Опис навчальної дисципліни, її мета, предмет вивчення та результати навчання

Прикладна статистика – це наука збору, обробки даних і аналізу отриманих результатів з метою прийняття коректних обґрунтованих рішень технічного, ділового, політичного персонального або іншого характеру.

Ця дисципліна передбачає використання статистичних методів для відповіді на практичні питання в різних галузях, таких як фінанси, охорона здоров'я, маркетинг та інженерія.

Прикладна статистика є однією з фундаментальних дисциплін в освітній програмі бакалаврів з комп'ютерних наук та системного аналізу, оскільки абсолютна більшість бакалаврських дипломів аналізує часові ряди фінансових і виробничих процесів.

Дана дисципліна включає:

- вивчення та оволодіння методами збору та обробки статистичних даних у формі часових перерізів рядів;*
- вивчення та оволодіння методами побудови статистичних і математичних моделей;*
- вивчення критеріальної бази для дослідження якості даних і адекватності створених моделей;*
- уміння аналізувати як структура даних впливає на розв'язання задачі статистичного моделювання, прогнозування та формування статистичних висновків.*
- уміння перевіряти відповідність вхідних статистичних даних постановці задачі статистичного і математичного моделювання на основі отриманих даних;*

- уміння розробляти статистичні моделі досліджуваних процесів в економіці та фінансах;
- уміння доводити коректність та ефективність побудованих статистичних і математичних моделей на розв'язуванні прикладних задач;
- уміння порівнювати різні методи моделювання, прогнозування, оцінювання ризиків та формування статистичних висновків.

Вивчення навчальної дисципліни націлено на формування, розвиток та закріплення у здобувачів таких загальних та фахових **компетентностей**:

ЗК 1 – Здатність до абстрактного мислення, аналізу і синтезу.

ЗК 2 – Здатність застосувати знання в практичних ситуаціях.

ЗК 3 – Знання та розуміння предметної області та розуміння професійної діяльності.

ЗК 6 – Здатність вчитися й оволодівати сучасними знаннями.

ЗК 7 – Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

ЗК 9 – Здатність працювати в команді.

ЗК 10 – Здатність бути критичним та самокритичним.

ЗК 12 – Здатність оцінювати та забезпечувати якість виконуваних робіт.

ЗК 14 – Здатність реалізовувати свої права та обов'язки як члена суспільства, усвідомлювати цінності громадянського (вільного демократичного) суспільства та необхідність його сталого розвитку, верховенства права, прав і свобод людини і громадянина в Україні.

ЗК 15 – Здатність зберігати та примножувати моральні, культурні, наукові цінності і досягнення суспільства на основі розуміння історії та закономірностей розвитку предметної області системного аналізу, її місця у загальній системі знань про природу і суспільство та у розвитку суспільства, техніки і технологій, використовувати різні види та форми рухової активності для активного відпочинку та ведення здорового способу життя.

ФК 4 – Здатність визначати основні чинники, які впливають на розвиток фізичних, економічних, соціальних процесів, відокремлювати в них стохастичні та невизначені фактори, формулювати ці фактори у вигляді випадкових або нечітких величин, векторів, процесів та досліджувати залежність між ними.

ФК 9 – Здатність представляти математичні аргументи і висновки з них з якістю і точністю в таких формах, які проходять для занять в аудиторіях як усно, так і в письмовій формі.

ФК 10 – Здатність виконувати експериментальні та спостережувані дослідження і аналізувати дані, отримані в них.

Внаслідок вивчення курсу студент повинен бути здатний продемонструвати такий **програмний результат навчання ОПП**:

ПРО3 - Вміти визначати ймовірнісні розподіли стохастичних показників та факторів, що впливають на характеристики досліджуваних процесів, досліджувати властивості та знаходити характеристики багатовимірних випадкових векторів та використовувати їх для розв'язання прикладних задач, формалізувати стохастичні показники та фактори у вигляді випадкових величин, векторів, процесів.

ПРО6 - Знати та вміти застосовувати основні методи постановки та вирішення задач системного аналізу в умовах невизначеності цілей, зовнішніх умов і конфліктів, ПРО9 - Вміти створювати ефективні алгоритми для обчислювальних задач системного аналізу та систем підтримки прийняття рішень.

ПР12 - Застосовувати методи і засоби роботи з даними і знаннями, методи математичного, логіко-семантичного, об'єктного та імітаційного моделювання, технології системного і статистичного аналізу.

ПР13 - Проектувати, реалізовувати, тестувати, впроваджувати, супроводжувати,

експлуатувати програмні засоби роботи з даними і знаннями в комп'ютерних системах і мережах.

ПР14 - Розуміти і застосовувати на практиці методи статистичного моделювання і прогнозування, оцінювати вихідні дані.

ПР15 - Розуміти українську та іноземну мови на рівні, достатньому для обробки фахових інформаційно-літературних джерел, професійного усного і письмового спілкування, написання текстів за фаховою тематикою.

ПР16 - Розуміти і реалізовувати свої права і обов'язки як члена суспільства, усвідомлювати цінності вільного демократичного суспільства, верховенства права, прав і свобод людини і громадянина в Україні, дотримуватися академічної доброчесності.

ПР17 - Зберігати та примножувати досягнення і цінності суспільства на основі розуміння місця предметної області у загальній системі знань, використовувати різні види та форми рухової активності для ведення здорового способу життя.

2. Пререквізити та постреквізити дисципліни (місце в структурно-логічній схемі навчання за відповідною освітньою програмою)

Вибірковий модуль прикладна статистика вимагає від студентів комплексу знань, умінь та навичок, отриманих при вивченні дисциплін:

Дисципліні ПС передуює кредитний модуль «Теорія ймовірностей», «Математична статистика» передуює.

Дисципліна ПС забезпечує наступні навчальні дисципліни у програмі підготовки фахівця: «Теорія інформації і кодування» (ПО 6), «Теорія випадкових процесів» (ПО 10), «Стаціонарні випадкові процеси» (ПО 13), «Теорія прийняття рішень» (ЗО 13), «Основи системного аналізу» (ЗО 12), «Моделювання складних систем» (ПО 15), «Аналіз часових рядів» (ПО 8) Системний аналіз стохастичних розподілених процесів» (ПО 16).

3. Зміст навчальної дисципліни

Кредитний модуль включає теми, подані нижче.

РОЗДІЛ 1. Вступні зауваження. Етапи збору та обробки статистичних даних.

Тема 1.1. Основні типи статистичних моделей для формального опису статистичних даних.

Типи статистичних моделей. Вимоги до статистичних даних, оцінок параметрів моделей і моделей у цілому. Мета обробки статистичних даних. Формування статистичного висновку на основі результатів обробки даних.

Тема 1.2. Чотири етапи аналізу статистичних даних.

Методи збору даних. Підготовка даних для побудови моделей. Нормування, фільтрація, виявлення екстремальних значень та їх обробка. Заповнення пропусків даних. Критерії якості даних. Побудова статистичної моделі та аналіз її адекватності за допомогою множини критеріїв.

РОЗДІЛ 2. Розрахунок та аналіз описових статистик. Прийняття статистичних рішень, перевірка гіпотез.

Тема 2.1. Опис положення спостережень у ряду розподілу, групування даних.

Процентильні ранги і процентилі, z-оцінки і T-оцінки. Групування даних, групова таблиця частот. Статистичні параметри згрупованих даних. Квартілі, децилі, кумулятивна крива. Дисперсія і стандартне відхилення згрупованих даних.

Тема 2.2. Прийняття статистичних рішень, ризик і перевірка гіпотез.

Послідовність статистичної перевірки гіпотез. Перевірка гіпотез у задачах трьох типів; односторонні і двосторонні критерії. Альтернативне формулювання процедури перевірки гіпотез. Оцінювання надійності правил перевірки статистичних гіпотез.

Тема 2.3. Регресійний аналіз статистичних даних.

Попередня підготовка даних і оцінювання структури регресійної моделі. Провідні індикатори та їх використання. Методи оцінювання параметрів регресійних моделей. Переваги і недоліки. Оцінювання адекватності лінійних регресійних моделей. Поняття розв'язку регресійного рівняння, аналіз його збіжності.

Тема 2.4. Нелінійна регресія – модель двійкового вибору. Часові перерізи даних.

Формування структури та оцінювання параметрів нелінійної моделі. Метод Монте-Карло для марковських ланцюгів та його практичне застосування. Приклади застосування регресійних моделей.

РОЗДІЛ 3 Прогнозування на основі статистичних моделей.

Тема 3.1. Умовні та безумовні статистичні характеристики процесів. Прогнозування без знаходження розв'язків різницевих рівнянь.

Прогнозування динаміки розвитку процесів за допомогою рівнянь низького порядку. Побудова функцій прогнозування, аналіз статистичних характеристик похибок оцінок прогнозів.

Тема 3.2. Критерії якості оцінок прогнозів.

Навчальна і тестова вибірки; перехресна валідація. Коефіцієнт Тейла і середня абсолютна похибка у процентах, інші статистичні характеристики. Багатокрокові прогнози та їх аналіз.

Тема 3.3. Прогнозування ймовірнісного розподілу даних – аналітичний підхід.

Прогнозування розподілу за регресійною моделлю. Формування прогнозного розподілу за теоремою Байєса. Застосування методу Монте-Карло для марковських ланцюгів для прогнозування ймовірнісного розподілу даних.

Самостійна робота. Вивчення дисципліни крім лекцій включає такі види СРС: підготовка до аудиторних занять; виконання та оформлення завдань комп'ютерного практикуму, підготовка до захисту виконаних завдань комп'ютерного практикуму.

4. Навчальні матеріали та ресурси

Базова:

1. Бідюк П.І., Терентьев О.М., Просьянкіна-Жарова Т.І. Прикладна статистика. – Київ: НТУУ «КПІ», 2012. – 510 с.
2. Gelman A., Carlin J.B., Stern H.S., Rubin D.B. Bayesian Data Analysis. – New York: CRC Press, 2004. – 670p.
3. Бідюк П.І., Меньяйленко О.С., Половцев О.В. Методи прогнозування. – Луганськ: Альма-матер, 2008. – 305 с.
4. Бідюк П.І., Харрінгтон Т., Ткач Б.П. Математична статистика. – Київ: Видавничий Дім «Персонал», 2017. – 348 с
5. Бідюк П.І., Романенко В.Д., Тимошук О.Л. Аналіз часових рядів. – Київ: «Політехніка», 2012. – 608 с.
6. Грін Г. Економетричний аналіз. – Київ: «Основи», 2010. – 1154 с.
7. Коваленко І.П. Математична статистика. – Київ: Видавничий Дім «Слово», 2012. – 496 с.
8. Pole A., West M., Harrison J. Applied Bayesian forecasting and time series analysis. – London: Chapman & Hall/CRC, 1994. – 409 p.
9. Gooijer J.G. Elements of nonlinear time series analysis and forecasting. – Berlin: Springer, 2017. – 620 p.
10. Shumway R.H., Stoffer D.S. Time series analysis and its applications. – New York: Springer, 2006. – 588 p.
11. Commandeur J.J.F., Koopman S.J. State space time series analysis. – Oxford: Oxford University Press, 2007. – 189 p.
12. Small M. Applied nonlinear time series analysis. – Singapore: World Scientific Publishing, 2005. – 261 p.

Допоміжна:

1. Коваленко І.П. Математична статистика у прикладах і задачах. – Київ: Слово, 2012. – 490 с.
2. Згуровський М.З., Бідюк П.І., Терентьев О.М., Просянкіна-Жарова Т.І. Байєсівські мережі у системах підтримки прийняття рішень. – Київ: Едельвейс, 2014. – 300 с.
3. Кузнецова Н.В., Бідюк П.І. Теорія і практика аналізу фінансових ризиків: системний підхід. – Київ: Ліра-К, 2020. – 400 с.
4. Бідюк П.І., Тимошук О.Л., Коваленко А.Є., Коршевнюк Л.О. СППР: проектування і реалізація. – Київ: НТУУ «КПІ ім. Ігоря Сікорського», 2020. – 605 с.
5. Рассел Стюарт, Норвіг Пітер. Искусственный интеллект: современный подход, 2-е изд.; Издательский дом «Вильямс», 2007. – 1408 с.
6. Karatzas I., Shreve S.E. Methods of financial mathematics. – New York: Springer, 1998. – 408 p.
7. Press S.J. Subjective and objective Bayesian statistics. – Hoboken, New Jersey: John Wiley & Sons, Inc., Publication, 2003. – 560 p.

Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

У гугл-класі будуть надані завдання з детальними інструкціями та необхідним матеріалом, які необхідно вчасно виконувати.

Студентам рекомендується вести докладний конспект лекцій і фіксувати основні результати лабораторних занять.

Важливим аспектом якісного засвоєння матеріалу, відпрацювання прийомів і алгоритмів вирішення основних завдань дисципліни є самостійна робота, що дає можливість перетворити отримані знання в об'єкт власної творчої діяльності. Самостійна робота включає в себе читання літератури, виконання огляду літератури за вибраною тематикою, виконання звітів по завданнях комп'ютерного практикуму та підготовку до їх захисту, а також підготовка до написання модульної контрольної роботи та складання заліку.

Крім того, важливою частиною процесу навчання є активне використання отриманих знань у практичних завданнях та проектах. Самостійна робота також може включати участь у групових обговореннях, виконання додаткових завдань для розширення розуміння матеріалу, а також пошук і аналіз додаткових джерел інформації з питань, пов'язаних з дисципліною.

Для організації та проведення занять з комп'ютерного практикуму розроблено Методичні вказівки по дисципліні:

Прикладна статистика: Навчально-методичний посібник. Призначений для здобувачів ступеня бакалавра спеціальностей 122 Комп'ютерні науки, 124 Системний аналіз.. Укладач: Бідюк П.І., Данилов В.Я., Жиров О.Л. – Київ: 2023. – 186 с.

Політика та контроль

6. Політика навчальної дисципліни (освітнього компонента)

Усі роботи студенти мають прикріплювати в особистому кабінеті гугл-класу. Дедлайни кожного завдання позначені в кожному завданні у гугл-класі. Роботи мають бути виконані з дотриманням академічної доброчесності. Політика та принципи академічної доброчесності, етична поведінка студентів визначені у Кодексі честі <https://kpi.ua/code>.

7. Види контролю та рейтингова система оцінювання результатів навчання (PCO)

Семестровий контроль: залік

Поточний контроль: фронтальний (усний, письмовий), завдання комп'ютерного практикуму.

Рейтингова система оцінювання включає всі види тестування: контрольну роботу, виконання комп'ютерного практикуму та якість захисту завдань практикуму. Кожний студент отримує свій підсумковий рейтинг з дисципліни.

Рейтинг студента з кредитного модуля складається з балів, які він отримує за:

- написання модульної контрольної роботи;
- робота на практичних заняттях та захист завдань комп'ютерного практикуму;
- відповіді на заліку.

Табл. 1. Система рейтингових (вагових) балів:

Метод оцінювання		Максимальна оцінка в балах
ЗКП	1	20
	2	30
	3	40
МКР		10
Залік		40
Підсумковий рейтинг	$(ЗКП+МКР)*0,6+Залік$	100

Умовою атестації з календарного контролю є:

- Перша атестація – задача та захист першої лабораторної роботи.
- Друга атестація – задача трьох лабораторних та захист двох.

Умовою допуску до заліку є:

- виконання і захист запланованих лабораторних робіт,
- написання модульної контрольної роботи.

Після оцінювання відповідей здобувача на заліку викладач підсумовує стартові бали $(ЗКП+МКР)*0,6$ балів та бали за залік (40 балів), зводить до рейтингової оцінки (оцінювання результатів навчання здійснюється за 100-бальною шкалою) та переводить до оцінок за університетською шкалою (Таблиця відповідності рейтингових балів оцінкам за університетською шкалою).

Критерії нарахування балів за контрольні заходи:

- ~“відмінно”: 95 -100% - здобувач виявив всебічні, систематичні та глибокі знання навчального матеріалу з дисципліни; продемонстрував уміння вільно виконувати всі завдання, передбачені програмою; засвоїв основну та додаткову літературу; проявив творчі здібності в розумінні, в логічному, чіткому, стислому та ясному трактуванні навчального матеріалу; засвоїв взаємозв'язок основних понять дисципліни, їх значення для подальшої професійної діяльності
- ~“дуже добре”: 85-94% - здобувач виявив систематичні знання навчального матеріалу з дисципліни вище середнього рівня; продемонстрував уміння добре виконувати всі завдання, передбачені програмою, допустивши незначні помилки; засвоїв основну та додаткову літературу; засвоїв взаємозв'язок основних понять дисципліни, їх значення для подальшої професійної діяльності
- ~“добре”: 75-84% - здобувач виявив загалом добрі знання навчального матеріалу при виконанні передбачених програмою завдань, але припустив ряд помітних помилок; засвоїв основну літературу; показав систематичний характер знань з дисципліни; здатний до їх

самостійного використання та поповнення в процесі подальшої навчальної роботи і професійної діяльності

~ "задовільно": 65-74% - здобувач виявив знання основного навчального матеріалу з дисципліни в обсязі, необхідному для подальшого навчання та майбутньої професійної діяльності; ознайомився з основною літературою; впорався з виконанням завдань, передбачених програмою, але припустив значну кількість помилок або недоліків на запитання при співбесіді, тестуванні та при виконанні завдань тощо, принципи з яких може усунути самостійно

~ "достатньо": 60-64% - здобувач виявив знання основного навчального матеріалу з дисципліни в мінімальному обсязі, необхідному для подальшого навчання та майбутньої професійної діяльності; ; ознайомився з основною літературою; в основному виконав завдання, передбачені програмою, але припустив помилки у відповіді на запитання при співбесідах, тестуванні та при виконанні завдань тощо, які він може усунути лише під керівництвом та за допомогою викладача

~ "незадовільно": 30-59% - здобувач мав значні прогалини в знаннях основного навчального матеріалу; допускав принципові помилки при виконанні передбачених програмою завдань, але спроможний самостійно доопрацювати програмний матеріал і підготуватися для перездачі дисципліни

~ "незадовільно": 0-29% - здобувач не мав знань зі значної частини навчального матеріалу з дисципліни; допускав принципові помилки при виконанні більшості передбачених програмою завдань або не виконував ці завдання

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

Бали	Оцінка
100...95	Відмінно
94...85	Дуже добре
84...75	Добре
74...65	Задовільно
64...60	Достатньо
Менше 60	Незадовільно
Не виконано умови допуску до екзамену	Не допущено

Додаткова інформація з дисципліни (освітнього компонента)

- можливість зарахування сертифікатів проходження дистанційних чи онлайн курсів за відповідною тематикою.

Інформаційні ресурси

- Електронний кампус Національного технічного університету України ім.Ігоря Сікорського» <https://ecampus.kpi.ua/>

Робочу програму навчальної дисципліни (силабус) склала:

Старший викладач кафедри математичних методів системного аналізу, доктор філософії **Левенчук Людмила Борисівна**

Ухвалено кафедрою ММСА (протокол № 13 від 05.06. 2024)

Погоджено Методичною комісією НН ІПСА (протокол № 10 від 24.06.2024)