

Scientific journal
PHYSICAL AND MATHEMATICAL EDUCATION
 Has been issued since 2013.

ISSN 2413-158X (online)
 ISSN 2413-1571 (print)

Науковий журнал
ФІЗИКО-МАТЕМАТИЧНА ОСВІТА
 Видається з 2013.

<http://fmo-journal.fizmatsspu.sumy.ua/>



Чемерис О.А., Прус А.В. Статистико-ймовірнісна складова змісту підготовки фахівців з інформаційних технологій. Фізико-математична освіта. 2020. Випуск 1(23). Частина 2. С. 83-88.

Chemeris O., Prus A. Statistical and probability component of the content of training of information technology specialists. Physical and Mathematical Education. 2020. Issue 1(23). Part 2. P. 83-88.

DOI 10.31110/2413-1571-2020-023-1-2-013
 УДК 519.25:51-7

О.А. Чемерис

Житомирський державний університет імені Івана Франка, Україна
 olgachemerys@i.ua
 ORCID: 0000-0002-7099-1095

А.В. Прус

Житомирський державний університет імені Івана Франка, Україна
 pruswork@gmail.com
 ORCID: 0000-0002-8869-2544

СТАТИСТИКО-ЙМОВІРНІСНА СКЛАДОВА ЗМІСТУ ПІДГОТОВКИ ФАХІВЦІВ З ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

АНОТАЦІЯ

Формулювання проблеми. Вивчення елементів статистики у професійній підготовці майбутніх фахівців з комп'ютерних технологій сприяє реалізації прикладної спрямованості навчання. У статті наведено приклади використання математичних та статистичних методів аналізу для конкретних досліджень та дано методичні рекомендації, зокрема, наведено приклади обчислення й подання числових характеристик вибірки та знаходження залежностей між елементами варіаційних рядів.

Матеріали і методи. Аналіз науково-методичної літератури з математичної статистики, узагальнення та теоретичне моделювання; визначення основних статистичних характеристик вибіркової сукупності (за допомогою офісних програм середовища GeoGebra); регресійний аналіз результатів дослідження для встановлення взаємозв'язку між даними; комп'ютерні технології для визначення числових характеристик.

Результати. У статті розкрито питання сучасних статистичних досліджень для різних галузей, результати яких було використано при викладанні змістового модуля «Математична статистика» для майбутніх фахівців з інформаційних технологій. Аналіз даних допомагає виявити корисну інформацію та зробити правильні висновки. Застосування комп'ютерних пакетів дозволить легше та швидше одержати потрібний й результат. Широкий огляд методів обчислень завдань зі статистики дозволяє набутти досвіду щодо практики обчислень в цій дисципліні. Наведено приклади розв'язання задач подальшою візуалізацією як початкових, так і кінцевих даних. Використано програмні середовища MS Excel та GeoGebra.

Висновки. Власний досвід навчання студентів дав можливість сформулювати деякі рекомендації для викладання змістового модуля «Математична статистика» та вести елементи аналізу даних, з якими стикаються фахівці інформаційних технологій у практичній діяльності. Огляд сучасних статистичних пакетів дозволить переглянути методи закріплення інформації та урізноманітнити практичне вивчення основних питань. У подальших публікаціях плануємо сформулювати методичні поради для застосування багатофакторного аналізу на прикладі авторських задач.

КЛЮЧОВІ СЛОВА: аналіз даних, описова статистика, ящик з вусами, коефіцієнт кореляції, пряма регресії.

ВСТУП

Постановка проблеми. Сьогодні вивчення математики є актуальним, оскільки все більше спеціальностей потребують застосувань як загальних математичних знань, так і спеціальних практичних навичок і умінь. Роботодавці та інші стейкхолдери на полегливо вимагають від працівників знань теорії ймовірностей, прикладної математики, статистики, основ математичного аналізу, математичної логіки тощо.

Впродовж останніх років спостерігається тенденція щодо збільшення попиту на кваліфікованих ІТ-фахівців, адже в структурі будь-якого підприємства спостерігаємо залучення від одного та більше працівників цієї сфери діяльності. Сьогодні маємо біля 150 закладів вищої освіти в Україні, які сприяють підготовці фахівців з комп'ютерних наук. Зокрема, в Житомирському державному університеті імені Івана Франка реалізується освітньо-професійна програма «Сучасні інформаційні технології та програмування» для підготовки здобувачів першого рівня вищої освіти спеціальності 122 Комп'ютерні науки (Ленчук, 2019). Програма має прикладний характер, орієнтована на формування на уково-технічного світогляду майбутнього фахівця, здатного розв'язувати спеціалізовані завдання, зокрема, із застосування інформаційних

технологій. Під «інформаційними технологіями» розуміють сукупність методів, виробничих і програмно-технологічних засобів, об'єднаних у технологічний ланцюжок, що забезпечує збирання, зберігання, обробку, висновки і поширення інформації, вони призначені для зниження трудомісткості процесів використання інформаційних ресурсів (Спірін, 2005).

Серед фахових компетенцій, якими має оволодіти випускник бакалаврату є СК 2 – здатність до виявлення закономірностей певних явищ, застосування методів в статистичній обробці даних, яка повністю може бути реалізована при вивченні дисципліни «Теорія ймовірностей, ймовірнісні процеси та математична статистика» (Ленчук, 2019).

Вивчення елементів статистики у професійній підготовці майбутніх фахівців з комп'ютерних технологій сприяє реалізації прикладної спрямованості навчання, зокрема, розвиваються вміння щодо аналізу випадкових подій, їх результативності та оцінки ймовірностей, висування гіпотез, прогнозування ситуацій та їх оптимального розв'язання (Пасічник, 2011). Важливо мати огляд застосування методів в різних областях, адже сучасні аналітичні та статистичні методи застосовуються в найрізноманітніших областях: в економіці, фінансах, соціології, психології, лінгвістиці, біології, екології, хімії, промисловості тощо, історично статистичним методам притаманне взаємозбагачення взаємопроникнення знань.

Актуальність дослідження. Серед науковців, які не лише збирали і класифікували дані, а аналізували їх та описували закономірності, згадуємо бельгійського вченого Ламбера Кетле, який одним з перших показав, що випадковості, які спостерігаються в живій природі, внаслідок їхньої повторюваності виявляють певну тенденцію, яку можна описати мовою математики (Гмурман, 2003). Біометричні методи Френсиса Гальтона і Карла Пірсона стали основою математичної статистики (Конет, 2010). Розподіли Стьюдента та Фішера-Снедекора використовують для визначення довірчих інтервалів невідомих параметрів розподілу та перевірки критеріїв значимості тощо.

Окреслимо питання сучасних статистичних досліджень для різних галузей, результати яких ми використали при викладанні змістового модуля: огляд програмних засобів статистичного аналізу даних (Роїк М. В., Присяжнюк О. І., Денисюк В. О.) (Роїк&Денисюк&Присяжнюк, 2017); застосування прикладних програм статистичного аналізу щодо результатів наукових досліджень (Мунтян О. А., Мунтян М. Л., Яровенко А. Г.) (Мунтян&Мунтян&Яровенко, 2017); огляд сучасних методів та технологій статистичного аналізу даних у гуманітарних науках (Кислова М. А., Савельєв В. В., Середа Ю. В.) (Кислова, 2017; Савельєв, 2018; Середа, 2013). Метою аналізу даних є виявлення корисної інформації та правильність висновків. Аналіз даних може мати багато аспектів та підходів, реалізовуватися за допомогою різних інструментів – у тому числі математичних, статистичних, за допомогою різноманітних способів візуалізації, але головне його завдання – укріплення інформації.

Мета статті – зробити огляд основних методів в аналізі статистичних даних; надати методичні рекомендації щодо практичної підготовки майбутніх фахівців з інформаційних технологій при опануванні змістового модуля «Математична статистика».

МЕТОДИ ДОСЛІДЖЕННЯ

Аналіз науково-методичної літератури, систематизація та узагальнення, аналіз та порівняння, теоретичне моделювання; визначення основних статистичних характеристик вибіркової сукупності; регресивний аналіз результатів; комп'ютерні технології.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Прикладна частина змістового модулю «Математична статистика» присвячена вивченню основних методів аналізу статистичних даних, які найчастіше застосовуються в дослідженнях та на практиці, а також формуванню практичних навичок застосування цих методів за допомогою комп'ютерних програм аналізу даних.

Перелік статистичних пакетів, які використовують професіонали, досить великий, наприклад: SAS, BMDP, STADIA, STATGRAPHICS, SPSS, Statistica; BIOSTAT, MESOSAUR, DATASCOPE тощо (Роїк, 2017), для навчання та кожні рекомендують Minitab, MatLab, STATA, S-PLUS, BMDP, DATASCOPE тощо.

Лідером з можливістю проведення якісної обробки даних є статистичний пакет SPSS (Statistical Package for the Social Sciences), він має вбудований інтелектуальний засіб для пропозиції щодо вибору статистики для початкових даних. Програма створює графіки та діаграми з можливістю копіювання та подальшого використання. На жаль еплатним, хоча є демоверсія для навчання. StatSoft Statistica — рекомендують для статистичних досліджень будь-якої складності, це надійний засіб для обробки даних, який гарно подає 3D графіки та ілюстрації; є безкоштовна демоверсія. Для аналізу даних застосовують також спеціальні мови програмування, зокрема, R, Python, приклади наводимо на практичних заняттях в спеціальних комп'ютерних класах (Майборода, 2020).

Microsoft Excel містить деякі статистичні функції, що є вбудованими формулами, але розрахунки за ними не визнаються авторитетними науковими журналами, та кожне можливо якісно візуалізувати інформацію. Але MS Excel та GeoGebra добре підходять для введення даних, попередніх статистичних обчислень, для побудови деяких видів діаграм. При встановленні вбудованого модуля даних, який є в усіх останніх версіях Microsoft Excel, можна проводити остаточний статистичний аналіз.

В усіх типах дослідження найчастіше використовується описова статистика, методи якої дозволяють відшукати найбільш типові значення змінної (мода, медіана, середнє арифметичне) та продемонструвати, як сильно розкидані значення відносно типового (розмах вибірки, середнє квадратичне відхилення), тобто кількісно підсумувати та оцінити масив даних. Статистичні графіки надзвичайно різноманітні, їх часто використовують для зображення варіаційних рядів: гістограма (сходиноковий графік для інтервального варіаційного ряду, чітко видно моду), полігон (варіаційний ряд у вигляді ламаної лінії), огіва, кумулята (криві на графіках підсумків в частот, добре визначають медіану), діаграми розсіювання (наочно бачимо залежність між змінними), ящик з вусами (компактно подає інформацію про вибірку, застосовують для порівнянь), матричні графіки (для описання залежності між багатьма змінними), обличчя Чернова (схема візуального подання багатоваріантних даних у вигляді людського обличчя) тощо. Ефективними у статистичних дослідженнях є комп'ютерні засоби побудови графічних зображень у спеціальних пакетах.

Підсумування за п'ятьма значеннями, або п'ятиточкова характеристика вибірки – це вид описової статистики, що надає інформацію про набір спостережень, і, як видно із самої назви – складається із п'яти важливих показників: (мінімальне значення в спостереженні; перший (або нижній) квартиль; медіана (середнє значення); третій (або верхній) квартиль; максимальне значення).

Нагадаємо визначення понять, якими будемо оперувати:

- розмах (R) вибірки – це різниця між максимальним і мінімальним значеннями елементів вибірки:

$$R = x_{\max} - x_{\min};$$

- квартилі Q_1, Q_2, Q_3 – це значення досліджуваної характеристики, як відтинають послідовно чверть, половину та тричверті кількості менших за величину значень: Q_1 – нижній (перший квартиль), Q_2 – медіана (M_e (другий квартиль), Q_3 – верхній (третій квартиль);

- інтерквартильний розмах IKR – це різниця між верхнім та нижнім квартилями: $IKR = Q_3 - Q_1$;

- мода M_o – це значення, що найчастіше зустрічається у вибірці.

П'ятиточкову характеристику подають як коробчасту діаграму, який називають "ящик з вусами" віданглійського терміну "box-and-whisker plot", або "boxplot" (Роїк, 2017). Діаграма була представлена Джоном Тьюкі в 1970-х роках з подальшою модифікацією. Цей рисунок є прямокутником із перетинкою і сторонами, що паралельні осям координат, та відрізками, паралельними горизонтальній осі, які відходять від середин бічних сторін прямокутника. Прямокутник із відрізками розташовується так, щоб бічні сторони проєктувались на точки горизонтальної осі, що є квартилями (довжина коробки), а кінці горизонтальних відрізків проєктуються на точки, що є мінімальним та максимальним елементами вибірки. Вуса – це значення, які лежать, відповідно, лівіше та правіше від першого та третього квартиля. Викиди – це значення, які знаходяться поза межами інтервалу $1,5 \cdot IKR$.

Розглянемо наступну задачу та побудуємо для даних у умові «box-and-whiskers plot».

Задача 1. Знайдіть усі необхідні величини для побудови "ящика з вусами" та зобразіть його. Елементами вибірки є бали за іспит студентів групи фізико-математичного факультету з дисципліни «Програмування»: 63, 75, 86, 66, 85, 94, 65, 74, 74, 65, 63, 63, 75, 86, 94, 82, 63, 74.

Розв'язання. Перш, ніж побудувати коробкову діаграму, потрібно знайти найменше та найбільше значення вибірки, а також усі три квартилі. Для цього сформуємо групувану вибірку, для побудови якої спочатку запишемо варіаційний ряд (варіанти x_i розташуємо в порядку зростання та вкажемо їх частоти появи n_i):

x_i	63	65	66	74	75	82	85	86	94
n_i	4	2	1	3	2	1	1	2	2

Об'єм вибірки $N = \sum_i n_i = 18$, $x_{\min} = 63$, $x_{\max} = 94$. Отже, розмах вибірки: $R = x_{\max} - x_{\min} = 94 - 63 = 31$.

Далі визначимо три квартилі для цієї вибірки. Оскільки $N = 18$, то за перший квартиль Q_1 беремо член варіаційного ряду з порядковим номером $x_5 = 65$. За третій квартиль Q_3 беремо член варіаційного ряду з порядковим номером $x_{14} = 85$. Інтерквартильний розмах $IKR = Q_3 - Q_1 = 85 - 65 = 20$.

Другий квартиль Q_2 (медіана) дорівнює 74. Отже, п'ятиточковою характеристикою цієї вибірки є п'ять значень: 63; 65; 74; 85; 94. Побудуємо «box-and-whiskers plot» за даними задачі в середовищі GeoGebra (рис. 1).

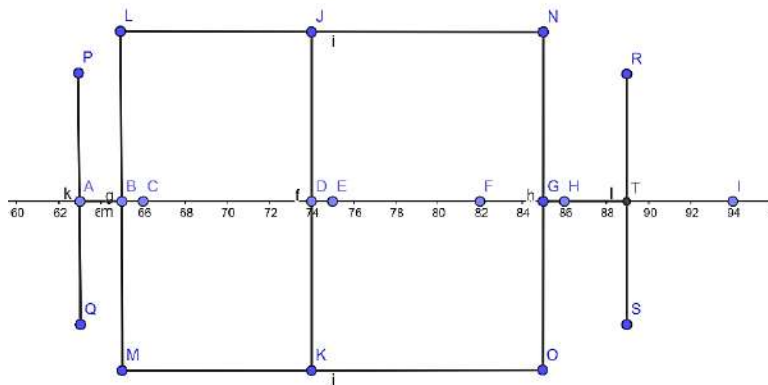


Рис. 1. «Box-and-whiskers plot» до задачі 1 (<https://www.geogebra.org/materials/yours>)

Статистичний розподіл вибірки є асиметричним, з чітко вираженим медіальним інтервалом. Маємо унімодальний випадок (можна визначити за рядом розподілу значення для моди $M_o = 63$, тобто є бал, який зустрічається найчастіше в порівнянні з іншими). Результат іспиту в 94 бали є викидом статистичного ряду.

Використання вбудованого модуля Аналіз даних, зокрема, інструменту «Описателна статистика», дозволить зекономити час та віднайти усі характеристики автоматично (рис. 2).

Усі явища та процеси, які нас оточують, взаємопов'язані, тому важливо навчитись вивчати та пояснювати взаємозв'язки та залежності між ними. Існує цілий ряд процедур для визначення статистичної залежності, коли зміна однієї досліджуваної величини впливає на розподіл іншої. Тому завдання математичної статистики полягає в кількісному оцінюванні взаємозалежностей щодо наявності, на пряму, форм впливу різних факторів на інші. Отже, в питання змістового модуля дисципліни включені методи кореляційного та регресійного аналізу.

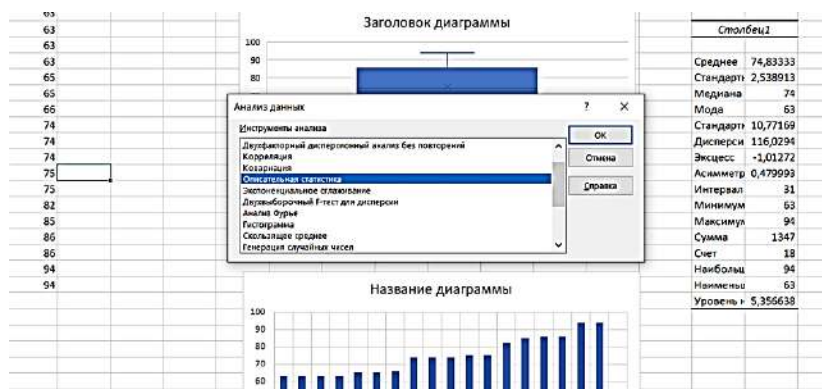


Рис. 2. Робота з вбудованим модулем «Аналіз даних» (Microsoft Excel 2016)

Крім стандартних завдань за навчальними посібниками для математичної статистики, формулюємо власні задачі. Дані для таких задач беремо із різних джерел: сайт Міністерства фінансів України (<https://minfin.com.ua/ua/>); Лондонська торгівельна біржа (<https://www.lme.com/>), результати опитувань соціологічної лабораторії факультету соціології Київського національного університету імені Тараса Шевченка (<https://sociology.knu.ua/uk/nauka/>); власні дослідження студентів фізико-математичного факультету.

Задача 2. Встановити зв'язок та перевірити залежність між коливаннями валют в Україні (євро, долар) за період 01.05.-26.05.2020 року.

Дані брали з сайту Міністерства фінансів України (для євро – <https://minfin.com.ua/ua/currency/nbu/eur/2020-05-20>, для долару – <https://minfin.com.ua/ua/currency/nbu/usd/2020-05-20>). Результати маємо в таблиці 1.

Таблиця 1

Курси НБУ валют (євро, долар) за період 01.05.-26.05.2020

Дати	Курс НБУ євро до гривні	Курс НБУ долара до гривні	Дати	Курс НБУ євро до гривні	Курс НБУ долара до гривні
01.05.20	29.24240000	26.97140000	14.05.20	29.02910000	26.68970000
02.05.20	29.24240000	26.97140000	15.05.20	28.78690000	26.67920000
03.05.20	29.24240000	26.97140000	16.05.20	28.78690000	26.67920000
04.05.20	29.29390000	26.95920000	17.05.20	28.78690000	26.67920000
05.05.20	29.50780000	26.97490000	18.05.20	28.76770000	26.63800000
06.05.20	29.26910000	26.97740000	19.05.20	28.79100000	26.57100000
07.05.20	29.10170000	26.93230000	20.05.20	29.02170000	26.51590000
08.05.20	28.91170000	26.81850000	21.05.20	29.14030000	26.58540000
09.05.20	28.91170000	26.81850000	22.05.20	29.46160000	26.76990000
10.05.20	28.91170000	26.81850000	23.05.20	29.46160000	26.76990000
11.05.20	28.91170000	26.81850000	24.05.20	29.46160000	26.76990000
12.05.20	29.09660000	26.81960000	25.05.20	29.16760000	26.75560000
13.05.20	29.08430000	26.79470000	26.05.20	29.31570000	26.87420000

Представимо розв'язання задачі в середовищі GeoGebra: після внесення даних, програма будує діаграму розсіяння, зображує пряму регресії та записує її рівняння (<https://www.geogebra.org/classic#spreadsheet>).

Кутковий коефіцієнт прямої $r = 0,31$ – це коефіцієнт кореляції, який характеризує зв'язок між випадковими величинами. За числовим значенням бачимо, що курси валют мають однакову фазу коливань, але залежність долара від євро невисока (рис. 3). Навіть можемо спрогнозувати, що при курсі євро до гривні в 30,0, курс долара становитиме 27,0716 (підтвердження маємо для курсів в валюту 20.03.2020 та 03.-04.04.2020).

До речі, при зміні змінних (незалежна – долар, залежна – євро) коефіцієнт кореляції дорівнює 0,9292, що говорить про сильний зв'язок між курсами долара та євро. Це підтверджено фінансовими аналітиками, адже курс євро в Україні формується від залежності євро/долар, а ось сам долар – вже на ринкових умовах (Мазуренко, 2019).

ОБГОВОРЕННЯ

Формування у випускника бакалаврату за спеціальністю 122 Комп'ютерні науки фахових компетенцій та програмних результатів передбачає включення у матеріал дисципліни питань по застосуванню методів статистичної обробки даних. Описова статистика та регресійний аналіз є найпростішими прикладами для аналізу вихідних даних. Правильне візуальне подання інформації дозволяє точно описати вихідні дані. Адже для фахівці в різних галузях вкрай важливим є вміння представляти результати своєї діяльності так, щоб це найкраще сприймалось іншими людьми – не фахівцями в цій галузі. Коробкова діаграма – це швидкий, компактний спосіб вивчення одного або декількох наборів даних у графічному вигляді, проста для побудови та в зручній формі подає важливі числові характеристики розподілу. Кореляційний аналіз даних обґрунтовує доцільність використання статистичних методів для обробки результатів досліджень будь-якої природи для встановлення залежності між факторами, а обчислення в спеціальних програмах підвищить рівень якості висновків.

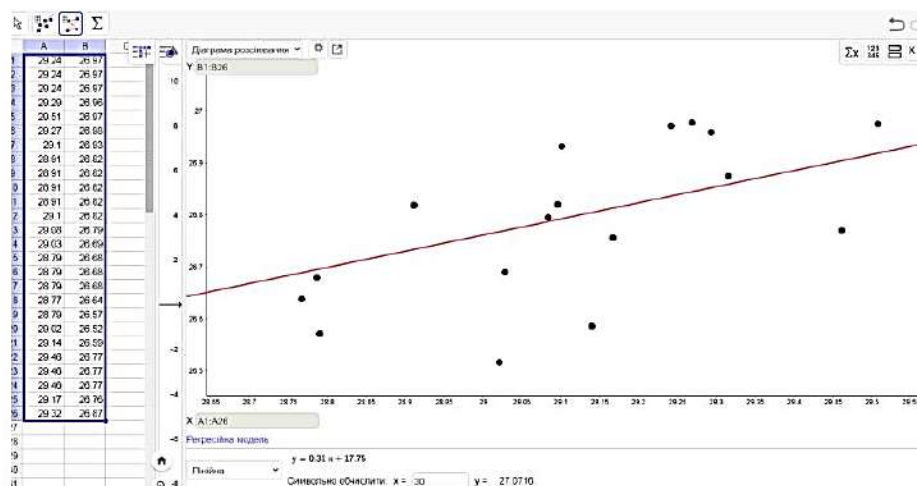


Рис. 3. Пряма регресії для дослідження курсів НБУ валют за 01.05.-26.05.2020

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШОГО ДОСЛІДЖЕННЯ

Власний досвід навчання студентів – майбутніх фахівців з інформаційних технологій – дав можливість сформулювати методичні рекомендації для змістового модуля «Математична статистика». Елементи аналітичної геометрії та дискретної математики, математичного аналізу лінійної алгебри, які вивчають студенти бакалавратуна першому курсі, дозволять використати матеріал дисциплін з метою глибокого розуміння фактів математичної статистики. Огляд сучасних статистичних пакетів дозволив переглянути методи закріплення інформації та урізноманітнити практичне вивчення основних питань дисципліни. У подальших публікаціях плануємо сформулювати методичні поради для застосування конкретних статистичних пакетів для пришвидшення одержання результатів експерименту та навести приклади авторських задач.

Список використаних джерел

1. Басюк Т. М., Думаєнський Н. О., Пасічник О. В. *Основи інформаційних технологій* : навч. посіб. / за ред. В. В. Пасічника. Львів : Новий Світ-2000, 2011. 390 с.
2. Гмурман В. Е. *Теория вероятностей и математическая статистика* : учеб. пос. 9-е изд. Москва : Высш. шк., 2003. 479 с.
3. Кислова М. А., Горшкова Г. А. Система комп'ютерної математики як складова мобільного навчального середовища з вищої математики. *Новітні комп'ютерні технології*. 2017. Том XV. С. 197-201.
4. Конет І. М. *Практикум з математичної статистики* : навч. посіб. Кам'янець-Подільський: Абетка-Світ, 2010. 212 с.
5. Мазуренко В. Долар проб'є рекордну позначку: аналітики озвучили прогнози. URL: <https://www.volynnews.com/news/all/dollar-probye-rekordnu-roznachku-analityky-ozvuchily-prohnozy> (Дата звернення 26.05.2020).
6. Майборода Р. Є., Сугакова О. В. Статистичний аналіз даних за допомогою пакету STATISTICA. URL: <http://matphys.rpd.univ.kiev.ua/downloads/courses/mmatstat/StatAn.doc> (Дата звернення 20.05.2020).
7. Мунтян О. А., Мунтян М. Л., Яровенко А. Г. Пакети прикладних програм статистичного аналізу результатів у наукових дослідженнях. *Сучасні інформаційні технології та інноваційні методи навчання у підготовці фахівців : методологія, теорія, досвід, проблеми*. Випуск 49. Київ-Вінниця : ТОВ фірма «Планер», 2017. С. 135-138.
8. Навчальна програма дисципліни самостійного вибору університету «Теорія ймовірностей, ймовірнісні процеси та математична статистика» для підготовки фахівців першого (бакалаврського) рівня вищої освіти галузі знань 12 Інформаційні технології спеціальності 122 Комп'ютерні науки за освітньо-професійною програмою «Сучасні інформаційні технології та програмування». Навчальні програми дисциплін кафедр алгебри та геометрії ЖДУ імені Івана Франка. 2019. URL: <https://zu.edu.ua/packages.asp?prefixs=155>. (Дата звернення 22.02.2020).
9. Роїк В. О., Денисюк О. І., Присяжнюк М. В. *Ефективна економіка*. 2017. № 7. URL: <http://www.economy.nayka.com.ua>. (Дата звернення 24.04.2020).
10. Савельєв В. *Статистика и котики*. Москва: АСТ, 2018. 189 с. ISBN 978-5-17-108287-1.
11. Середа Ю. В. Сучасні методи статистичного аналізу даних у гуманітарних науках : майстер-клас. URL: <http://i-soc.com.ua/ua/highschool/master-klas/metodi-stat-analizu-gum-nauk> (Дата звернення 19.03.2020).
12. Спірін О. М. Методологічні засади розвитку сучасних систем вищої освіти. *Вісник Житомирського державного університету імені Івана Франка*. 2005. № 20. С. 104-109.

References

1. Basiuk, T. M., Dumanskyi, N. O. & Pasichnyk, O. V. (2011). *Osnovy informatsiinykh tekhnolohii [Fundamentals of information technology]*. Lviv : Novyi Svit-2000 [in Ukrainian].
2. Gmurman, V. E. (2003). *Teorija veroyatnostej i matematicheskaja statistika [Probability theory and mathematical statistics]*. Moskva : Vyssh. shk [in Russian].
3. Kislova, M. A. & Horschikova, H. A. (2017). Systema kompiuternoi matematyky yak skladova mobilnoho navchalnoho seredovshcha z vyshchoi matematyky [The system of computer mathematics as a component of the mobile learning environment].

- environment in higher mathematics]. *Novitni kompiuterni tekhnologii – The latest computer technology*, 15, 197-201 [in Ukrainian].
4. Konet, I. M. (2010). *Praktykum z matematychnoi statystyky [Workshop on mathematical statistics]*. Kamianets-Podilskyi : Abetka-Svit [in Ukrainian].
 5. Mazurenko, V. (2019). Dollar probie rekordnu poznachku: a nalitky ozvuchylyprohnozy [The dollar will break the record mark: analysts voiced forecasts]. Retrieved from <https://www.volynnews.com/news/all/dolar-probye-rekordnu-poznachku-analitky-ozvuchyly-prohnozy> [in Ukrainian].
 6. Maiboroda, R. Ye. & Suhakova, O. V. (2015). Ctatystychnyi analiz danykh za dopomohoiu paketu STATISTICA [Statistical analysis of data using the STATISTICA package]. Retrieved from <http://matphys.rpd.univ.kiev.ua/downloads/courses/mmatstat/StatAn.doc> [in Ukrainian].
 7. Muntian, O. A., Muntian, M. L. & Yarovenko, A. H. (2017). Pakety prykladnykh prohram statystychnoho analiz rezultativ naukovykh doslidzhen [Packages of applied programs of statistical analysis of research results]. *Suchasni informatsiini tekhnologii ta innovatsiini metodyky navchannia u pidhotovtsi fakhivtsiv : metodolohiia, teoriia, dosvid, problem – Modern information technologies and innovative teaching methods in training: methodology, theory, experience, problems*, 49, 135-138. [in Ukrainian].
 8. Navchalna prohrama dystsypliny samostiinoho vyboru universytetu «Teoriia ymovirnostei, ymovirnisni protsesy ta matematychna statystyka» dlia pidhotovky fakhivtsiv pershoho (bakalavrskoho) rivnia vyschoi osvity haluznan 12 Informatsiini tekhnologii spetsialnosti 122 Kompiuterni nauky za osvitho-profesiinoiu prohramoiu «Suchasni informatsiini tekhnologii ta prohramuvannia». [Curriculum of the discipline of independent choice of the university "Probability theory, probability processes and mathematical statistics" for training specialists of the first (bachelor's) level of higher education in the field of knowledge 12 Information technology specialty 122 Computer science educational program "Modern information technology and programming"] Navchalni prohramy dystsyplin kafedry alhebrya heometrii ZhDU imeni Ivana Franka. (n.d.). [zu.edu.ua](https://zu.edu.ua/packages.asp?prefiks=155). Retrieved from <https://zu.edu.ua/packages.asp?prefiks=155> [in Ukrainian].
 9. Roik, V. O., Denisiuk, O. I. & Prysiazhniuk, M. V. (2017). Efektyvna ekonomika [Efficient economy], 7. Retrieved from <http://www.economy.nayka.com.ua> [in Ukrainian].
 10. Savel'ev, V. (2018). *Statistika i kotiki [Statistics and seals]*. Moskva : AST [in Russian].
 11. Sereda, Yu. V. (2013). Suchasni metodystatystychnoho analizu danykh u humanitarnykh naukakh : maister-klas [Modern methods of statistical data analysis in the humanities]. Retrieved from <http://i-soc.com.ua/ua/highschool/master-klas/metodi-stat-analizu-gum-nauk> [in Ukrainian].
 12. Spirin, O. M. (2005). Metodolohichni zasady rozvytku suchasnykh system vyschoi osvity [Methodological principles of development of modern systems of higher education]. *Visnyk Zhytomyrskoho derzhavnogo universytetu imeni Ivana Franka – Bulletin of Zhytomyr State University named after Ivan Franko*, 20, 104–109 [in Ukrainian].

STATISTICAL AND PROBABILITY COMPONENT OF THE CONTENT OF TRAINING OF INFORMATION TECHNOLOGY SPECIALISTS

O.A. Chemeris, A.V. Prus

The Zhytomyr State University after I. Franko, Ukraine

Abstract.

Formulation of the problem. The study of elements of statistics in the training of future specialists in computer technology contributes to the implementation of the applied orientation of education. The article provides examples of the use of mathematical and statistical methods of analysis for specific studies and provides guidelines, in particular, examples of calculation and presentation of numerical characteristics of the sample and finding the relationships between the elements of variation series.

Materials and methods. Analysis of scientific and methodological literature on mathematical statistics, generalization, and theoretical modeling; determination of the main statistical characteristics of the sample (using office programs and GeoGebra environment); regression analysis of research results to establish the relationship between the data; computer technology to determine numerical characteristics.

Results. The article reveals the issues of modern statistical research for various fields, the results of which were used in teaching the content module "Mathematical Statistics" for future specialists in information technology. Data analysis helps to identify useful information and draw the right conclusions. Using computer packages will make it easier and faster to get the desired result. A broad overview of methods for calculating problems in statistics allows you to gain experience in the practice of computing in this discipline. Examples of problem-solving with subsequent visualization of both initial and final data are given. MS Excel and GeoGebra software environments are used.

Conclusions. The students' own experience allowed them to formulate some recommendations for teaching the content module "Mathematical Statistics" and to conduct elements of data analysis, which are encountered by specialists in information technology in practice. The review of modern statistical packages will allow us to reconsider methods of fixing of the information and to diversify practical studying of the basic questions. In future publications, we plan to formulate methodological advice for the application of multifactor analysis on the example of the author's problems.

Keywords: data analysis, descriptive statistics, mustache box, correlation coefficient, regression line.