

Козловський Богдан,
здобувач другого (магістерського) рівня вищої освіти
фізико-математичного факультету,
Житомирський державний університет імені Івана Франка,
м. Житомир, Україна
Наконечна Оксана,
кандидат технічних наук,
доцент кафедри інформаційних технологій,
Одеський державний аграрний університет,
м. Одеса, Україна

МЕТОДОЛОГІЧНІ ОСНОВИ ТА ПРИКЛАДНЕ ЗАСТОСУВАННЯ КРИТЕРІЮ χ^2 ПІРСОНА

Математична статистика є важливою галуззю сучасної науки, яка надає методи для збору, обробки, аналізу та інтерпретації даних. Це дозволяє дослідникам знаходити закономірності, оцінювати достовірність отриманих результатів та приймати на їх основі обґрунтовані рішення. Такі підходи широко застосовуються у різних галузях, від медицини до соціології, що підкреслює універсальність статистичних методів.

З розвитком цифрових технологій та збільшенням обсягу даних, статистичні методи стали невід'ємною частиною емпіричних досліджень у багатьох сферах. Сьогодні дослідники та практики зіштовхуються з потребою у статистично достовірних інструментах, які допомагають підтвердити гіпотези, оцінити закономірності та зробити прогнози на основі зібраних даних. Одним з ключових методів математичної статистики є критерій χ^2 Пірсона, який використовується для перевірки відповідності емпіричних частот до теоретично очікуваних, особливо в аналізі категоріальних даних. Це робить його надзвичайно корисним для аналізу розподілів та асоціацій між змінними, що є частими завданнями у соціальних науках, біомедицині та ринкових дослідженнях.

Сучасні технології та розвиток нових методів обробки даних, зокрема в рамках штучного інтелекту та машинного навчання, значно розширили межі застосування критерію χ^2 Пірсона. Завдяки цьому, він набув нової цінності для наукових досліджень і аналізу, стаючи інструментом для оцінки відповідності даних у великих інформаційних масивах та побудови точних прогнозів.

Дослідження критерію χ^2 Пірсона та його сфер застосування є актуальним, оскільки цей метод дозволяє не лише глибше зрозуміти принципи його роботи, але й оцінити його ефективність у різноманітних галузях, де аналіз категоріальних даних відіграє ключову роль. Використання χ^2 Пірсона сприяє підвищенню надійності й точності статистичних висновків, що особливо важливо в умовах роботи з великими обсягами даних. Це, своєю чергою, розширює можливості для практичного застосування цього критерію в аналізі даних, прогнозуванні тенденцій та розробці рішень на основі об'єктивної статистичної інформації. Тому тема дипломної роботи є актуальною.

Мета дослідження полягає у дослідженні методологічних основ

Секція 4. Технології розробки інформаційних систем

застосування критерію χ^2 Пірсона, а також дослідити його значущість у різних галузях для оцінки достовірності даних та підтвердження статистичних гіпотез.

Виклад матеріалу. Наукове дослідження є ключовим елементом науково-технічного прогресу, оскільки воно забезпечує суспільство новими знаннями та інноваційними рішеннями. Це систематичний процес, що включає виявлення і аналіз закономірностей, що дозволяє глибше зрозуміти об'єкти та процеси, а також формує наукову базу для практичного застосування. Основні етапи дослідження, такі як вибір теми, формулювання мети, збір і обробка даних, є критично важливими для досягнення об'єктивних результатів та забезпечення наукової обґрунтованості.

Вибір теми дослідження є першим і визначальним кроком, що задає напрямок усій роботі. Дослідник має враховувати актуальність теми, її значення для науки та власний інтерес. Чітке визначення об'єкта і предмета дослідження допомагає зосередитися на найважливіших аспектах проблеми, що в свою чергу сприяє більш глибокому аналізу та отриманню результатів.

Мета дослідження, що визначає основні досягнення, формулюється як ціль, до якої прагне дослідник. Завдання, які деталізують цю мету, структурують дослідницький процес і забезпечують послідовність дій, що дозволяє рухатися до кінцевого результату. Крім того, наукове дослідження виконує кілька функцій: описову, пояснювальну, прогностичну та перетворювальну, що підкреслює його різноманітність і можливості впливу на розвиток суспільства [1].

Вибір методів для обробки даних є важливим етапом, оскільки різні дослідницькі завдання вимагають відповідних критеріїв і методів [2]. Основними цілями статистичного аналізу є підтвердження гіпотез, виявлення взаємозв'язків між змінними та прогнозування. Дослідник може вибрати між двома основними схемами організації дослідження: аналізом однієї групи даних або порівнянням двох наборів даних.

При виборі статистичного критерію враховується кількість вибірок і їх характер (залежні або незалежні). Зокрема, для виявлення відмінностей між значеннями параметрів застосовують критерії Стьюдента, Манна-Уїтні, Пірсона та інші, а для дослідження кореляційних зв'язків – коефіцієнти Пірсона, Спірмена тощо [3]. Правильний підхід передбачає ретельне планування дослідження і вибір методів обробки на початкових етапах, однак можливий також вибір критерію після отримання експериментальних даних на основі завдань і доступних методів.

Критерій χ^2 Пірсона є одним із найпоширеніших статистичних методів для аналізу відмінностей у розподілах ознак, особливо при роботі з категоріальними даними. Його застосовують, коли потрібно порівняти частоти в різних наборах даних або дослідити зв'язок між категоріальними змінними. Критерій χ^2 дозволяє оцінити, чи спостережувані відмінності між розподілами є статистично значущими [4].

Обмеження застосовності критерію Пірсона [4]:

1) обсяг вибірки $n \geq 30$; із зростанням n точність застосування критерію

Секція 4. Технології розробки інформаційних систем

підвищується; при вибірках з $n < 30$ метод дає дуже наближені значення, що може бути частково компенсовано, якщо критичне значення вибиратиметься для рівня значущості $p \leq 0.01$;

2) якщо кількість градацій ознаки g , то обсяг вибірки має задовольняти умові $n \geq 5g$;

3) вибрані градації ознаки (інтервали значень ознаки, які вважаються різними) повинні охоплювати всі можливі значення параметра в обох вибірках (іншими словами, сума частот в обох вибірках повинна обов'язково дорівнювати 1 або, що те саме, сума спостережень по градаціях повинна дорівнювати загальному числу спостережень); встановлені інтервали не повинні перекриватись;

4) якщо градацій лише дві («виконано» – «не виконано», «згоден» – «не згоден», «так» – «ні» тощо), необхідно вносити поправку (вона називається поправкою на безперервність). У таких випадках краще застосовувати критерій Фішера.

Загальні схеми розрахунку $\chi^2_{\text{експ}}$ дещо відрізняються для ситуацій, коли проводиться порівняння експериментального розподілу з теоретичним або двома експериментальними.

Зіставлення експериментального розподілу з теоретичним [4]:

1) визначити кількість вимірювань, що потрапили в кожную з g допустимих градацій ($n_{ie}, i = 1 \dots g$);

2) для того теоретичного розподілу, з яким проводиться порівняння експериментального, знайти теоретичні кількості вимірювань, що потрапляють у кожную градацію ознаки n_{it} :

$$n_{it} = n \cdot p_{it}, \quad (1)$$

де p_{it} – ймовірність попадання в інтервал i ; n – обсяг експериментальної вибірки; обчислення p_{it} є інтегральною задачею, вирішення якої потребує явного виду теоретичної функції розподілу. Для рівномірного розподілу:

$$n_{it} = \frac{n}{g}. \quad (2)$$

3) значення $\chi^2_{\text{експ}}$ знаходиться за формулою:

$$\chi^2_{\text{експ}} = \sum_{i=1}^g \frac{(n_{ie} - n_{it})^2}{n_{it}}. \quad (3)$$

Вираз (1.6) можна переписати у зручніший для обчислень:

$$\chi^2_{\text{експ}} = \sum_{i=1}^g \frac{(n_{ie})^2}{n_{it}} - n. \quad (4)$$

Зіставлення двох експериментальних розподілів [4]. При порівнянні двох експериментальних вибірок їх обсяги, у загальному випадку, можуть відрізнятися – позначимо їх n_1 та n_2 ; число градацій ознаки, залишається однаковим для обох вибірок і становить g ; тоді кількість вимірювань у кожній вибірці, що потрапляють у градацію під i номер позначимо n_{i1} та n_{i2} ($i = 1 \dots g$). Значення $\chi^2_{\text{експ}}$ знаходиться за формулою:

$$\chi^2_{\text{експ}} = n_1 n_2 \sum_{i=1}^g \left[\frac{1}{n_{i1} + n_{i2}} \left(\frac{n_{i1}}{n_1} - \frac{n_{i2}}{n_2} \right)^2 \right]. \quad (5)$$

У роботах [2-3] присвячених статистичній обробці експериментальних

Секція 4. Технології розробки інформаційних систем

результатів, $\chi_{\text{експ}}^2$ обчислюється через частоти потрапляння вимірювань у той чи інший інтервал (градацію) – такий підхід точніше передає зміст критерію, проте, для практичного використання формули (3) та (5) більш зручні, оскільки ґрунтуються безпосередньо на експериментальних даних та не вимагають проміжних розрахунків частот; при необхідності, не важко перейти від n_{i1} та n_{i2} до відповідних частот:

$$f_{ki} = \frac{n_{ki}}{n_k}; (k = 1, 2).$$

Оскільки в розрахункову формулу $\chi_{\text{експ}}^2$ входять квадрати різниць частот, цей критерій фіксує лише існування відмінностей у розподілах, але не напрямок зміни ознаки. Іншими словами, однакова відмінність розподілів може бути отримана як у випадку, коли в експериментальних даних частка показників зростає порівняно з контрольними, так і у випадку, коли вона знижується. Це означає, що зміни в експериментальних даних можуть відображати як позитивні, так і негативні тенденції. Таким чином, однакова відмінність розподілів може мати різні інтерпретації залежно від контексту дослідження та змін у даних.

Для виявлення напряму зсуву потрібно скористатися іншими статистичними способами, зокрема, критерії Макнамари (M), Вілкоксона (T) і знака (G -критерій) можуть бути застосовані для оцінки достовірності зсуву рівня ознаки у двох вибірках залежних даних. Ці непараметричні методи дозволяють перевірити гіпотезу щодо зсуву показників між двома пов'язаними вибірками (наприклад, до та після впливу певного фактора), що може вказувати на наявність зсуву та його напрям [2].

Також критерій тенденцій Пейджа (L) може використовуватись для аналізу зміни ознаки у трьох і більше вибірках, що є корисним при дослідженні спрямованого зсуву рівня ознак.

Для оцінки зсуву в кількох вибірках можна використовувати параметричні та непараметричні критерії дисперсійного аналізу (однофакторного або двофакторного), такі як критерій Фішера (F -критерій).

Критичне значення ($\chi_{\text{кр}}^2$) визначається за числом ступенів свободи (ν) даної задачі, яке на одиницю менше від кількості можливих градацій ознаки (g):

$$\nu = g - 1.$$

Далі для ν та p знаходиться критичні значення критерію Пірсона. За умови $\chi_{\text{експ}}^2 \geq \chi_{\text{кр}}^2$ застосовується експериментальна гіпотеза; при зворотному співвідношенні – вона відкидається. Зі зростанням ν зростає і $\chi_{\text{кр}}^2$.

На відміну від інших методів, критичне значення критерію Пірсона не залежить від обсягу експериментальних вибірок і визначається лише числом ступенів свободи. У той же час, при збільшенні обсягу вибірки найчастіше зростає $\chi_{\text{експ}}^2$ – за цієї причини при одних і тих же співвідношеннях частот велика вибірка зможе забезпечити прийняття експериментальної гіпотези, а менша – ні.

Критерій χ^2 Пірсона у наукових дослідженнях має широке застосування, наприклад, у медичних дослідженнях він допомагає визначити, чи пов'язані певні біомаркери з результатами діагностики, а в соціологічних – аналізувати залежність між рівнем доходів та освітою.

Секція 4. Технології розробки інформаційних систем

Критерій Колмогорова-Смірнова є альтернативним методом, який використовується для перевірки відповідності вибіркового розподілу теоретичному або для порівняння розподілів двох вибірок, але його частіше застосовують до неперервних даних. Він вимірює максимальну різницю між емпіричними функціями розподілу та може використовуватися для невеликих вибірок [1].

Критерій χ^2 Пірсона є одним із основних інструментів у статистичному аналізі, широко використовуваним у наукових дослідженнях для перевірки гіпотез, що стосуються асоціацій між категоріальними змінними. Основні сфери застосування цього критерію включають тести на асоціацію між змінними, перевірку адекватності статистичних моделей, а також порівняння розподілів частот у різних категоріях.

У наукових працях критерій Пірсона застосовується для виявлення статистично значущих зв'язків між змінними. Наприклад, він може використовуватися для аналізу тестових завдань у психодіагностиці або для оцінки якості веб-порталів [5], де перевіряється відповідність моделей реальним даним. У медичних дослідженнях критерій допомагає визначити відмінності у клінічному перебігу захворювань серед різних вікових груп, а в соціальних науках – для аналізу взаємозв'язків між освітою та політичними поглядами.

Критерій χ^2 також відіграє важливу роль у порівнянні частот у різних категоріях, що дозволяє дослідникам виявити значущі відмінності в розподілах характеристик, таких як рівень освіти або умови праці. З розвитком нових технологій, зокрема штучного інтелекту, критерій Пірсона знаходить нові сфери застосування, включаючи валідацію моделей та аналіз розподілів похибок [6-8].

Таким чином, критерій χ^2 Пірсона є потужним інструментом, що дозволяє дослідникам отримувати цінну інформацію про зв'язки між категоріальними змінними у різних галузях, від соціальних наук до медицини і екології.

Необхідно зазначити на важливості вимірювань і показників для застосування критерію χ^2 Пірсона в різних сферах досліджень, зокрема в освіті, науці та техніці. Вимірювання є основою статистичного аналізу, оскільки дозволяють зв'язувати показники з визначеними шкалами значень. Для забезпечення надійності результатів необхідно використовувати чітко визначені міри та шкали вимірювання. Також в даному методі важлива валідність показників, адже не всі вони можуть бути визнані такими без виконання основних умов вимірювання.

Ключовими елементами в статистичних дослідженнях є генеральні сукупності та вибірки. Правильний вибір типу вимірювальної шкали (номінативна, рангові, інтервальні, пропорційні) впливає на достовірність даних, особливо у соціальних науках.

Важливими елементами експериментального дослідження є статистичні гіпотези та критерії, зокрема формулювання нульової та альтернативної гіпотез.

Програмні засоби статистичних обчислень відіграють важливу роль у сучасному аналізі даних, забезпечуючи дослідникам потужні інструменти для проведення складних статистичних аналізів і моделювання в різних сферах,

Секція 4. Технології розробки інформаційних систем

таких як соціальні науки, економіка, медицина та інженерія. Пакети, такі як SPSS, SAS, Stata та MATLAB, пропонують широкий спектр можливостей для обробки даних, виконання статистичних тестів, моделювання та візуалізації результатів. Загалом, вибір програмного забезпечення залежить від специфіки дослідження, обсягу даних та вимог до аналізу, проте всі ці інструменти сприяють покращенню якості досліджень та прийняття обґрунтованих рішень на основі аналізу даних.

Практична реалізація критерію χ^2 Пірсона. Дослідна робота здійснювалася на базі одної із ЗОШ 1-3ст. Житомирської області за даними 2022-2024 н.р., як респонденти виступали учні 10 та 11 класів.

Мета дослідної роботи – визначити наскільки учні здатні до самоаналізу своїх математичних знань, самоспостереження та самооцінки своєї діяльності на уроках математики.

Методи дослідження: емпіричні - тестування, анкетування, аналіз продуктів навчальної діяльності учнів, педагогічний експеримент; вивчення та узагальнення досвіду викладання шкільного курсу «Математика» в 10-11 класах; статистичні - обробка емпіричних даних, кількісний та якісний аналіз результатів дослідження, статистична обробка результатів за допомогою критерію χ^2 Пірсона для перевірки гіпотез та визначення статистичної значущості отриманих результатів.

Загалом у дослідній роботі взяли участь 62 учня (дані 30 осіб контрольного 2022-2023н.р. року та 32 особи експериментального 2023-2024 року). Саме в цей період розпочалось активне впровадження концепції навчання нової Української школи у старших класах. Тому дані за 2022-2023 н.р. будуть базовими (контрольними) до впровадження концепції, за цими даними можна визначити початковий рівень учнів до застосування математичних знань. І відповідно, дані за 2023-2024 н.р. будемо вважати експериментальними, в цей період учні навчаються за згаданою концепцією. Таким чином, порівняння даних дозволить чітко визначити, чи є зміни в навчальних результатах, і чи це пов'язано із запровадженою концепцією.

Для оцінки рівня математичних знань учнів за рефлексивним критерієм використано опитувальник діагностики рефлексивності. Такий опитувальник складається із запитань, які спрямовані на виявлення наступних аспектів: здатність до аналізу власних дій і вчинків; оцінка власних досягнень та помилок; вміння коригувати поведінку на основі минулого досвіду; здатність прогнозувати наслідки своїх дій; уміння порівнювати свої дії з діями інших та робити висновки.

Інтерпретація відповідей відбувалась за таким принципом: відповіді типу «А» оцінювались в 3 бали, «В – в 2 бали, «С» - в 1 бал, підсумовувались бали кожного учасника та розподілялись на рівні рефлексивності: низький (4-6 балів), середній (7 – 9 балів) та високий (10 – 12 балів). Це дозволило встановити розподіл учнів за кожним питанням у рамках цих діапазонів.

На підставі цих показників було сформовано гіпотези для оцінки діагностики рефлексивності учнів: H_0 - немає статистично значущої різниці в

Секція 4. Технології розробки інформаційних систем

самоаналізу своїх математичних знань та самооцінки діяльності на уроках математики в залежності від інших змінних; Н1 - існує статистично значуща різниця в самоаналізу своїх математичних знань та самооцінки діяльності на уроках математики в залежності від інших змінних.

Результати діагностики рефлексивності у контрольній та експериментальній групах сформувались наступним чином (рис. 1 - 2).

Здійснивши описову статистику, було визначено, що середні оцінки на перше (nk1) та друге питання (nk2) для контрольної групи вказують на вищу частоту спроб аналізувати свої дії після виконання складних завдань (41,7%) у порівнянні з обговоренням досягнень із вчителем (33,9%).

Розраховані значення χ^2 для кожного з питань показали, що отримані показники перевищують критичні значення ($\chi^2_{кр}=5,991$), що дозволяє стверджувати наявність статистично значущої різниці між групами.

Таким чином, аналіз анкетування показує, що учні, які працюють за експериментальною методикою, виявляють вищий рівень здатності до самоаналізу та ефективніше використовують досвід минулих завдань для покращення своїх результатів.

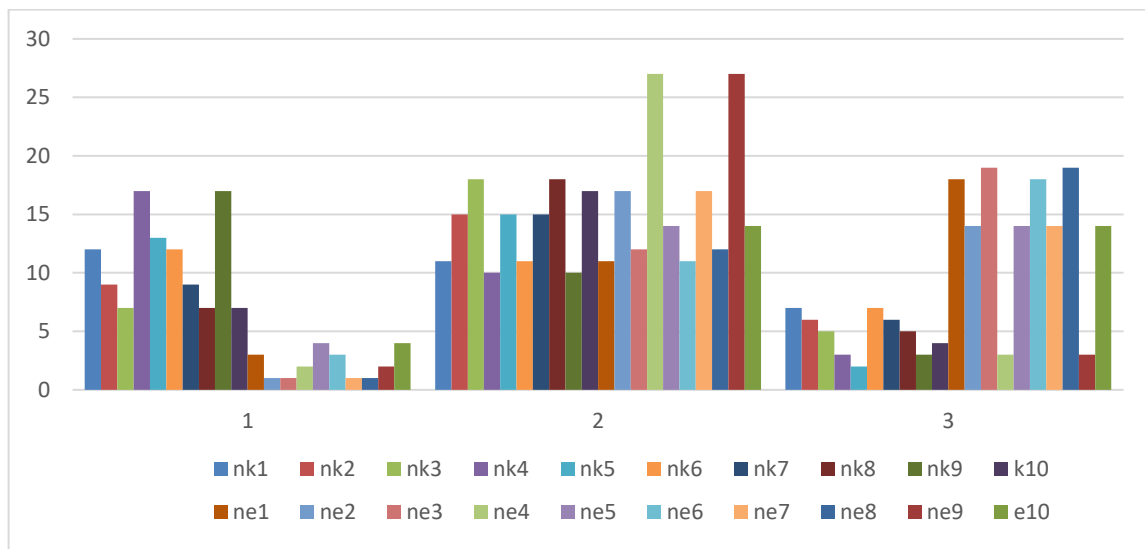


Рис. 1. Початкові дані діагностики рефлексивності

Дж. розроблено автором

Секція 4. Технології розробки інформаційних систем

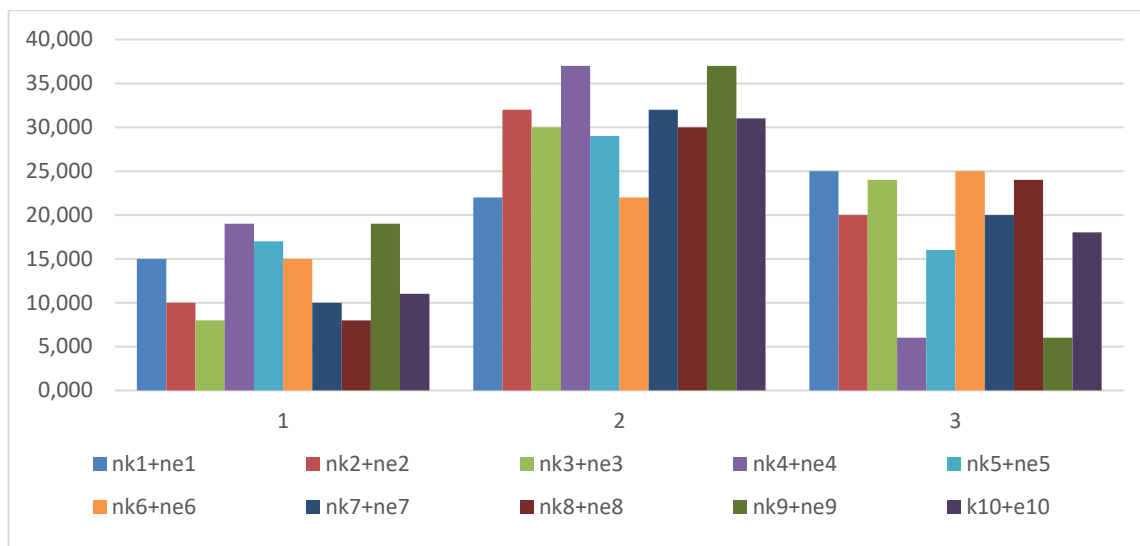


Рис. 2. Результати діагностики рефлексивності

Дж. розроблено автором

У процесі виконання наукового дослідження було використано програму Microsoft Excel. Дане програмне забезпечення забезпечило ефективну обробку даних, проведення статистичного аналізу, а також візуалізацію отриманих результатів. Excel дозволив створити таблиці та графіки, що сприяло кращому розумінню та інтерпретації отриманих результатів, а також підвищило наочність представлення інформації.

На основі отриманих результатів, можна сформулювати такі рекомендації для вдосконалення навчального процесу для підвищення рівня самоаналізу учнів у процесі вивчення математики:

1. Варто організувати систематичні рефлексивні вправи після виконання складних завдань, під час яких учні можуть аналізувати свої дії, що і буде формувати навички самоконтролю та самовдосконалення.

2. Рекомендується навчати учнів планувати свої дії перед виконанням складних завдань. Це можна зробити через надання шаблонів або алгоритмів для формування чіткої послідовності дій, що сприятиме організованому підходу до вирішення задач.

3. Надавати підтримку учням у процесі самостійного аналізу помилок, що допоможе їм розвивати критичне мислення. Для цього можна запропонувати картки або бланки для аналізу помилок, де учні зможуть записувати, у чому полягала помилка та як її виправити.

4. Підтримка використання учнями нових підходів для вирішення задач (наприклад, наводити альтернативні методи вирішення задач) дозволить їм знаходити оптимальні шляхи досягнення результату. Для цього варто розробити інтерактивні вправи або завдання з використанням різних стратегій, що стимулюватиме учнів до експериментування з методами.

5. Регулярний аналіз та вдосконалення на основі попередніх робіт може допомогти учням глибше розуміти свої сильні та слабкі сторони. Тому рекомендується проводити заняття з аналізу виконаних робіт з метою вдосконалення.

6. Організація групових занять, де учні зможуть обговорити свої досягнення

Секція 4. Технології розробки інформаційних систем

та недоліки з іншими, сприятиме розвитку комунікативних навичок та обміну досвідом, що допоможе краще засвоїти нові знання та навички.

7. Розробка завдань з реальними задачами допоможе учням бачити практичну користь від аналізу своїх дій. Це сприятиме розвитку критичного мислення та відповідальності за власні рішення.

Висновок. Для статистичної обробки результатів використовувався критерій χ^2 Пірсона. Його застосування дозволяє стверджувати, що в експериментальній та контрольних групах на момент закінчення дослідного експерименту існують статистично достовірні відмінності.

Таким чином, практична реалізація критерію χ^2 Пірсона продемонструвала, як цей статистичний метод може допомогти в прийнятті обґрунтованих рішень, а також може стати корисним для аналізу даних здобутків учнів старшої школи до застосування математичних знань, що також підсилює роль критерію χ^2 у педагогічних дослідженнях та сприяє розвитку математичної грамотності серед молоді.

Список використаних джерел та літератури

1. Навчальний посібник щодо вивчення дисципліни "Методи прикладного статистичного аналізу" / Укл.: Г.Г. Швачич, В.С. Коноваленков, О.В. Соболенко, Т. М. Заборова, В. І. Христян, Є. Є. Єгорцева – Дніпропетровськ: НМетАУ, 2017. 178 с.

2. Моцний Ф.В. Аналіз непараметричних і параметричних критеріїв перевірки статистичних гіпотез. Частина І. Критерії узгодження Пірсона і Колмогорова / Ф.В. Моцний // Статистика України. 2018. № 4. С. 14-24. URL: <http://surl.li/ukhzt>

3. Математичні методи в психології: курс лекцій. Мультимедійне навчальне видання. Х.: НУЦЗУ, 2020. 141 с. URL: <http://surl.li/zgqtxs>

4. Самойчук К.О., Верхоланцева В.О. Методи теоретичних і експериментальних досліджень. URL: <http://surl.li/yukljl>

5. Гикавчук, М.С. Інформаційна технологія аналізу конкурентоздатності веб-порталів [Текст] / М. С. Гикавчук, С. С. Петровський, Т. К. Скрипник // Вісник Хмельницького національного університету. Технічні науки. 2019. №6. С. 120-124. URL: <http://surl.li/uhqyu>

6. Полупанов В.М., Каретніков Р.В. Критерії Пірсона в застосуванні до аналізу n-факторного масиву вимірів при інженерних рішеннях в сільському господарстві. URL: <http://surl.li/ukibs>

7. Стожок А.Ю. Інструментарій дослідження соціальної мобільності персоналу. Проблеми сучасних трансформацій. Серія: економіка та управління / Економіка та управління підприємствами (за видами діяльності). 2024. № 13. URL: <http://surl.li/elnnhe>

8. Юр'єва Л.М., Шорніков А. В. Новий інструмент у діагностиці кіберзалежності / Український Вісник Психоневрології. 2021. URL: <http://surl.li/cvtevb>