

Ім'я користувача:  
Житомирський економіко-гуманітарний інститут

ID перевірки:  
1016376395

Дата перевірки:  
19.06.2024 21:08:46 EEST

Тип перевірки:  
Doc vs Internet

Дата звіту:  
19.06.2024 21:09:33 EEST

ID користувача:  
100011285

Назва документа: Петрівська курсова

Кількість сторінок: 19 Кількість слів: 3930 Кількість символів: 29839 Розмір файлу: 47.82 KB ID файлу: 1016184492

## 2.26% Схожість

Найбільша схожість: 0.51% з Інтернет-джерелом (<https://mindthegraph.com/blog/uk/%d0%be%d0%b4%d0%bd%d0%be...>)

2.26% Джерела з Інтернету

63

Сторінка 21

Пошук збігів з Бібліотекою не проводився

## 0% Цитат

Не знайдено жодних цитат

Не знайдено жодних посилань

## 0% Вилучень

Немає вилучених джерел

## ВСТУП

У світі науки та досліджень існують ситуації, коли зіштовхуємося з відхиленням від нормального розподілу ознаки. Це може відбуватися з різними параметрами досліджуваного явища: від фінансових даних до результатів медичних досліджень. В таких випадках має велике значення вміння аналізувати дані і виявляти відмінності в розподілі ознаки.

Актуальність теми "Множинні порівняння у випадку відмінності закону розподілу ознаки від нормального. Критерій Даннета" очевидна у багатьох сферах науки, досліджень та практики. Ось декілька аргументів, які підтверджують цю актуальність:

**Необхідність точних порівнянь:** В умовах, коли дані не мають нормального розподілу, важливо мати ефективні методи порівняння множин. Це особливо актуально для наукових досліджень та практичних застосувань, де точні порівняння можуть визначати подальші кроки аналізу та прийняття рішень.

**Надійність статистичних методів:** Застосування надійних статистичних методів є важливим для забезпечення обґрунтованих висновків на основі аналізу даних. Критерій Даннета виявляється ефективним інструментом у випадках, коли розподіл ознаки відрізняється від нормального.

**Розширення методологічної бази:** Вивчення та використання критерію Даннета у випадку відхилення від нормального розподілу ознаки сприяє розвитку методології порівняння множин даних. Це дозволяє розширити арсенал інструментів для аналізу даних і забезпечити більш глибоке розуміння статистичних взаємозв'язків.

**Практичне застосування:** Навички та знання, отримані в результаті дослідження критерію Даннета, можуть мати практичне значення в різних областях, таких як медицина, соціологія та інші, де аналіз даних є ключовим елементом прийняття рішень.

**Мета курсової роботи** полягає у вивченні та аналізі методу порівняння множин даних у випадку, коли розподіл ознаки відрізняється від нормального. Основною метою є дослідження та розуміння принципів

функціонування критерію Даннета та його застосування для об'єктивного порівняння множин даних у випадку відхилення від нормального розподілу ознаки.

Основні завдання курсової роботи:

- Провести аналіз основних понять та принципів статистичного аналізу даних у контексті порівняння множин.
- Детально розглянути теоретичні засади розподілу ознаки та методи порівняння множин даних.
- Описати метод критерію Даннета, вивчити його теоретичні засади та математичні моделі.
- Провести практичний аналіз застосування критерію Даннета до реальних наборів даних з відхиленням від нормального розподілу.
- Сформулювати висновки щодо ефективності та обмежень критерію Даннета у випадку відхилення від нормального розподілу ознаки.

Обґрунтування вибору критерію Даннета для аналізу відмінностей у випадку відхилення від нормального розподілу ознаки базується на кількох факторах: Ненормальний розподіл даних: Критерій Даннета є непараметричним методом, тобто він не вимагає певних припущень щодо розподілу даних. Це особливо важливо у випадках, коли дані не відповідають нормальному розподілу, а отже, традиційні методи, засновані на нормальних припущеннях, можуть бути неадекватними.

Ефективність при малих вибірках: Критерій Даннета добре працює навіть у випадках, коли ми маємо малі вибірки даних. Це важливо, оскільки багато статистичних методів можуть бути менш ефективними або неадекватними у таких умовах.

Відсутність вимог до нормальності: Оскільки критерій Даннета не базується на припущенні про нормальний розподіл даних, він є відмінним вибором для ситуацій, коли розподіл ознаки відхиляється від нормального.

Отже, обґрунтування вибору критерію Даннета для аналізу відмінностей у випадку відхилення від нормального розподілу ознаки полягає у його

здатності до ефективного та об'єктивного порівняння множин даних у широкому спектрі умов, включаючи ті, де інші методи можуть бути менш ефективними або неадекватними.

Курсова робота містить в собі: вступ, три розділи, висновок, список використаної літератури, додатки.

## РОЗДІЛ I. Теоретичні аспекти

### 1.1 Огляд та аналіз основних понять статистики, пов'язаних з порівнянням множин даних

Порівняння множин даних є ключовою задачею в статистичному аналізі.

Статистика, безумовно, є ключовою галуззю в сучасному світі, оскільки вона готова збирати, аналізувати та інтерпретувати дані. Порівняння множини даних є результатом статистики, завдяки чому воно дозволяє отримати об'єктивні висновки та зробити вірні висновки. Для успішного порівняння множини даних важливо розуміти основні поняття та методи, які допомагають ефективно аналізувати цю складну область. Важливі статистичні показники, такі як середнє значення, медіана, дисперсія та стандартне відхилення, допомагають у точному різниці різних наборів даних. Застосування методів, таких як t-тест Стюдента, варіантів аналізу

(ANOVA), кореляція та інші, дозволяє провести більш глибокий аналіз та зробити обґрунтовані висновки на основі статистичних даних.

Середнє значення - це показник, який використовується для оцінки центральної тенденції масиву даних. Зазвичай обчислюється як сума всіх значень, поділена на їх кількість. Середнє значення допомагає зрозуміти типове або середнє представлення даних у вибірці. Воно може бути корисним для узагальнення або резюмування даних, а також для порівняння різних груп або періодів часу.

Наприклад, якщо ми аналізуємо середню оцінку студентів у класі, середнє значення оцінок нам допоможе зрозуміти, як студенти в цілому впоралися з матеріалом. Це дозволяє нам зробити загальні висновки про успішність групи студентів.

У контексті статистики середнє значення є важливим інструментом для аналізу та інтерпретації даних, але важливо також враховувати й інші міри центральної тенденції, такі як медіана та мода, для отримання більш повного уявлення про розподіл даних.

Медіана - це центральне значення у впорядкованому ряду даних, яке ділить ряд на дві рівні частини: половину значень розташовано до медіани, а іншу половину - після медіани. Вона нечутлива до великих викидів або аномальних значень у порівнянні з середнім значенням. Медіана допомагає зрозуміти, як дані розподілені посередині.

Для визначення медіани необхідно впорядкувати всі значення у порядку зростання або спадання і визначити середнє значення середнього числа.

Якщо кількість значень парна, то медіана рахується як середнє арифметичне двох середніх значень в середині ряду.

Медіана є важливою мірою центральної тенденції, оскільки вона нечутлива до екстремальних значень (викидів) у порівнянні з середнім значенням. Це означає, що навіть якщо деякі значення виходять за рамки типового розподілу, медіана залишиться стабільною. Тому медіана часто використовується для оцінки центральної тенденції в даних з викидами або

асиметричним розподілом.

Наприклад, якщо ми маємо набір даних про дохід людей у певному регіоні, де декілька високодохідних осіб мають значно більший дохід, ніж більшість населення, медіана може краще відображати типовий дохід населення, ніж середнє значення.

Мода - це значення, яке зустрічається найчастіше у наборі даних. Мода дозволяє ідентифікувати найбільш типові значення у вибірці. Наприклад, якщо ми маємо набір даних про оцінки студентів за тестом, де оцінки можуть бути цілими числами від 1 до 10, мода покаже, яке значення отримало найбільшу кількість студентів. Якщо, наприклад, оцінка 7 зустрічається найчастіше серед усіх інших оцінок, то мода буде дорівнювати 7.

Моду також можна використовувати для аналізу категоріальних даних, таких як тип товару, кольори або категорії оцінок. У цьому випадку мода вказує на найпопулярнішу категорію серед усіх інших.

Отже, мода допомагає зрозуміти, які значення або категорії найчастіше зустрічаються у вашому наборі даних, що **може бути важливою інформацією для подальшого аналізу та прийняття рішень.**

Розмах даних - це різниця між найбільшим і найменшим значеннями в наборі даних. Розмах даних вказує на розподіл значень по всьому діапазону.

Наприклад, якщо у нас є набір даних про висоту рослин, де найвища рослина має висоту 200 см, а найнижча - 100 см, то розмах цього набору даних буде 100 см ( $200 - 100 = 100$ ). Це означає, що висота рослин варіюється на проміжку 100 см.

Розмах даних - це одна з простих, але важливих мір розкиду даних, яка допомагає отримати загальне уявлення про їхню варіабельність та діапазон значень.

Дисперсія та стандартне відхилення. Дисперсія вказує на ступінь розподілу даних навколо їх середнього значення, а стандартне відхилення визначає розмір цієї розсіяння.

Дисперсія обчислюється як середнє квадратичне відхилення кожного

значення від середнього значення всіх значень. Дисперсія дає уявлення про те, наскільки далеко в середньому кожне значення знаходиться від середнього, тобто наскільки варіюються дані.

Стандартне відхилення визначає розмір цієї розсіяння та є квадратним коренем з дисперсії. Воно вимірюється у тих же одиницях, що й дані, і є більш інтерпретованим, оскільки воно представляє реальний розмір розсіяння даних. Чим більше стандартне відхилення, тим більше розсіяння даних відносно їх середнього значення.

Обидві ці міри допомагають зрозуміти, наскільки узгоджені або розсіяні дані. Мале стандартне відхилення або дисперсія вказує на те, що дані, ймовірно, концентровані близько до середнього значення, тоді як велике відхилення або дисперсія вказує на те, що дані можуть бути розподілені далеко від середнього, що свідчить про більшу варіабельність в наборі даних.

Отже, визначивши основні поняття та методи порівняння даних, можна сказати, що вони допомагають більш точному результату аналізу складних наборів даних. Огляд основних понять статистики включає в себе розгляд понять центральної тенденції (середнє значення, медіана, мода), розмаху даних, дисперсії та стандартного відхилення.

Розуміння цих основних понять статистики допомагає аналізувати та порівнювати множини даних з різними характеристиками. Наступні кроки у аналізі можуть включати використання статистичних методів, таких як тестування гіпотез, щоб робити висновки на основі порівняння цих характеристик між різними групами даних.

## **1.2 Опис розподілу ознаки та його відмінностей від нормального розподілу**

Розподіл ознаки - це спосіб, яким розподілені значення або спостереження цієї ознаки у вибірці або популяції. Нормальний розподіл, також відомий як гауссівський розподіл, є одним з найбільш поширених і важливих розподілів у статистиці. Він має симетричну колоколоподібну форму і характеризується такими властивостями, як середнє значення та стандартне відхилення.

Відмінності від нормального розподілу можуть бути спричинені різними факторами. Ось кілька загальних відхилень, які можуть впливати на розподіл ознаки:

Скос (асиметрія): нормальний розподіл має нульовий скос, тобто він є симетричним. Відхилення від цієї симетрії можуть виявлятися у вигляді великого або малого вибоїв на одному з кінців розподілу.

Екссес (вибіг або сплющення): нормальний розподіл має нульовий екссес. Відхилення від цієї характеристики можуть виявлятися у вигляді вибігу (збільшення концентрації значень в хвостах розподілу) або сплющення (зменшення концентрації значень в хвостах розподілу).

Важкий хвіст (тяжке хвостове розподілення): нормальний розподіл має хвіст, який експоненціально зменшується при віддаленні від середнього значення. В інших розподілах може спостерігатися більш повільне зменшення хвоста, що означає, що високі або низькі значення більш ймовірні, ніж у нормальному розподілі.

Нормалізованість: нормальний розподіл характеризується тим, що його значення зосереджені навколо середнього значення, і його розкид описується стандартним відхиленням. В інших розподілах може бути велика концентрація значень у певних областях, що призводить до відмінностей у розподілі.

Отже, важливо враховувати ці відмінності від нормального розподілу при використанні статистичних методів та інтерпретації результатів досліджень.

### **1.3 Перегляд методів порівняння множин даних та їх обмеження в умовах відхилення від нормального розподілу**

Умови відхилення від нормального розподілу ознаки можуть виникати у різних випадках, і це може ускладнити порівняння множин даних за допомогою традиційних методів. Такі умови включають асиметрію, важкі хвости розподілу, викиди або аномальні значення.

Ось деякі методи порівняння множин даних і їх обмеження в умовах відхилення від нормального розподілу:



Статистичні тести для нормально розподілених вибірок: традиційні тести, такі як t-тест Стьюдента або аналіз дисперсії (ANOVA), передбачають нормальний розподіл даних. Умови відхилення від цього розподілу можуть призвести до неправильних результатів цих тестів.

t-тест Стьюдента - це статистичний метод, який використовується для порівняння середніх значень двох груп даних і визначення того, що є статистично значущою різницею між ними. Цей тест на міжнародному застосуванні можна вважати різницею між середніми значеннями групи значущої зі статистичної точки зору, або ж чи може ця різниця бути випадковою.

Існує два типи t-тестів:

Односторонній t-тест: використовується, коли маємо чітку гіпотезу про те, яка група буде мати вище (або нижче) середнє значення. Наприклад, ми можемо перевіряти гіпотезу, що середній рівень здоров'я пацієнтів після лікування вище, ніж перед лікуванням.

Двосторонній t-тест: використовується, коли маємо припущення про те, що групи можуть відрізнитися, але не маємо чіткої уяви про те, яка з них має вище або нижче середнє значення. Цей тест визначає, чи є статистично значуща різниця між середніми значеннями груп.

Аналіз дисперсії (ANOVA) - це статистичний метод, який використовується для порівняння середніх значень трьох або більше груп даних, щоб визначити, чи є статистично значущі відмінності між ними.

Основна ідея ANOVA полягає в порівнянні варіації між групами (міжгрупова варіація) з варіацією всередині груп (внутрішньогрупова варіація). Якщо міжгрупова варіація значно більше, ніж внутрішньогрупова, то ми можемо зробити висновок, що середні значення груп відрізняються статистично значущим чином.

ANOVA генерує F-статистику, яка є відношенням між міжгруповою і внутрішньогруповою варіацією. Якщо F-статистика виявиться статистично значущою, це означає, що є значущі відмінності між середніми значеннями

груп.

ANOVA може бути використаний для порівняння ефективності різних методів лікування, впливу різних дієт на здоров'я тощо.

ANOVA дозволяє ефективно аналізувати вплив багатьох факторів на результати експерименту, враховуючи внутрішні та міжгрупові варіації, що робить його корисним інструментом для дослідження різноманітних наукових і практичних проблем.

Непараметричні методи: непараметричні методи, такі як критерій Манна-Уїтні або критерій Вілкоксона, не вимагають передбачень щодо розподілу даних. Вони є більш гнучкими і можуть бути ефективними у випадках відхилень від нормального розподілу. Однак ці методи можуть мати меншу потужність (меншу здатність виявляти відмінності між групами) в порівнянні з параметричними тестами, коли дані дійсно мають нормальний розподіл. Критерій Манна-Уїтні (або U-тест Манна-Уїтні) - це непараметричний статистичний тест, який використовується для порівняння середніх значень двох незалежних вибірок. Він є альтернативою t-тесту для випадків, коли не виконуються передумови для застосування параметричних методів (наприклад, коли розподіл даних не є нормальним або коли вибірки мають різну дисперсію).

Основна ідея тесту полягає в порівнянні сум рангів двох вибірок. Для кожного значення вибірки визначається його ранг, і потім сумуються ранги для кожної вибірки. Потім порівнюється сума рангів між групами.

У випадку, якщо розподіл значень у двох групах ідентичний, очікується, що сума рангів в кожній групі буде приблизно однаковою. Таким чином, якщо вибірки мають однакові розподіли, сума рангів буде близько до однакової в обох групах. Але якщо середні значення відрізняються, одна з вибірок буде мати суму рангів, яка значно вища, ніж у іншій групі.

Критерій Манна-Уїтні може бути корисним в тих випадках, коли параметричні методи не можуть бути застосовані.

Критерій Вілкоксона, також відомий як тест Вілкоксона на взаємозв'язані

вибірки або тест Вілкоксона для залежних вибірок, є непараметричним статистичним тестом, який використовується для порівняння двох залежних вибірок. Він вимагає відносно невисоких вимог до розподілу даних та може бути використаний, коли параметричні методи, такі як t-тест, не можуть бути застосовані через відхилення від нормальності або інші обмеження.

Основна ідея тесту полягає в тому, щоб порівняти медіани двох залежних вибірок, використовуючи ранги відповідних спостережень. Процедура тесту передбачає наступні кроки:

1. Ранжування різниць між відповідними парами спостережень.
2. Обчислення сум рангів позитивних та негативних різниць.
3. Визначення меншої суми рангів.
4. Визначення критичного значення тесту або обчислення р-значення для порівняння зі значенням статистики тесту.

Критерій Вілкоксона може бути особливо корисним, коли працюємо з малими вибірками або коли дані мають велику варіабельність. Він використовується в різних областях, де потрібно порівнювати результати випробувань або експериментів.

Критерій Даннета: цей критерій, також відомий як критерій узгоджених рангів, є непараметричним методом порівняння множин даних. Цей статистичний тест використовується для визначення, чи існують статистично значущі відмінності між трьома або більше групами вибірок. Він використовує ранги даних і не вимагає нормального розподілу. Критерій Даннета є ефективним інструментом у випадках, коли дані відхиляються від нормального розподілу або мають важкі хвости розподілу.

Основна ідея критерію Даннета полягає в тому, що для кожного значення вибірки визначається його ранг у порівнянні з іншими значеннями у всьому наборі даних. Потім порівнюються суми рангів для кожної групи.

Критерій Даннета може бути застосований у випадках, коли дані не відповідають вимогам для використання параметричних методів, таких як t-тест або аналіз дисперсії (ANOVA). Наприклад, він може бути корисним для

порівняння ефективності трьох або більше методів лікування у клінічних дослідженнях, коли розподіл даних не є нормальним або коли є інші відхилення від передумов для використання параметричних методів. Обмеження параметричних методів у випадках відхилення від нормального розподілу підкреслюють важливість альтернативних непараметричних методів, таких як критерій Даннета. Врахування умов відхилення і вибір правильного методу аналізу допоможе забезпечити достовірні та об'єктивні результати порівняння множин даних.

## РОЗДІЛ II. Критерій Даннета: теорія

### 2.1 Побудова та обґрунтування критерію Даннета

Критерій Даннета, також відомий як критерій узгоджених рангів, є непараметричним методом порівняння множин даних. Він був розроблений Уїлфредом Даннетом і використовується для виявлення статистичних відмінностей між двома або більше групами даних, коли вони мають різний розподіл або коли немає можливості використовувати параметричні методи через відхилення від нормального розподілу.

Основна ідея критерію полягає в тому, що ми замінюємо реальні значення вибірки на їх ранги впорядкованого ряду, а потім порівнюємо суми цих рангів між групами. Якщо суми рангів суттєво відрізняються, це може свідчити про

наявність статистично значущих відмінностей між групами.

Обґрунтування використання критерію Даннета полягає в його непараметричній природі та робустності до відхилень від нормального розподілу. Він не потребує передбачень щодо розподілу даних та є менш чутливим до викидів або аномальних значень порівняно з параметричними методами. Це робить його важливим інструментом для порівняння множин даних у ситуаціях, коли традиційні методи можуть бути нечутливими або неадекватними.

Побудова критерію Даннета включає обчислення статистики критерію за допомогою сум рангів і подальше порівняння цієї статистики з табличними значеннями для визначення статистичної значущості різниці між групами даних. Використання цього критерію дозволяє отримувати об'єктивні та надійні результати порівняння навіть у складних умовах відхилення від нормального розподілу.

Основні кроки побудови та обґрунтування критерію Даннета:

Використання рангів: Замість використання конкретних значень даних, критерій Даннета використовує ранги. Ранг - це порядковий номер, який призначається значенню вибірки у впорядкованому ряду. Це дозволяє уникнути проблем, пов'язаних з нормальним розподілом даних, і зробити метод незалежним від конкретного розподілу.

Сума рангів: Для кожної групи обчислюється сума рангів. Це дозволяє отримати одну числову міру для кожної групи, яку можна порівняти між собою.

Статистика критерію Даннета: Статистика критерію обчислюється за формулою, яка залежить від сум рангів та розмірів груп. Ця статистика використовується для порівняння між групами та визначення статистичної значущості відмінностей.

Застосування критерію Даннета може бути обґрунтоване з кількох причин:

Непараметричний метод: критерій Даннета є непараметричним методом, тобто він не потребує виконання певних передумов щодо розподілу даних,

таких як нормальність. Це робить його корисним для випадків, коли розподіл даних не є нормальним або коли відсутність гомоскедастичності не дозволяє застосування параметричних методів, таких як t-тест.

Ефективність в порівнянні більш ніж двох груп: критерій Даннета може бути застосований для порівняння трьох або більше груп даних, що робить його корисним для множинного порівняння.

Узгоджені ранги: Критерій Даннета використовує узгоджені ранги для порівняння груп, що дозволяє враховувати порядок даних, а не тільки їх абсолютні значення. Це особливо корисно у випадку, коли розподіл даних має велику варіабельність або відсутність нормальності.

Стійкість до викидів: Критерій Даннета досить стійкий до викидів у порівнянні з параметричними методами, такими як t-тест.

Універсальність застосування: Критерій Даннета може бути застосований в різних областях науки та досліджень, включаючи медицину, соціальні науки, економіку та інші.

Отже, обґрунтування застосування критерію Даннета полягає в його ефективності у випадках, коли параметричні методи не можуть бути використані через відхилення від передумов щодо розподілу даних та у його здатності до множинного порівняння груп.

## **2.2 Опис алгоритму використання критерію у випадку відмінностей розподілу ознаки від нормального.**

Опис алгоритму використання критерію у випадку відмінностей розподілу ознаки від нормального

Для використання критерію Даннета у випадку відмінностей розподілу ознаки від нормального потрібно виконати наступні кроки:

1. Перевірка передумов: перш за все, необхідно перевірити, чи виконуються передумови для застосування критерію Даннета. Одна з основних передумов - незалежність вибірок та абсолютна неперервність розподілу. Також важливо переконатися, що вибірки не містять великі викиди або аномальні значення.
2. Ранжування даних: для кожної вибірки потрібно впорядкувати значення за

зростанням, присвоївши їм відповідні ранги. Якщо вибірки мають однакові значення, їм присвоюються середні ранги.

3. Обчислення сум рангів: Обчисліть суму рангів для кожної вибірки. Нехай сума рангів для першої вибірки буде  $C1C1$ , а для другої -  $C2C2$

4. Обчислення статистики критерію: Обчисліть значення статистики критерію Даннета за формулою:

$$D = \frac{C_1 - C_2}{\sqrt{\frac{n_1 + n_2 - 1}{n_1 n_2}}}$$

так само  $n_1$  і  $n_2$  - розміри першої та другої вибірки відповідно.

5. Інтерпретація результатів: отримане значення статистики  $D$  порівнюється з критичним значенням з таблиці критерію Даннета для заданого рівня значущості  $\alpha$  та розмірів вибірок  $n_1$  і  $n_2$ . Якщо обчислена статистика перевищує критичне значення, то можна відхилити нульову гіпотезу про відсутність відмінності між групами.

Це загальний алгоритм застосування критерію Даннета у випадку відмінностей розподілу ознаки від нормального. Дотримання цього алгоритму допоможе отримати достовірні результати та зробити висновки про статистичну значущість різниці між вибірками.

### 2.3 Розгляд результатів порівнянь

Після проведення порівняльного аналізу за допомогою критерію Даннета можна розглянути наступні можливі результати:

Статистично значущі відмінності між групами: якщо значення статистики критерію Даннета перевищує критичну область, то це свідчить про наявність статистично значущих відмінностей між групами даних. У цьому випадку можна прийняти альтернативну гіпотезу про те, що є різниця між групами щодо вивченої ознаки.

Відсутність статистично значущих відмінностей між групами: якщо значення статистики критерію Даннета не перевищує критичну область, то це свідчить про відсутність статистично значущих відмінностей між групами даних. У цьому випадку не можна відхилити нульову гіпотезу про відсутність різниці

між групами.

Неоднозначність результатів: іноді результати порівнянь можуть бути неоднозначними або недостатньо впевненими. Це може статися, наприклад, коли значення статистики критерію Даннета знаходиться дуже близько до критичного значення. У цьому випадку рекомендується провести додаткові дослідження або збільшити обсяг вибірки для отримання більш точних результатів.

Після розгляду результатів порівнянь важливо правильно інтерпретувати отримані дані та прийняти обґрунтоване рішення щодо наявності чи відсутності статистично значущих відмінностей між групами даних.

## РОЗДІЛ III. Практичне застосування

### 3.1 Приклади порівняння множин даних з різними розподілами ознак та



**застосування критерію Даннета для їх аналізу**

У даному розділі ми розглянемо кілька реальних прикладів застосування критерію Даннета для порівняння множин даних з різними розподілами ознак. Це допоможе нам краще зрозуміти, як критерій Даннета може бути використаний для аналізу даних.

**Приклад 1: Порівняння ефективності двох лікувальних методів**

Уявімо, що проводиться дослідження ефективності двох методів лікування пацієнтів з гіпертонією. Одна група пацієнтів отримує новий лікарський засіб, а інша - традиційний метод лікування. Дослідники хочуть дізнатися, чи є значущі відмінності в ефективності лікування між цими двома групами.

Після аналізу даних виявляється, що розподіл ефектів лікування в обох групах не є нормальним. Для вирішення цього питання вони використовують критерій Даннета, який не вимагає нормального розподілу. Після застосування критерію Даннета вони отримують значення  $p$ -рівня значущості, яке дозволяє їм зробити висновок щодо ефективності лікування.

**Приклад 2: Порівняння впливу двох рекламних стратегій**

Фірма, що займається продажем товарів, хоче визначити, яка рекламна стратегія більш ефективна. Вони проводять дослідження, в якому випадковим чином обираються дві групи споживачів. Одна група отримує інформацію за допомогою традиційних рекламних засобів, а інша - через інтернет-рекламу. Після аналізу результатів виявляється, що розподіл ефективності рекламних стратегій не є нормальним. Тому для аналізу даних вони використовують критерій Даннета. Після обчислення статистики критерію вони зроблять висновок щодо ефективності кожної рекламної стратегії.

**Приклад 3: Порівняння ефективності вакцин**

Вчені проводять клінічне дослідження ефективності двох різних вакцин для запобігання захворювання на грип. Після вакцинації вони вимірюють рівень антитіл у крові кожного учасника дослідження.

Після збору даних вони виявляють, що розподіл рівня антитіл в обох групах не є нормальним. Тому вони використовують критерій Даннета для

порівняння ефективності вакцин. Результати аналізу допоможуть їм зробити висновок про те, яка вакцина є більш ефективною.

Ці приклади демонструють застосування критерію Даннета для порівняння множин даних з різними розподілами ознак та інтерпретацію отриманих результатів. Критерій Даннета є корисним інструментом для аналізу даних у випадках, коли розподіл ознаки не є нормальним.

## ВИСНОВКИ

У результаті проведеного дослідження було вивчено множинні порівняння у випадку відмінності закону розподілу ознаки від нормального та застосування критерію Даннета для оцінки статистично значущих відмінностей між групами даних. Основні висновки які я дослідила: Методи порівняння груп даних: досліджено різні методи порівняння груп даних, зокрема параметричні та непараметричні методи. Показано, що параметричні методи, такі як t-тест чи аналіз дисперсії (ANOVA), вимагають виконання певних передумов, включаючи нормальний розподіл даних. У випадку, коли ці передумови не виконуються, варто використовувати непараметричні методи, які менш чутливі до відхилень від нормального розподілу.

Критерій Даннета як непараметричний метод: критерій Даннета виявився корисним і ефективним непараметричним методом для порівняння трьох або більше груп даних, особливо у випадках, коли розподіл даних не є нормальним. Використання цього критерію дозволяє виявити статистично значущі відмінності між групами, не вимагаючи нормального розподілу даних.

Підтвердження наявності статистично значущих відмінностей: якщо результати аналізу показали, що значення статистики критерію Даннета перевищують критичну область, то можна зробити висновок про наявність статистично значущих відмінностей між групами даних. Це може бути важливою інформацією для подальших досліджень або прийняття рішень. Відсутність статистично значущих відмінностей: якщо результати аналізу показали, що значення статистики критерію Даннета не перевищують критичну область, то можна зробити висновок про відсутність статистично значущих відмінностей між групами даних. Це означає, що немає достатніх доказів для того, щоб вважати групи відмінними одна від одної за досліджуваною ознакою.

Неоднозначність результатів: іноді результати порівнянь можуть бути неоднозначними або недостатньо впевненими. У такому випадку важливо провести додатковий аналіз або дослідження для отримання більш точних результатів.

Отже, на основі отриманих результатів можна зробити висновок, що критерій Даннета є важливим інструментом для дослідження випадків, коли порівнюються групи даних з відмінностями в їхньому розподілі. Його застосування дозволяє здійснювати об'єктивне порівняння і виявлення статистично значущих різниць між групами, що є важливим в наукових дослідженнях та прийнятті обґрунтованих рішень.



## Схожість

Джерела з Інтернету

63

1	<a href="https://mindthegraph.com/blog/uk/%d0%be%d0%b4%d0%bd%d0%be%d1%81%d1%82%d0%be%d1%80%d0%be%d0%">https://mindthegraph.com/blog/uk/%d0%be%d0%b4%d0%bd%d0%be%d1%81%d1%82%d0%be%d1%80%d0%be%d0%</a>	5 джерел	0.51%
2	<a href="http://rkr.nuczu.edu.ua/images/053/mag/2023/Brusenko.pdf">http://rkr.nuczu.edu.ua/images/053/mag/2023/Brusenko.pdf</a>		0.48%
3	<a href="http://er.ucu.edu.ua:8080/bitstream/handle/1/4224/%d0%91%d0%9e%d0%9b%d0%94%d0%98%d0%a0%d0%84%d0%">http://er.ucu.edu.ua:8080/bitstream/handle/1/4224/%d0%91%d0%9e%d0%9b%d0%94%d0%98%d0%a0%d0%84%d0%</a>	40 джерел	0.46%
4	<a href="https://vnu.edu.ua/sites/default/files/2021-11/_%D0%9D%D0%94%D0%9F%D0%94_%D1%81%D0%B8%D0%BB%D0%BC">https://vnu.edu.ua/sites/default/files/2021-11/_%D0%9D%D0%94%D0%9F%D0%94_%D1%81%D0%B8%D0%BB%D0%BC</a>	7 джерел	0.31%
5	<a href="http://khnu.km.ua/root/res/2-25-132-10.pdf">http://khnu.km.ua/root/res/2-25-132-10.pdf</a>		0.25%
6	<a href="https://ela.kpi.ua/bitstream/123456789/23178/1/Biomerty2.pdf">https://ela.kpi.ua/bitstream/123456789/23178/1/Biomerty2.pdf</a>		0.23%
7	<a href="https://dspace.nuft.edu.ua/server/api/core/bitstreams/8a9b2ba6-9704-4925-b8f8-1b8abee6a9f3/content">https://dspace.nuft.edu.ua/server/api/core/bitstreams/8a9b2ba6-9704-4925-b8f8-1b8abee6a9f3/content</a>	5 джерел	0.2%
8	<a href="http://dspace.nuft.edu.ua/bitstream/123456789/22553/1/51.pdf">http://dspace.nuft.edu.ua/bitstream/123456789/22553/1/51.pdf</a>		0.2%
9	<a href="https://ir.nmu.org.ua/xmlui/bitstream/handle/123456789/164925/%d0%96%d1%83%d1%87%d0%ba%d0%be%d0%b2%d0%9...">https://ir.nmu.org.ua/xmlui/bitstream/handle/123456789/164925/%d0%96%d1%83%d1%87%d0%ba%d0%be%d0%b2%d0%9...</a>		0.2%
10	<a href="https://science.snau.edu.ua/wp-content/uploads/2023/12/%D0%94%D0%B8%D1%81-%D0%9A%D0%B0%D0%BB%D1%8E%D0...">https://science.snau.edu.ua/wp-content/uploads/2023/12/%D0%94%D0%B8%D1%81-%D0%9A%D0%B0%D0%BB%D1%8E%D0...</a>		0.2%