**Гіпотези про параметри нормальних розподілів**

**1. Гіпотеза про сподівання**

Нехай  (1) – ряд незалежних спостережень проведених в однакових умовах над незалежною нормальною статистичною змінною . Потрібно перевірити нульову гіпотезу  про те, що математичне сподівання генеральної сукупності , з якої отримана ця вибірка, дорівнює параметру .

, де – деяка стала.

На основі вибірки (1) обчислюємо середнє та варіансу

, .

Розглянемо статистику:

. (2)

Англійський статистик В.Госсет (Gosset) у 1908 році в роботі під псевдонімом Стьюдент довів, що статистика  має розподіл із густиною

, .

Густину називають **густиною розподілу Стьюдента**з ступенями вільності: .

При  отримаємо густину Коші

.

На підставі густини розподілу  та статистики Стьюдента (2) будується критерій істотності для відхилення вибіркового середнього значення  від гіпотетичного значення .

Зазначимо, що при  густина  збігається до нормально розподіленої густини.

На основі розподілу Стьюдента визначаємо критичну область для заданого рівня значущості  та числа ступенів вільності . Вона складається з двох частин, симетричних відносно ординат, кожна з яких має площу по /2. Статистику (2) називають **статистикою Стьюдента**. Для статистики Стьюдента складені таблиці критичних значень для різних ступенів вільності (додаток 7).

Сформулюємо алгоритм перевірки нульової гіпотези за допомогою критерію Стьюдента:

1. Вибираємо рівень значущості .
2. Обчислюємо за формулою (2) емпіричне значення статистики Стьюдента, визначивши попередньо за формулою (1) середнє вибіркове та варіансу.
3. При /2 та кількості ступенів вільності  з додатка 7 знаходимо критичне значення цієї статистики.
4. Якщо , то нульову гіпотезу про те, що , приймаємо. У протилежному випадку вважаємо, що вона суперечить експериментальним даним і може бути відхилена.

|  |  |  |  |
| --- | --- | --- | --- |
| **1,1** | **A(x)** | **B(y)** | **Різниця z=x-y** |
| 1 | 1,9 | 0,7 | 1,2 |
| 2 | 0,8 | -1,6 | 2,4 |
| 3 | 1,1 | -0,2 | 1,3 |
| 4 | 0,1 | -1,2 | 1,3 |
| 5 | -0,1 | -0,1 | 0,0 |
| 6 | 4,4 | 3,4 | 1,0 |
| 7 | 5,5 | 3,7 | 1,8 |
| 8 | 1,6 | 0,8 | 0,8 |
| 9 | 4,6 | 0,0 | 4,6 |
| 10 | 3,4 | 2,0 | 1,4 |

У класичній роботі про - розподілСтьюдентнаводить такі дані про додаткові години сну в 10 пацієнтів, викликаними снодійними препаратами  та .

Перевірити, чи істотна різниця між дією двох снотворних засобів  та . Якщо припустити, що різниця між додатковими годинами сну викликана дією засобів  та , тоді  буде нормально розподіленою вибіркою.

Формулюємо нульову гіпотезу:  – немає істотної різниці між дією двох снотворних засобів.

Доведення проведемо за допомогою критерію Стьюдента*.*

Розглянемо різницю . На основі експериментальних даних знаходимо середню різницю. Та стандарт

.

Відношення Стьюдента дорівнює

.

Критичне значення знаходимо з таблиці для  та числа ступенів вільності  . Оскільки то гіпотезу про еквівалентність снодійних засобів відкидаємо, тобто існує істотна різниця між дією двох снотворних засобів.

**2. Інтервал довіри для невідомого сподівання**

На основі області прийому гіпотези про сподівання нормальної величини можемо записати співвідношення, тобто при визначенні критичної області для критерію Стьюдента одержуємо таке співвідношення

.

Умова не зміниться, якщо її замінити на еквівалент

,

,

.

Отже з імовірністю  випадковий інтервал:

 (2)

накриває невідоме сподівання  нормально розподіленої генеральної сукупності. Цей інтервал називають **інтервалом довіри,** або **довірчим інтервалом** для сподівання , при рівні значущості .

**Приклад.** Проведене 31 спостереження над нормально розподіленою статистичною змінною , на основі яких одержали середнє вибіркове.

Знайти 90% інтервал довіри для невідомого сподівання генеральної сукупності.

У цьому випадку . З таблиці, при таких вихідних даних, беремо  (додаток 7)

Шуканим інтервалом згідно формули (2) є

, який з ймовірністю 90 % містить невідоме сподівання цієї генеральної сукупності.

**3. Порівняння сподівань двох нормальних розподілів**

Нехай  ряд незалежних спостережень над незалежною нормальною статистичною змінною , а  ряд незалежних спостережень над незалежною нормальною статистичною змінною . Перевірити гіпотезу про те, що математичне сподівання популяції, з яких одержані ці вибірки мають однакове математичне сподівання, тобто: .

На основі вибірок знаходимо середні



та варіанси

.

Виберемо рівень значущості α та розглянемо статистику:

.

У 1925 році Р.Фішер довів, що статистика має розподіл Стьюдента з кількістю ступенями вільності . Наступна частина аналогічна попередньо розглянутому випадку.

**Приклад**. В одному класі з 20 дітей випадково вибирають 10, яким щоденно почали видавати апельсиновий сік. Решта 10 щоденно отримували молоко. Через деякий час зафіксоване таке збільшення ваги дітей (у фунтах):

*I група: 4 2,5 3,5 4 1,5 1 3,5 3 2,5 3,5*

*II група:1,5 3,5 2,5 3 2,5 2 2 2,5 1,5 3*

Середній приріст ваги в першій групі дорівнює 2,9 а в другій групі – 2,4. Чи істотна ця різниця?

Формулюємо нульову гіпотезу  : нема істотної різниці між збільшенням ваги у 2-х групах. Використаємо критерій Стьюдента.



При  і : .

Гіпотезу приймаємо, бо 

## **4. Гіпотеза про дисперсію нормальної популяції**

Дисперсія характеризує точність машин і приладів, точність технологічного процесу, похибку показань вимірювального приладу і таке інше. Тому важливо приймати рішення про гіпотези відносно дисперсії.

Нехай  ряд незалежних спостережень над нормально розподіленою статистичною змінною .

Потрібно перевірити гіпотезу про те, що дисперсія нормальної популяції з якої взята вибірка дорівнює , тобто, .

За даними вибірки обчислимо її середнє та варіансу:

.

Виберемо рівень значущості  та розглянемо статистику . Якщо гіпотеза вірна, то ця статистика має розподіл , де  число ступенів вільності. На основі розподілу статистики  визначаємо критичну область для гіпотези. Очевидно, що для гіпотези будуть сприятливі ті випадки, коли значення статистики  близьке до 1. Тому критична область для статистики складається із двох частин: дуже малих і дуже великих значень змінної . При рівні значущості α нижнім критичним значенням є , а верхнім .

Якщо емпіричне значення статистики  знаходиться між  і , то гіпотезу про те, що дисперсія дорівнює  приймаємо, а в інших випадках відхиляємо.

# **Приклад.** Стандартні екзамени проходили кілька років із сподіванням a=70 і дисперсією . Для деякої групи 25 студентів за результатами їх успішності отримали середнє значення сумарної їх успішності і варіансу . Чи є підстави сумніватися, що дисперсія сумарної успішності всіх студентів буде дорівнювати 9? .

# Знаходимо .

При  і  .

Підстав сумніватися немає. Гіпотезу приймаємо.

## **5. Інтервал довіри для невідомої дисперсії нормальної популяції**

Розподіл статистики  дає можливість при заданому рівні значущості  на підставі вибірки вказати з імовірністю 1- інтервал довіри для невідомої дисперсії .

При вибраному числу значущості  і числу степенів вільності  для яких виконується співвідношення

,

отримаємо інтервал довіри для дисперсії нормальної популяції

.

Отже з імовірністю 1-α випадковий інтервал  накриває невідоме значення дисперсії генеральної сукупності. Цей інтервал називають  **інтервалом довіри**, або **довірчим інтервалом** для дисперсії  при рівні значущості α.

**Приклад**. Проведене 31 спостереження над нормально розподіленою статистичною змінною . Одержали .

Знайти 90 % інтервал довіри до дисперсії генеральної сукупності.

За умовою задачі  і . Тому з таблиці [] маємо .

Шуканий інтервал  є 90 %-им інтервалом довіри для невідомої дисперсії генеральної сукупності.