



Нотатки до завдання 6

p-значення та H_0

У статистиці часто перевіряють гіпотези:

- **Нульова гіпотеза (H_0):** Зазвичай стверджує, що **немає ефекту**, різниця випадкова.
- **Альтернативна гіпотеза (H_1):** Є **справжній ефект**, зміна, зв'язок тощо.

"Різниця" — це відхилення від того, що ми очікуємо, якщо H_0 істинна.

P-значення — це число, яке показує, **наскільки ймовірно**, що результат, який ми бачимо в даних, міг би з'явитися **випадково, якщо нульова гіпотеза є істинною**.

P-значення — це ймовірність отримати такі або більш екстремальні результати, як у твоїх даних, якщо H_0 є істинною.

p-value	Інтерпретація
< 0.05	Результат статистично значущий — різниця навряд чи випадкова, відхиляємо H_0
≥ 0.05	Недостатньо доказів проти H_0 — ми не відхиляємо H_0

Чим менше p-значення, тим більше підстав відкинути H_0

Приклади

P-значення	Сила доказів проти H_0
0.005	Дуже сильні
0.1	Слабкі
0.32	Дуже слабкі
0.56	Практично ніякі
0.94	Майже неможливо, що H_0 хибна

Дослідження	P-значення	Сила доказів проти H_0
A	0.002	Дуже сильні
B	0.2	Дуже слабкі

Формальні рішення

Якщо p-значення є малим:

- **ВІДХИЛЯЄМО H_0**
- Вибірка була би дуже екстремальною, якщо би H_0 була правильною (то **такий результат, який ми отримали — дуууже малоймовірний**, майже як диво).
- Результати є статистично значущими (тобто навряд чи є випадковістю)
- Є докази правильності H_1

Якщо p-значення не є малим:

- **НЕ ВІДХИЛЯЄМО H_0**
- Вибірка не була би надто екстремальною, якщо би H_0 була правильною
- Результати не є статистично значущими
- Або H_0 або H_a можуть бути правильними

Рівень значущості

Рівень значущості, позначається як α (альфа), це **межа**, яку ми встановлюємо заздалегідь, щоб вирішити, коли відкинути нульову гіпотезу H_0 .

Формально,

α — це максимальна ймовірність того, що ми помилково відкинемо H_0 , якщо насправді вона **істинна**. Тобто ймовірність допустити помилку першого роду.

Рівень значущості альфа, є пороговим значенням, p-значення нижчі за нього є достатньо малими, щоб ми відхилили нульову гіпотезу.

- Якщо p-значення є меншим за альфа, результати є *статистично значущими*, і ми відхиляємо нульову гіпотезу на користь альтернативної
- Якщо p-значення не є меншим за альфа, результати не є статистично значущими, тож ми не можемо зробити висновок на основі тесту

- Часто альфа = 0.05 без обумовлення, якщо не вказано інше

Обираємо рівень значущості альфа

- Зазвичай, $\alpha = 0.05$
- Якщо помилка I типу (відхилити правильну нульову гіпотезу) є гіршою, ніж помилка II типу, можна обрати менше $\alpha = 0.01$
- Якщо помилка II типу (не відхилити хибну нульову гіпотезу) є гіршою, ніж помилка I типу, можна обрати $\alpha = 0.10$

Статистична VS Практична значущість

- При малому розмірі вибірки навіть велика різниця може бути незначущою (адже це може статися через брак даних)
- При великому розмірі вибірки навіть дуже маленька різниця може бути значущою (тобто ефект не є випадковістю)
- Статистично значущий результат не завжди є практично значущим, особливо при великому розмірі вибірки (те, що результат **не випадковий**, ще не означає, що він **корисний** або **важливий** у реальному житті)

T-test

T-тест (або **тест Стьюдента**) — це **статистичний тест**, який дозволяє перевірити, **чи дійсно є різниця між середніми значеннями** двох груп, чи ця різниця могла виникнути випадково.

Python

```
from scipy import stats

# Прикладні дані
group1 = [23, 25, 22, 20, 21] # Наприклад, оцінки до
group2 = [27, 29, 26, 28, 30] # Оцінки після

# Незалежний T-тест для 2 груп
t_stat, p_value = stats.ttest_ind(group1, group2)
```

```
print("T-статистика:", t_stat)
print("p-value:", p_value)
```

Результат тесту

Результатом **T-тесту** є **p-value**, а також **t-статистика**.

1. T-статистика (t-statistic):

- Це числове значення, яке оцінює, наскільки сильно відрізняються середні значення двох груп, порівняно з їхнім стандартним відхиленням.
- Чим більше значення t-статистики, тим більше різниця між групами в порівнянні з їх варіацією.

2. P-value:

- Це ймовірність того, що спостережувана різниця між групами могла виникнути випадково, тобто, ймовірність того, що нульова гіпотеза (що між групами немає різниці) правильна.
- Якщо **p-value** менше за рівень значущості (зазвичай 0.05), то відкидаємо нульову гіпотезу і вважаємо результат статистично значущим. Інакше ми не відкидаємо нульову гіпотезу.
- **t-статистика** допомагає зрозуміти, наскільки сильно змінюється середнє значення однієї групи порівняно з іншою.
- **p-value** дає ймовірність того, що ця різниця є випадковою, тобто, чи можна відкинути нульову гіпотезу про відсутність різниці. (якщо значення мале → різниця навряд чи є випадковістю, тобто різниця є фактом → відхиляємо гіпотезу про відсутність різниці).

Приклад

```
t_stat, p_value = stats.ttest_ind(group1, group2)
print(f"T-статистика: {t_stat}")
print(f"P-значення: {p_value}")
```

Якщо:

- **t-статистика = 2.5**

- **p-value = 0.03**

Це означає, що різниця між середніми значеннями двох груп значуща, оскільки p-value менше 0.05, і ми можемо відкинути нульову гіпотезу про відсутність різниці.

Знак t-значення

- **Додатне t** — середнє в першій групі **більше** за другу.
- **Від'ємне t** — середнє в першій групі **менше** за другу.
- **Близьке до нуля t** — середні значення майже однакові.