

ĐẠI HỌC BÁCH KHOA HÀ NỘI
TRƯỜNG CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG (SOICT)
KHOA KỸ THUẬT MÁY TÍNH

Sign Test

Kiểm định Dấu: Từ Lý thuyết đến Python

GVHD: TS. Đoàn Phong Tùng

Nhóm thực hiện:

Đào Duy Tân	20251334M
Nguyễn Đức Dũng	20251240M
Trương Tuấn Nghĩa	20251196M
Hồ Hải Anh	20252102M
Đinh Lâm Nghị	20241116M

Hà Nội, Tháng 12/2025

Mục lục Tổng quan I

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

④ Demo Python/Jupyter

⑤ FAQs

⑥ Kết luận

⑦ Câu hỏi & Trả lời

Mục lục |

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

④ Demo Python/Jupyter

⑤ FAQs

⑥ Kết luận

⑦ Câu hỏi & Trả lời

Sign Test trong Lịch sử - John Arbuthnot (1710s) |

John Arbuthnot (1667-1735)

- Bác sĩ Hoàng gia Anh
- Nhà toán học, nhà văn
- Nghiên cứu: Tỷ lệ sinh nam/nữ ở London

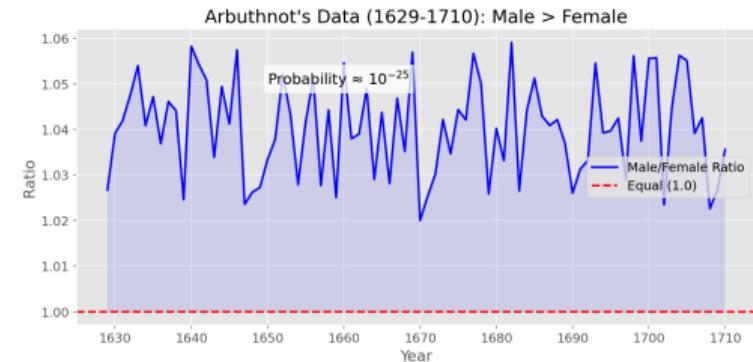
Phát hiện đáng chú ý:

- 82 năm liên tiếp (1629-1710)
- Số trẻ em nam sinh > Số trẻ em nữ
- Ngẫu nhiên hay có ý nghĩa?

Kết luận lịch sử:

- Nếu ngẫu nhiên: $P = (1/2)^{82} \approx 10^{-25}$
- → **Bác bỏ giả thuyết ngẫu nhiên!**
- Một trong những kiểm định đầu tiên

Nguồn: Arbuthnot, J. (1710). "An argument for Divine Providence..."



Ý nghĩa

Ứng dụng sớm nhất của thống kê

Các Ứng dụng Lịch sử Nổi tiếng I



Y học - 1747

James Lind

Điều trị bệnh Scurvy

- 12 thủy thủ
- 6 phương pháp
- So sánh cam/chanh

Tâm lý - 1930s

J.B. Rhine

Thí nghiệm ngoại cảm

- Đoán bài Zener
- Đúng/Sai?
- Kiểm tra ESP

Kinh tế - 1960s

Stock Analysis

Dự đoán thị trường

- Tăng/giảm?
- So với baseline
- Skilled vs random

Tại sao Sign Test vẫn Quan trọng Hôm nay? I

Ưu điểm vượt thời gian:

① Đơn giản

- Chỉ đếm dấu +/-
- Dễ giải thích

② Robust

- Không sợ outliers
- Không cần chuẩn hóa
- OK với mẫu nhỏ

③ Linh hoạt

- Áp dụng rộng
- Dữ liệu thứ bậc OK

Ứng dụng hiện đại:

Marketing (A/B)

Clinical Trials

User Experience

Quality Control



Mẫu chốt

300+ năm tuổi nhưng vẫn hữu ích!

Mục lục I

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

④ Demo Python/Jupyter

⑤ FAQs

⑥ Kết luận

⑦ Câu hỏi & Trả lời

Sign Test là gì? I

Định nghĩa 1 (Kiểm định Dấu - Sign Test)

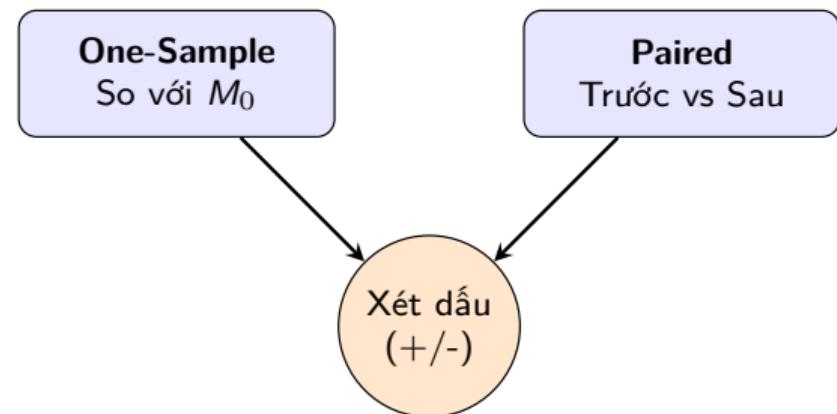
Kiểm định phi tham số sử dụng **dấu (+/-)** để kiểm tra median hoặc so sánh mẫu phụ thuộc.

Hai dạng chính: 1. One-sample

- Kiểm tra: Median = M_0 ?
- H_0 : Median = M_0

2. Paired

- So sánh trước-sau
- H_0 : Median(Diff) = 0



Côt lõi: Quy về nhị thức

Nguyên lý Hoạt động - 4 Bước I

1. Tính hiệu số (d_i)



2. Gán dấu (+, -, 0)



3. Đếm S^+ (Loại bỏ 0)



4. Kiểm định (Binomial)

① Tính hiệu số

- One: $d_i = X_i - M_0$
- Paired: $d_i = X_i - Y_i$

② Xác định dấu

- $d_i > 0 \rightarrow S^+$
- $d_i < 0 \rightarrow S^-$
- $d_i = 0 \rightarrow \text{Loại}$

③ Thống kê

- $S^+ \sim \text{Binom}(n, 0.5)$

④ P-value

- So với $\alpha = 0.05$

Giả thuyết & Công thức I

Giả thuyết & Công thức II

Cặp giả thuyết:

Hypotheses

- $H_0: \text{Median}(D) = 0$
- $H_1: \text{Median}(D) \neq 0$ (2-tailed)

Mức ý nghĩa: $\alpha = 0.05$

Thống kê:

Test Statistic

$$S^+ \sim \text{Binomial}(n, 0.5)$$

Quyết định:

- $p < \alpha \rightarrow$ Bác bỏ H_0
- $p \geq \alpha \rightarrow$ Giữ H_0

Mục lục I

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

④ Demo Python/Jupyter

⑤ FAQs

⑥ Kết luận

⑦ Câu hỏi & Trả lời

Ví dụ 1: Y học - Thuốc giảm đau I

Ví dụ 1: Y học - Thuốc giảm đau II

Bối cảnh:

- 10 bệnh nhân bị đau mãn tính
- Đo điểm đau (VAS scale 0-10) trước và sau dùng thuốc mới
- Câu hỏi: Thuốc có giảm đau không?

Dữ liệu:

ID	Trước	Sau	Hiệu	Dấu
1	8	6	+2	+
2	7	7	0	(loại)
3	9	5	+4	+
4	6	5	+1	+
5	8	7	+1	+
6	7	8	-1	-
7	9	6	+3	+
8	8	6	+2	+
9	7	5	+2	+
10	9	7	+2	+

Phân tích:

- Loại bỏ ID 2 (hiệu = 0)
- $n = 9$ (sau loại ties)
- $S^+ = 8$ (số dấu +)
- $S^- = 1$ (số dấu -)

Kiểm định:

- H_0 : Thuốc không hiệu quả
- H_1 : Thuốc giảm đau (one-tailed)
- $\alpha = 0.05$

Kết quả:

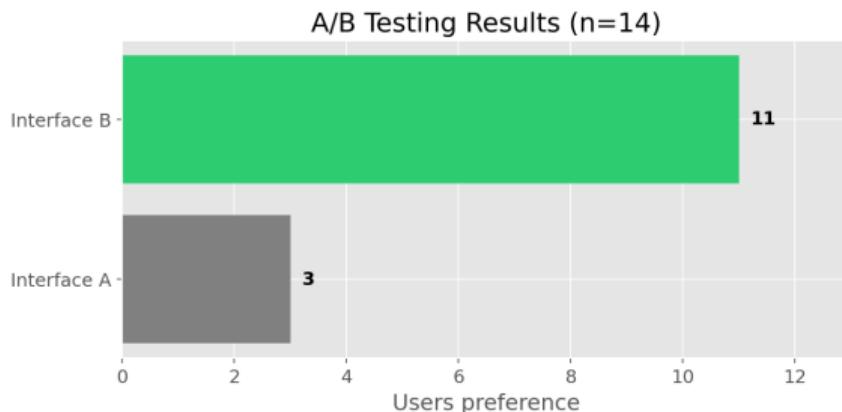
- p-value = **0.039**
- **0.039 < 0.05**
- → **Bắc bỏ H_0**
- Có bằng chứng thuốc hiệu quả!

Ví dụ 2: Marketing - A/B Testing Website I

Ví dụ 2: Marketing - A/B Testing Website II

Tình huống:

- Công ty test 2 giao diện website (A vs B)
- 15 người dùng test cả 2 phiên bản
- Đánh giá mức độ hài lòng (1-5 sao)
- **Câu hỏi:** Giao diện B có tốt hơn A không?



Dữ liệu tóm tắt:

- Số người thích B hơn A: **11**
- Số người thích A hơn B: **3**
- Số người không phân biệt: **1** (loại)
- $n = 14$ (sau loại ties)

Sign Test:

- $S^+ = 11$ (thích B hơn)
- p-value = **0.059** (two-tailed)

Quyết định:

- Với $\alpha = 0.05$: **Chưa đủ bằng chứng** (biên giới!)
- Với $\alpha = 0.10$: Có bằng chứng B tốt hơn
- → Có thể cần thêm dữ liệu

Ví dụ 3: Tâm lý - Nghiên cứu Sở thích Màu sắc I

Ví dụ 3: Tâm lý - Nghiên cứu Sở thích Màu sắc II

Câu hỏi nghiên cứu:

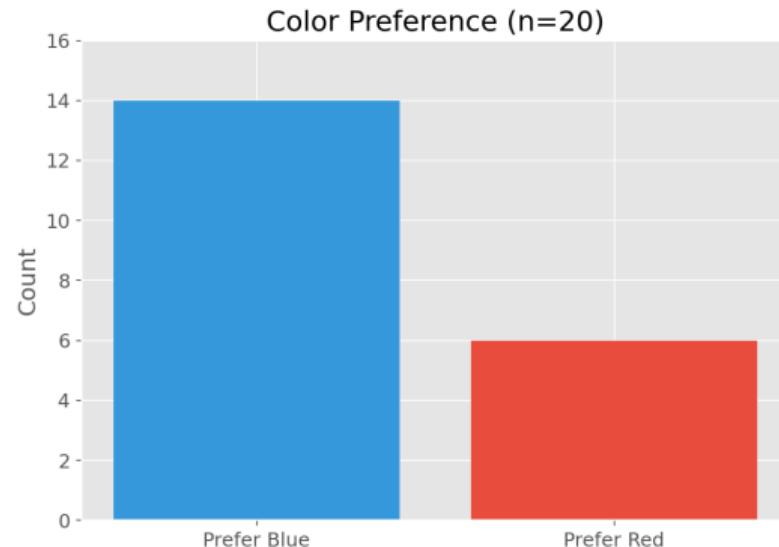
- Người ta có thích màu xanh hơn màu đỏ không?
- (So với ngẫu nhiên 50-50)

Thiết kế nghiên cứu:

- 20 người tham gia
- 10 items (vật phẩm)
- Mỗi item: Chọn 1 trong 2 màu
- Đếm số lần chọn màu xanh

Kết quả:

- 14/20 người chọn xanh nhiều hơn đỏ
- 6/20 người chọn đỏ nhiều hơn xanh
- Không có ties (ai cũng có preference)



Phân tích Sign Test:

- H_0 : Không có preference ($p = 0.5$)
- H_1 : Có preference ($p \neq 0.5$)

Mục lục |

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

④ Demo Python/Jupyter

⑤ FAQs

⑥ Kết luận

⑦ Câu hỏi & Trả lời

Demo Setup - Ví dụ Thuốc giảm cân

Bài toán:

- 12 người tham gia thử nghiệm thuốc giảm cân
- Đo cân nặng trước và sau 2 tháng sử dụng
- Câu hỏi: Thuốc có hiệu quả không?

Chuẩn bị Jupyter Notebook:

Listing 1: Cell 1: Import libraries

```
import numpy as np
import pandas as pd
from scipy.stats import binomtest
import matplotlib.pyplot as plt

# Du lieu: Can nang 12 nguoi
data = {
    'ID': range(1, 13),
    'Truoc_kg': [85, 90, 78, 92, 88, 76, 95, 82, 79, 91, 87, 84],
    'Sau_kg': [82, 88, 79, 90, 85, 77, 93, 80, 78, 89, 86, 83]
}

df = pd.DataFrame(data)
print(df.head())
```

In [1]: df.head()

Nhóm SV HUST (HUST)

Kiểm định Dâu (Sign Test)

Demo Step 1: Tính hiệu số và xác định dấu

Listing 2: Cell 2: Calculate differences and signs

```
# Buoc 1: Tinh hieu (truoc - sau)
df['Hieu_kg'] = df['Truoc_kg'] - df['Sau_kg']

# Buoc 2: Xac dinh dau
def assign_sign(x):
    if x > 0: return '+'
    elif x < 0: return '-'
    else: return '0'

df['Dau'] = df['Hieu_kg'].apply(assign_sign)
df_filtered = df[df['Hieu_kg'] != 0]

# Hien thi ket qua
print(df_filtered[['ID', 'Truoc_kg', 'Sau_kg', 'Hieu_kg', 'Dau']])
```

ID	Truoc	Sau	Hieu	Dau
1	85	82	3	+
2	90	88	2	+
3	78	79	-1	-
4	92	90	2	+
5	88	85	3	+

Demo Step 2: Thống kê & Kiểm định

Listing 3: Cell 3: Perform Sign Test

```
# Dem so dau +
n_plus = (df_filtered['Hieu_kg'] > 0).sum()
n_total = len(df_filtered)
print(f"So_dau+: {n_plus}/{n_total}")

# Kiem dinh Binomial (one-tailed: greater)
result = binomtest(n_plus, n_total, p=0.5, alternative='greater')
print(f"P-value: {result.pvalue:.4f}")

# Ket luan
alpha = 0.05
if result.pvalue < alpha:
    print(f"BAC BO HO (p={result.pvalue:.4f}<{alpha})")
    print("=> CO BANG CHUNG: Thuoc co hieu qua!")
else:
    print(f"GIU HO")
```

So dau +: 9/11

P-value: 0.0327

Demo Step 3: Visualization

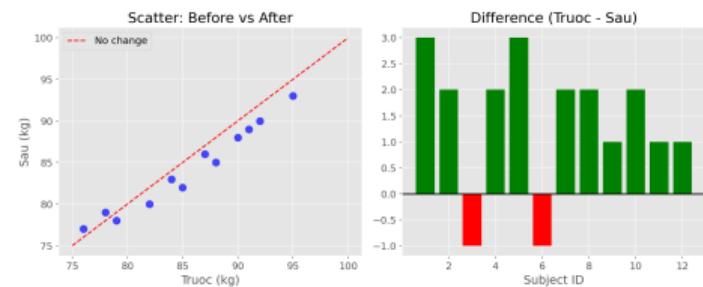
Listing 4: Cell 4: Visualize results

```
fig, axes = plt.subplots(1, 2, figsize=(12, 5))

# Plot 1: Scatter truoc vs sau
ax1 = axes[0]
ax1.scatter(df['Truoc_kg'], df['Sau_kg'], s=100)
ax1.plot([75, 100], [75, 100], 'r--')
ax1.set_xlabel('Truoc(kg)')
ax1.set_ylabel('Sau(kg)')

# Plot 2: Bar chart hieu so
ax2 = axes[1]
colors = ['green' if x > 0 else 'red' for x in df['Hieu_kg']]
ax2.bar(df['ID'], df['Hieu_kg'], color=colors)
ax2.axhline(0, color='black', lw=2)

plt.tight_layout()
plt.show()
```



Insight từ plots

- Hầu hết điểm dưới đường chéo
- Đa số bars màu xanh (dương)
- Chỉ 2 người tăng cân

Mục lục |

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

④ Demo Python/Jupyter

⑤ FAQs

⑥ Kết luận

⑦ Câu hỏi & Trả lời

FAQs - Câu hỏi Thường gặp (1/2) |

FAQs - Câu hỏi Thường gặp (1/2) II

Q1: Sign Test vs t-test - Khi nào dùng cái nào?

Tiêu chí	Dùng Sign Test	Dùng t-test
Phân phối	Bất kỳ	Gần chuẩn
Outliers	OK	Nhạy cảm
Cỡ mẫu	$n \geq 10$	$n \geq 30$ tốt hơn
Power	Thấp hơn 5-10%	Cao hơn
Quan tâm	Median	Mean

Q2: Tại sao loại bỏ hiệu = 0?

- Không mang thông tin về hướng
- Giảm n nhưng tăng độ chính xác
- Standard practice trong Sign Test

Q3: Sign Test có yếu hơn t-test?

- Đúng! Power thấp hơn nếu data chuẩn
- Nhưng robust hơn nhiều với vi phạm
- Trade-off: Power vs Robustness

FAQs - Câu hỏi Thường gặp (2/2) |

FAQs - Câu hỏi Thường gặp (2/2) II

Q4: Cỡ mẫu tối thiểu là bao nhiêu?

- **Lý thuyết:** $n \geq 5$ (sau khi loại zeros)
- **Thực tế:** $n \geq 10$ để p-value đáng tin cậy
- **Khuyến nghị:** Càng lớn càng tốt! (như mọi test)

Q5: Làm sao biết kết quả có ý nghĩa thực tế?

- P-value chỉ nói về **statistical significance**
- Cần xem thêm:
 - **Effect size:** Bao nhiêu % có dấu +?
 - **Magnitude:** Hiệu số trung bình là bao nhiêu?
 - **Context:** Có quan trọng trong thực tế không?

Q6: Python có thư viện nào tốt?

- `scipy.stats.binomtest` (recommended)
- `statsmodels.stats`
- Tự code cũng dễ!

Q7: Sign Test cho >2 nhóm?

- **Không!** Chỉ cho 1 sample hoặc paired
- Dùng **Friedman test** cho > 2 nhóm
- Hoặc **Cochran's Q** cho binary

Mục lục |

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

④ Demo Python/Jupyter

⑤ FAQs

⑥ Kết luận

⑦ Câu hỏi & Trả lời

Take-home Messages I

Take-home Messages II

1. Sign Test = Simple but Powerful

- 300+ năm lịch sử, vẫn cực kỳ hữu ích
- Dễ hiểu, dễ thực hiện
- Không cần giả định phức tạp

2. Khi nào nên dùng Sign Test?

- ✓ Dữ liệu **không phân phối chuẩn**
- ✓ Có **outliers** hoặc extreme values
- ✓ Quan tâm đến **median** hơn mean
- ✓ Mẫu nhỏ ($n = 10-30$)
- ✓ Cần phương pháp **robust**

3. Python makes it easy!

- Chỉ cần `scipy.stats.binomtest`
- Hiểu logic → Code 5 dòng
- Demo trong Jupyter Notebook rất trực quan

Tài liệu Tham khảo & Học thêm I

Tài liệu Tham khảo & Học thêm II

Sách giáo khoa:

- Hollander, M., Wolfe, D. A., & Chicken, E. (2013). *Nonparametric Statistical Methods* (3rd ed.). Wiley.
- Gibbons, J. D., & Chakraborti, S. (2020). *Nonparametric Statistical Inference* (6th ed.). CRC Press.
- Daniel, W. W. (2000). *Applied Nonparametric Statistics* (2nd ed.). Duxbury Press.

Tài liệu online:

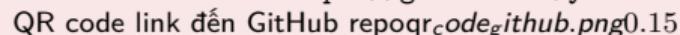
- SciPy Documentation: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.binomtest.html>
- Statistics How To: <https://www.statisticshowto.com/sign-test/>
- Penn State STAT 415: <https://online.stat.psu.edu/stat415/>

Lịch sử:

- Arbuthnot, J. (1710). "An argument for Divine Providence, taken from the constant Regularity observ'd in the Births of both Sexes". *Philosophical Transactions*, 27, 186-190.

GitHub Repository

Code & Notebooks: <https://github.com/yourusername/sign-test-demo>

QR code link đến GitHub repo qr_codegithub.png0.15

Mục lục I

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

④ Demo Python/Jupyter

⑤ FAQs

⑥ Kết luận

⑦ Câu hỏi & Trả lời

Danh sách Câu hỏi Thảo luận

Bấm vào câu hỏi để xem chi tiết:

Q1. So sánh: Sign Test vs Wilcoxon Signed-Rank?

Q2. Xử lý số liệu: Tại sao loại bỏ Ties (0)?

Q3. Dữ liệu: Dùng cho Likert scale được không?

Q4. Cỡ mẫu: Vấn đề khi mẫu nhỏ ($n < 10$)?

Q5. Báo cáo: Viết kết quả chuẩn academic?

Q6. Nâng cao: Multiple comparisons correction?

Q7. Công cụ: Các thư viện Python/R/SPSS?

Q8. Giả định: Assumptions nào bắt buộc?

Q9. Hiệu năng: Power Analysis & Sample Size?

Q10. Mở rộng: Các biến thể (McNemar, Friedman)?

Q1: So sánh Sign Test vs Wilcoxon I

Câu hỏi

Sign Test khác với Wilcoxon Signed-Rank Test như thế nào? Khi nào nên dùng cái nào?

Trả lời chi tiết:

Tiêu chí	Sign Test	Wilcoxon
Thông tin sử dụng	Chỉ dấu (+/-)	Dấu + Độ lớn (Rank)
Power	Thấp hơn (khoảng 64% của t-test)	Cao hơn (khoảng 95% của t-test)
Assumptions	Ít nhất (chỉ cần độc lập)	Thêm: phân phối đối xứng (symmetric)
Dễ hiểu	Rất trực quan	Phức tạp hơn về tính toán
Khi nào dùng	Dữ liệu rất lệch, nhiều Outliers	Dữ liệu gần đối xứng

Khuyến nghị:

- **Sign Test:** Khi dữ liệu rất "ồn" (noisy), nhiều giá trị ngoại lai cực đoan, hoặc chỉ quan tâm đến hướng thay đổi.
- **Wilcoxon:** Khi dữ liệu tương đối đối xứng và muốn đạt độ mạnh (power) cao hơn.

Q2: Xử lý Ties (Hiệu số = 0) |

Câu hỏi

Nếu có nhiều hiệu số bằng 0, có ảnh hưởng gì không? Có cách xử lý nào khác?

Trả lời:

Ảnh hưởng của ties:

- Giảm kích thước mẫu hiệu dụng (Effective sample size).
- Giảm sức mạnh thống kê (Statistical power).
- Cảnh báo:** Nếu $> 20\%$ dữ liệu là ties thì kết quả có thể thiếu tin cậy.

Phương pháp xử lý:

- Standard:** Loại bỏ hoàn toàn (Khuyên dùng).
- Chia đều:** Một nửa vào nhóm (+), một nửa vào nhóm (-).
- Random assign:** Gán ngẫu nhiên dấu.

Ví dụ minh họa:

Tổng $n = 20$, có 5 ties (25% số liệu là 0)

- Cách 1 (Loại):** Phân tích với $n = 15$.
- Cách 2 (Chia):** Tính 2.5 cho (+) và 2.5 cho (-).
- Cách 3 (Random):** Tung xu cho 5 trường hợp đó.

Lưu ý

Cách 1 (loại bỏ) là "standard practice" trong hầu hết các phần mềm thống kê (SPSS, R, Python).

Q3: Áp dụng cho data thứ bậc (Likert Scale) I

Câu hỏi

Sign Test có thể dùng cho dữ liệu thứ bậc (ordinal) như Likert scale (1-5) không?

Trả lời: ĐÚNG! Sign Test cực kỳ phù hợp cho ordinal data.

Ví dụ - Khảo sát hài lòng:

Thiết kế:

- Survey trước và sau cải tiến dịch vụ.
- Scale: 1 (Rất tệ) → 5 (Rất tốt).
- So sánh: Điểm Trước vs Điểm Sau.

Phân tích:

- Sau > Trước (VD: 3 lên 4) → Dấu (+)
- Sau < Trước (VD: 5 xuống 2) → Dấu (-)
- Sau = Trước → Loại

Tại sao phù hợp?

- ① Likert scale là ordinal, khoảng cách giữa 1-2 chưa chắc bằng 4-5.
- ② T-test yêu cầu tính Mean (cộng trừ), điều này không chặt chẽ với Likert.
- ③ Sign Test chỉ cần thứ tự (lớn hơn/nhỏ hơn), không cần đo khoảng cách.

Q4: Vấn đề với Sample size nhỏ |

Câu hỏi

Với $n = 6$ hoặc $n = 7$, Sign Test có đáng tin không? Cần làm gì?

Trả lời:

Vấn đề "Bước nhảy" P-value:

- Do phân phối nhị thức là rời rạc.
- P-values sẽ "nhảy" qua các mốc, khó đạt chính xác 0.05.

Ví dụ thực tế với $n=6$:

- Nếu 5 dấu (+): $p = 0.219$ (Quá lớn)
- Nếu 6 dấu (+): $p = 0.031$ (Nhỏ hơn 0.05)
- Không có p-value nào nằm giữa 0.031 và 0.219!

Giải pháp:

- ① **Exact test:** Luôn dùng Binomial Exact (không dùng xấp xỉ chuẩn).
- ② **Report exact p:** Báo cáo p-value chính xác thay vì chỉ nói < 0.05 .
- ③ **Power:** Chấp nhận rằng power sẽ rất thấp.
- ④ **Data:** Cố gắng thu thập thêm mẫu nếu có thể (ít nhất $n=10$).

Q5: Báo cáo kết quả (Reporting) I

Câu hỏi

Khi viết báo cáo khoa học hoặc luận văn, cần trình bày những thông số gì?

Checklist thông tin cần thiết:

① Sample size:

- Tổng số quan sát ban đầu (N).
- Số ties bị loại bỏ.
- Kích thước mẫu hiệu dụng (n).

② Test statistic:

- Số lượng dấu dương (S^+) và dấu âm (S^-).
- Tỷ lệ phần trăm (VD: 80% cải thiện).

③ P-value:

- Giá trị chính xác (VD: $p = 0.033$).
- Ghi rõ là One-tailed (1 phía) hay Two-tailed (2 phía).

Q5: Báo cáo kết quả (Reporting) II

Ví dụ câu viết mẫu (Standard Reporting):

"A Sign test was conducted to evaluate the effect of the new treatment. Of 12 participants, 1 showed no change and was excluded (effective n=11). Nine participants (81.8%) showed improvement, compared to 2 (18.2%) who worsened. The exact one-tailed Sign test indicated a significant improvement ($p = 0.033$), suggesting the treatment is effective."

Q6: Multiple Comparisons I

Câu hỏi

Nếu thực hiện nhiều Sign Tests (ví dụ: test 5 loại thuốc khác nhau trên cùng 1 nhóm), có cần điều chỉnh gì không?

Trả lời: CẦN! Đây là vấn đề kiểm định đa giả thuyết.

Vấn đề:

- Khi làm k tests độc lập với $\alpha = 0.05$.
- Xác suất mắc ít nhất 1 lỗi loại 1 (False Positive) tăng lên chóng mặt.
- Công thức: $P(\text{Error}) = 1 - 0.95^k$.
- Với 5 tests: $1 - 0.95^5 \approx 22.6\%$ (Rất cao!).

Giải pháp Correction:

- ① **Bonferroni:** Chia nhỏ α .

$$\alpha_{new} = \frac{0.05}{5} = 0.01$$

(Khá bảo thủ, khó đạt ý nghĩa thống kê).

- ② **Holm-Bonferroni:** Mạnh hơn Bonferroni, điều chỉnh theo từng bước (Sequential).

- ③ **FDR (Benjamini-Hochberg):** Kiểm soát tỷ lệ phát hiện sai, thường dùng trong Gen/Y sinh.

Q7: Công cụ thực hiện (Software Implementation) I

Câu hỏi

Có thư viện nào khác ngoài `scipy.stats` để làm Sign Test không? R và SPSS thì sao?

Tổng hợp các công cụ phổ biến:

Python:

- ① `scipy.stats.binomtest` (**Khuyến nghị nhất** - Chuẩn xác).
- ② `statsmodels.stats.descriptivestats.sign_test` (Trả về statistic và p-value).
- ③ `pingouin.sign_test` (Giao diện thân thiện kiểu Pandas).

R Language:

- ① `binom.test()` (Base R).
- ② `SIGN.test()` (Gói BSDA).
- ③ `SignTest()` (Gói DescTools).

Phần mềm thống kê (GUI):

- **SPSS:** Menu: Analyze → Nonparametric Tests → Legacy Dialogs → 2 Related Samples → Tick chọn "Sign".
- **SAS:** Dùng PROC FREQ với tùy chọn binomial.
- **Stata:** Dùng lệnh signtest var1 = var2.

Tip

Dù dùng công cụ nào, hãy đảm bảo bạn đang xem **Exact P-value** chứ không phải Asymptotic (xấp xỉ).

Q8: Kiểm tra Giả định (Assumptions) I

Câu hỏi

Sign Test là phi tham số (nonparametric) nhưng có assumptions nào cần check không?

Các giả định BẮT BUỘC:

① Independence (Quan trọng nhất):

- Các cặp quan sát phải độc lập với nhau.
- Ví dụ: Không thể đo 1 người 10 lần rồi coi là 10 mẫu độc lập.
- Nếu vi phạm → Test hoàn toàn sai.

② Random Sampling:

- Mẫu phải đại diện cho tổng thể (không bị bias).

③ Dữ liệu so sánh được:

- Biến số phải là Continuous (Liên tục) hoặc Ordinal (Thứ bậc).
- Không dùng cho dữ liệu Định danh (Nominal) chưa có thứ tự.

Những điều KHÔNG cần (Ưu điểm):

Q8: Kiểm tra Giả định (Assumptions) II

- ✓ KHÔNG cần phân phối chuẩn (Normality).
- ✓ KHÔNG cần phương sai đồng nhất (Homogeneity of variance).
- ✓ KHÔNG cần mẫu lớn.

Q9: Power Analysis & Sample Size I

Câu hỏi

Làm sao tính sample size cần thiết cho Sign Test? Power là bao nhiêu?

Power phụ thuộc vào 3 yếu tố:

- ① Sample size (n).
- ② Effect size (Tỷ lệ thực sự của dấu + trong quần thể, gọi là P_{real}).
- ③ Mức ý nghĩa α (thường là 0.05).

Ví dụ tính toán:

- Giả thuyết $H_0 : P = 0.5$.
- Thực tế thuộc tốt, $P_{real} = 0.7$ (70% người dùng sẽ đỡ).
- Muốn Power = 0.80 (80% khả năng phát hiện ra).

→ Kết quả tính toán: Cần $n \approx 28$.

Nếu thuộc cực tốt ($P_{real} = 0.8$) → Chỉ cần $n \approx 14$.

Công cụ tính Power:

- **Python:** statsmodels.stats.power (tùy biến class Binomial).
- **R:** Gói pwr hàm power.binom.test().
- **G*Power:** Phần mềm miễn phí phổ biến nhất cho sinh viên/nghiên cứu sinh.

Q9: Power Analysis & Sample Size II

Ghi nhớ

Nếu dữ liệu tuân theo phân phối chuẩn, Sign Test yếu hơn t-test. Nhưng nếu dữ liệu lệch, Sign Test có thể mạnh hơn!

Q10: Các biến thể và mở rộng (Extensions) I

Câu hỏi

Có phiên bản mở rộng nào của Sign Test không? Khi nào dùng?

Các "người anh em" của Sign Test:

① McNemar Test:

- Dùng cho dữ liệu **Nhị phân (Binary)**: 0/1, ĐẬU/Trượt, Bệnh/Khỏi.
- Thiết kế bảng 2x2 contingency.
- Vẫn là so sánh Trước/Sau (Paired).

② Friedman Test:

- Mở rộng cho **nhiều hơn 2 nhóm** (VD: Đo cân nặng tại T1, T2, T3).
- Dùng dữ liệu xếp hạng (Ranked data).
- Được coi là phiên bản phi tham số của ANOVA lặp lại (Repeated Measures ANOVA).

③ Cochran's Q Test:

- Giống Friedman nhưng dùng cho dữ liệu Nhị phân (Binary).
- Mở rộng của McNemar cho nhiều nhóm.

Q10: Các biến thể và mở rộng (Extensions) II

Sơ đồ quyết định (Decision Tree)

- Data liên tục/thứ bậc, 2 nhóm → **Sign Test** (hoặc Wilcoxon).
- Data nhị phân, 2 nhóm → **McNemar Test**.
- Data liên tục/thứ bậc, >2 nhóm → **Friedman Test**.
- Data nhị phân, >2 nhóm → **Cochran's Q**.

Liên hệ:

Trương Tuấn Nghĩa: nghia.tt251196m@sis.hust.edu.vn

GitHub: <https://github.com/yourusername/sign-test-demo>