



**TRƯỜNG ĐẠI HỌC
BÁCH KHOA HÀ NỘI**
HANOI UNIVERSITY
OF SCIENCE AND TECHNOLOGY



**VIỆN TOÁN ỨNG
DỤNG VÀ TIN HỌC**
SCHOOL OF APPLIED
MATHEMATICS AND INFORMATICS

Kiểm định Dấu (Sign Test)

Từ Lịch sử đến Thực hành với Python

Đào Duy Tân – 20251334M

Nguyễn Đức Dũng – 20251240M

Trương Tuấn Nghĩa – 20251196M

Hồ Hải Anh - 20252102M

Đinh Lâm Nghị - 20241116M

Tháng 12, 2025

ONE LOVE. ONE FUTURE

Lịch sử & Động lực

- Sign Test trong Lịch sử

- Tại sao Sign Test quan trọng?

Lý thuyết Cốt lõi

- Sign Test là gì?

- Nguyên lý hoạt động

- Giả thuyết & Công thức

Ví dụ Thực tế

- Ví dụ Y học

- Ví dụ Marketing

- Ví dụ Tâm lý học

Demo Python/Jupyter

- Setup Demo

- Tính hiệu số và dấu

- Thực hiện kiểm định

- Trực quan hóa

Lịch sử & Động lực

Lý thuyết Cốt lõi

Ví dụ Thực tế

Demo Python/Jupyter

FAQs

Kết luận

Câu hỏi & Trả lời

John Arbuthnot (1667-1735)

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

Imgarbuthnot_timeline.png0.65

Ý nghĩa

Ứng dụng sớm nhất của thống kê!

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

Imgplaceholder.png0.5

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

Lịch sử & Động lực

Sign Test trong Lịch sử

Tại sao Sign Test quan trọng?

Lý thuyết Cốt lõi

Ví dụ Thực tế

Demo Python/Jupyter

FAQs

Kết luận

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:

Imgplaceholder.png0.5

- ◇ **Marketing:** A/B testing
- ◇ **Y học:** Thử nghiệm lâm sàng
- ◇ **UX:** Nghiên cứu người dùng
- ◇ **QC:** Kiểm soát chất lượng

Mẫu chốt

300+ năm tuổi nhưng vẫn hữu ích!

Lịch sử & Động lực

Lý thuyết Cốt lõi

Sign Test là gì?

Nguyên lý hoạt động

Giả thuyết & Công thức

Ví dụ Thực tế

Demo Python/Jupyter

FAQs

Kết luận

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

Imgplaceholder.png0.5

Ý tưởng

Nếu H_0 đúng:

Số dấu + và - **bằng nhau**

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

Lịch sử & Động lực

Lý thuyết Cốt lõi

Sign Test là gì?

Nguyên lý hoạt động

Giả thuyết & Công thức

Ví dụ Thực tế

Demo Python/Jupyter

FAQs

Kết luận

Imgplaceholder.png0.5

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

- $S^+ \sim \text{Binom}(n, 0.5)$
- $n = \text{số sau loại ties}$

4. P-value

- So với $\alpha = 0.05$
- Quyết định bác bỏ H_0

Lịch sử & Động lực

Lý thuyết Cốt lõi

Sign Test là gì?

Nguyên lý hoạt động

Giả thuyết & Công thức

Ví dụ Thực tế

Demo Python/Jupyter

FAQs

Kết luận

Cặp giả thuyết:**Hypotheses**

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

hoặc

- ◇ $H_1: \text{Median}(D) > 0$ (1-tailed)
- ◇ $H_1: \text{Median}(D) < 0$ (1-tailed)

Mức ý nghĩa:

- ◇ $\alpha = 0.05$ hoặc 0.01

Imgplaceholder.png0.5

Thống kê:**Test Statistic**

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

- ◇ $S^+ = \text{số dấu dương}$
- ◇ $n = \text{số quan sát}$

P-value (2-tailed):

$$p = 2 \times \min(P(S^+ \leq s), P(S^+ \geq s))$$

Quyết định:

- ◇ $p < \alpha \rightarrow \text{Bác bỏ } H_0$
- ◇ $p \geq \alpha \rightarrow \text{Giữ } H_0$

Lịch sử & Động lực

Lý thuyết Cốt lõi

Ví dụ Thực tế

Ví dụ Y học

Ví dụ Marketing

Ví dụ Tâm lý học

Demo Python/Jupyter

FAQs

Kết luận

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**

Lịch sử & Động lực

Lý thuyết Cốt lõi

Ví dụ Thực tế

Ví dụ Y học

Ví dụ Marketing

Ví dụ Tâm lý học


Demo Python/Jupyter

FAQs

Kết luận

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?

So sánh 2 giao diện A và B bên cạnh nhau, có bar chart thể hiện 11/15 thích B hơn  *ab_testing_comparison.png* 0.95

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!)

Lịch sử & Động lực

Lý thuyết Cốt lõi

Ví dụ Thực tế

Ví dụ Y học

Ví dụ Marketing

Ví dụ Tâm lý học

Demo Python/Jupyter

FAQs

Kết luận

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)

Bar chart thể hiện số người thích xanh vs đỏ, với màu sắc tương ứng `color_preference.png`

Phân tích Sign Test:

- ◇ H_0 : Không có preference ($p = 0.5$)
- ◇ H_1 : Có preference ($p \neq 0.5$)
- ◇ $S^+ = 14$, $n = 20$
- ◇ Under H_0 : $S^+ \sim \text{Binomial}(20, 0.5)$

Kết luận:

- ◇ p-value = **0.115** (two-tailed)
- ◇ $0.115 > 0.05$
- ◇ \rightarrow **Giữ H_0**
- ◇ Không có preference rõ ràng

Lịch sử & Động lực

Lý thuyết Cốt lõi

Ví dụ Thực tế

Demo Python/Jupyter

Setup Demo

Tính hiệu số và dấu

Thực hiện kiểm định

Trực quan hóa

FAQs

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?

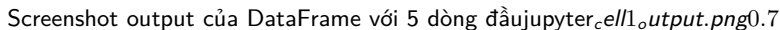
Chuẩn bị Jupyter Notebook:

```
[language=Python, caption=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())
```

Screenshot output của DataFrame với 5 dòng đầu  `jupyter_cell_output.png`

Lịch sử & Động lực

Lý thuyết Cốt lõi

Ví dụ Thực tế

Demo Python/Jupyter

Setup Demo

Tính hiệu số và dấu

Thực hiện kiểm định

Trực quan hóa

FAQs

[language=Python, caption=Cell 2: Calculate differences and signs] Buoc 1: Tinh hieu (truoc - sau)

```
df['Hieukg''] = df['Truockg''] - df['Saukg'']
```

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

```
df['Dau'] = df['Hieukg''].apply(assignsign)
```

Buoc 3: Loc bo hieu = 0 `dffiltered = df[df['Hieukg''] != 0]`

Hien thi ket qua `print(dffiltered[['ID', 'Truockg'', 'Saukg'', 'Hieukg'', 'Dau']])`

Screenshot bảng kết quả với cột Dấu, highlight dấu + màu xanh, dấu - màu

`đỏjupyteroutputtable.png0.75`

Lịch sử & Động lực

Lý thuyết Cốt lõi

Ví dụ Thực tế

Demo Python/Jupyter

Setup Demo

Tính hiệu số và dấu

Thực hiện kiểm định

Trực quan hóa

FAQs

```
[language=Python, caption=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 H0 (p=result.pvalue:.4f < alpha)")
print("=> CO BANG CHUNG: Thuoc co hieu qua!") else: print(f"GIU H0 (p=result.pvalue:.4f >=
alpha)") print("=> Chua du bang chung")
```

Screenshot output text với kết quả: Số dấu +: 9/11, P-value: 0.0327, BẮC BỎ

H0jupyter_output_tats.png0.6

Lịch sử & Động lực

Lý thuyết Cốt lõi

Ví dụ Thực tế

Demo Python/Jupyter

Setup Demo

Tính hiệu số và dấu

Thực hiện kiểm định

Trực quan hóa

FAQs

[language=Python, caption=Cell 4: Visualize results,

basicstyle=] 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, alpha =

0.6)ax1.plot([75, 100], [75, 100], 'r--', label =

Nochange')ax1.set_xlabel('CannangTruoc(kg)')ax1.set_ylabel('CannangSau(kg)')ax1.legend()

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, alpha = 0.6)ax2.axhline(0, color = 'black', lw =

2)ax2.set_xlabel('ID')ax2.set_ylabel('Hieuso(kg)')

plt.tight_layout()plt.show()

2 subplots: Scatter plot Trước vs Sau với đường chéo, Bar chart Hiệu số với màu xanh/đỏjupyter_{lots.png}0.95

Insight từ plots

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

Lịch sử & Động lực

Lý thuyết Cốt lõi

Ví dụ Thực tế

Demo Python/Jupyter

FAQs

Câu hỏi thường gặp

Kết luận

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

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

Lịch sử & Động lực

Lý thuyết Cốt lõi

Ví dụ Thực tế

Demo Python/Jupyter

FAQs

Kết luận

Tóm tắt

Tài liệu tham khảo

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

Lịch sử & Động lực

Lý thuyết Cốt lõi

Ví dụ Thực tế

Demo Python/Jupyter

FAQs

Kết luận

Tóm tắt

Tài liệu tham khảo

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_code_github.png`

Lịch sử & Động lực

Lý thuyết Cốt lõi

Ví dụ Thực tế

Demo Python/Jupyter

FAQs

Kết luận

Câu hỏi & Trả lời

Câu hỏi

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

Trả lời:

Tiêu chí	Sign Test	Wilcoxon
Thông tin sử dụng	Chỉ dấu (+/-)	Dấu + Độ lớn
Power	Thấp hơn	Cao hơn
Assumptions	Ít nhất	Thêm: symmetric distribution
Dễ hiểu	Rất dễ	Phức tạp hơn
Khi nào dùng	Outliers cực đoan	Data gần symmetric

Khuyến nghị:

- ◇ Sign Test: Khi data rất noisy, nhiều outliers, chỉ quan tâm hướng thay đổi
- ◇ Wilcoxon: Khi data gần symmetric và muốn power cao hơn

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 sample size hiệu dụng
- ◇ Giảm statistical power
- ◇ Nếu $> 20\%$ ties \rightarrow cần cẩn thận

Phương pháp xử lý:

1. **Standard:** Loại bỏ (khuyến nghị)
2. Chia đều vào + và -
3. Random assign

Ví dụ:

$n = 20$, có 5 ties (25%)

- ◇ **Cách 1:** Loại $\rightarrow n = 15$
- ◇ **Cách 2:** 2.5 vào +, 2.5 vào -
- ◇ **Cách 3:** Random split

Lưu ý

Cách 1 (loại bỏ) là standard practice và được khuyến nghị trong hầu hết trường hợp

Câu hỏi

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

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

Ví dụ - Khảo sát hài lòng:

Thiết kế:

- ◇ Survey trước và sau cải tiến dịch vụ
- ◇ Scale: 1 (rất không hài lòng) \rightarrow 5 (rất hài lòng)
- ◇ So sánh: Trước vs Sau

Phân tích:

- ◇ Sau $>$ Trước \rightarrow dấu +
- ◇ Sau $<$ Trước \rightarrow dấu -
- ◇ Sau = Trước \rightarrow loại

Tại sao phù hợp?

1. Likert scale = ordinal, không phải interval
2. Khoảng cách giữa 1-2 \neq 4-5
3. Sign Test chỉ cần thứ tự, không cần distance

Ưu điểm

Không cần giả định data là interval/ratio!

Câu hỏi

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

Trả lời:

Vấn đề với n nhỏ:

- ◇ Binomial distribution rời rạc
- ◇ P-values "nhảy vọt"
- ◇ Khó đạt significance level chính xác

Ví dụ với $n=6$:

- ◇ $S^+ = 5$: $p = 0.219$
- ◇ $S^+ = 6$: $p = 0.031$
- ◇ Không có p-value nào trong khoảng 0.031-0.219!

Giải pháp:

1. **Exact test:** Dùng binomial exact (luôn khuyến nghị)
2. **Report exact p:** Không làm tròn
3. **Consider power:** n nhỏ = low power
4. **Collect more data:** Nếu có thể!

Khuyến nghị

Sign Test vẫn VALID với n nhỏ, nhưng power thấp. Ít nhất cần $n \geq 5$ sau khi loại ties.

Câu hỏi

Khi viết báo cáo hoặc paper, cần báo cáo những gì về Sign Test?

Thông tin cần thiết:

1. **Sample size:**

- Original n
- Số ties loại bỏ
- Effective n (sau loại ties)

2. **Test statistic:**

- Số dấu + và -
- Tỷ lệ (optional)

3. **P-value:**

- Exact p-value
- One-tailed hay two-tailed

4. **Effect size:**

- Proportion of positive signs
- Median difference

Ví dụ câu viết tốt:

"A Sign test was conducted to evaluate whether the new medication reduced pain scores. Of 12 participants, 1 showed no change and was excluded, leaving $n=11$ for analysis. Nine participants (81.8%) showed improvement (positive signs), compared to 2 (18.2%) who showed worsening. The one-tailed Sign test indicated a significant improvement ($p = 0.033$), with a median pain reduction of 2 points on the VAS scale."

Câu hỏi

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

Trả lời: CẦN! Đây là vấn đề multiple testing.

Vấn đề:

- ◇ Làm k tests độc lập
- ◇ $\alpha = 0.05$ cho mỗi test
- ◇ Family-wise error rate > 0.05

Ví dụ với 5 tests:

$$P(\text{ít nhất 1 FP}) = 1 - 0.95^5 = 0.226$$

→ 22.6% chance có false positive!

Giải pháp correction:

1. **Bonferroni:** $\alpha^* = \alpha/k$
 - Conservative
 - Dễ áp dụng
2. **Holm:** Sequential Bonferroni
 - Less conservative
 - Recommended
3. **FDR (Benjamini-Hochberg):**
 - Controls false discovery rate
 - Modern approach

Câu hỏi

Có thư viện nào khác ngoài `scipy.stats` để làm Sign Test không? R có sẵn không?

Trả lời - Nhiều options:

Python:

1. `scipy.stats.binomtest` (**khuyến nghị**)
2. `statsmodels.stats.descriptivestats.sign_test`
3. `pingouin.sign_test` (user-friendly)

SPSS / SAS / Stata:

- ◇ **SPSS:** Analyze → Nonparametric → 2 Related Samples → Sign Test
- ◇ **SAS:** PROC FREQ with binomial option
- ◇ **Stata:** `signtest`

R:

1. `binom.test()` (base R)
2. `SIGN.test()` (BSDA package)
3. `SignTest()` (DescTools)

Tip

Hầu hết packages đều dùng exact binomial test - kết quả giống nhau!

Câu hỏi

Sign Test là nonparametric nhưng có assumptions nào cần check không?

Assumptions của Sign Test:

1. **Independence:** (QUAN TRỌNG NHẤT)
 - Mỗi observation độc lập
 - Paired design: mỗi cặp độc lập với nhau
 - Violation → test invalid
2. **Random sampling:**
 - Sample representative cho population
 - Không có selection bias
3. **Continuous or ordinal variable:**
 - Cần có thể so sánh được ($>$, $<$, $=$)
 - Nominal data → không phù hợp

Điều KHÔNG cần:

- ◇ ✓ Normality
- ◇ ✓ Homogeneity of variance
- ◇ ✓ Interval scale
- ◇ ✓ Large sample size

Câu hỏi

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

Power của Sign Test:

◇ Phụ thuộc vào:

1. Sample size (n)
2. Effect size (true proportion of + signs)
3. Significance level (α)

Ví dụ Power calculation:

Giả sử:

- ◇ True $p = 0.7$ (70% có dấu +)
- ◇ $\alpha = 0.05$ (two-tailed)
- ◇ Muốn power = 0.80

→ Cần $n = 28$

Nếu true $p = 0.8 \rightarrow$ Cần $n = 14$

Tools tính sample size:

Python:

- ◇ `statsmodels.stats.power`
- ◇ Custom code với binomial

R:

- ◇ `pwr` package
- ◇ `power.binom.test()`

Online:

- ◇ G*Power (free software)

Lưu ý

Sign Test có power thấp hơn t-test 5-10% khi data normal, nhưng robust hơn nhiều!

Câu hỏi

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

Variants của Sign Test:

1. McNemar Test:

- Binary outcome (success/fail)
- 2×2 contingency table
- Paired data

2. Cochran's Q:

- Extension cho >2 time points
- Binary repeated measures

3. Friedman Test:

- >2 related groups
- Ranked data
- Nonparametric ANOVA

Decision tree

- ◇ Continuous/ordinal, 2 groups → Sign Test
- ◇ Binary, 2 groups → McNemar
- ◇ Binary, >2 groups → Cochran's Q
- ◇ Continuous, >2 groups → Friedman

CÂU HỎI?

Cảm ơn các thầy cô và các bạn đã lắng nghe!

Liên hệ:

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

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