

# Kiểm định Dấu (Sign Test)

Từ Lịch sử đến Thực hành với 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

Môn: Thống kê Ứng dụng  
Trường Đại học Bách Khoa Hà Nội

Tháng 12, 2025

# 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

## Mục lục II

### 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) I

## John Arbuthnot (1667-1735)

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

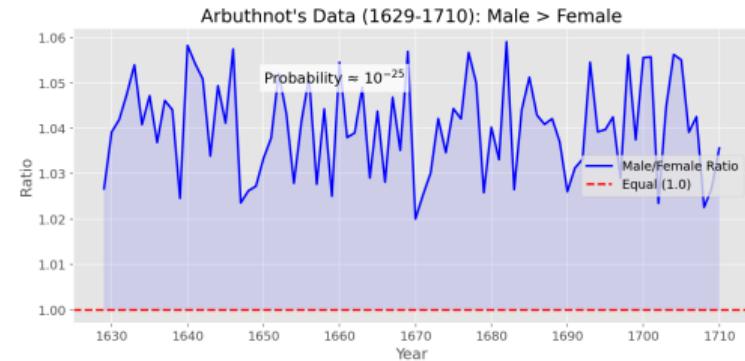
## Phát hiện đáng chú ý:

- 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

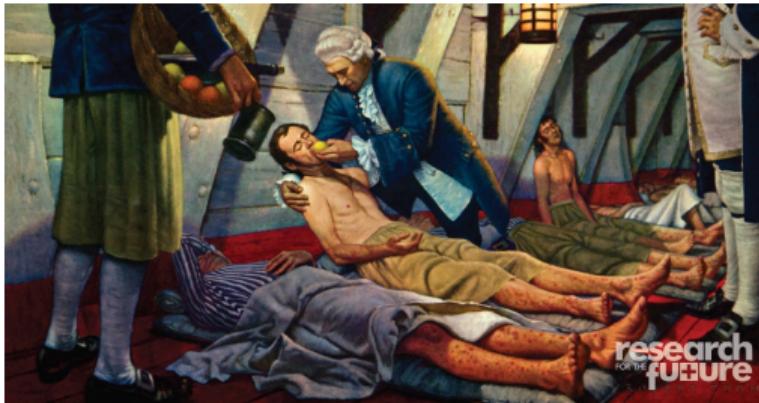
Nguồn: Arbuthnot, J. (1710). "An argument for Divine Providence..."



## 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

# Mục lục

## ① Lịch sử & Động lực

Sign Test trong Lịch sử

Tại sao Sign Test quan trọng?

## ② Lý thuyết Cốt lõi

## ③ Ví dụ Thực tế

## ④ Demo Python/Jupyter

## ⑤ FAQs

## ⑥ Kết luận

## ⑦ Câu hỏi & Trả lời

# 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

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

Sign Test là gì?

Nguyên lý hoạt động

Giả thuyết & Công thức

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)

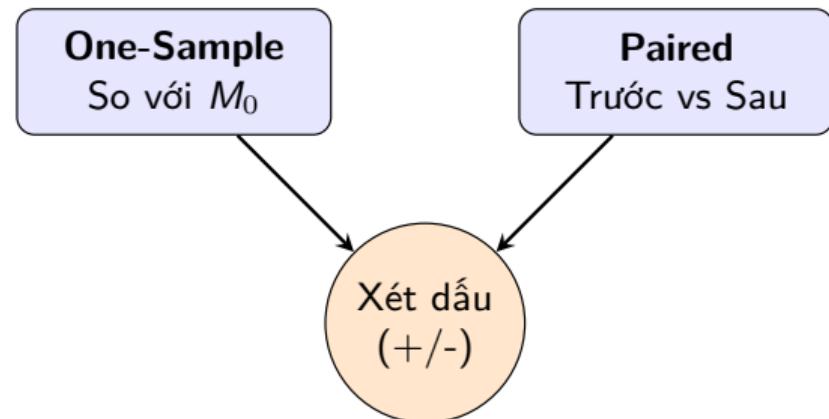
Kiểm định phi tham số sử dụng **dấu (+/-)** để kiểm tra median hoặc so sánh mẫu phụ thuộc.

### 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



**Cốt lõi:** Quy về nhị thức

# Mục lục

## ① Lịch sử & Động lực

## ② Lý thuyết Cốt lõi

Sign Test là gì?

Nguyên lý hoạt động

Giả thuyết & Công thức

## ③ Ví dụ Thực tế

## ④ Demo Python/Jupyter

## ⑤ FAQs

## ⑥ Kết luận

## ⑦ Câu hỏi & Trả lời

## 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

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

### ③ Thống kê

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

### ④ P-value

- So với  $\alpha = 0.05$

# Mục lục

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

Sign Test là gì?

Nguyên lý hoạt động

Giả thuyết & Công thức

③ Ví dụ Thực tế

④ Demo Python/Jupyter

⑤ FAQs

⑥ Kết luận

⑦ Câu hỏi & Trả lời

# 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

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

Ví dụ Y học

Ví dụ Marketing

Ví dụ Tâm lý học

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:

- $p\text{-value} = \textbf{0.039}$
- $\textbf{0.039} < 0.05$

# Mục lục

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

Ví dụ Y học

Ví dụ Marketing

Ví dụ Tâm lý học

④ Demo Python/Jupyter

⑤ FAQs

⑥ Kết luận

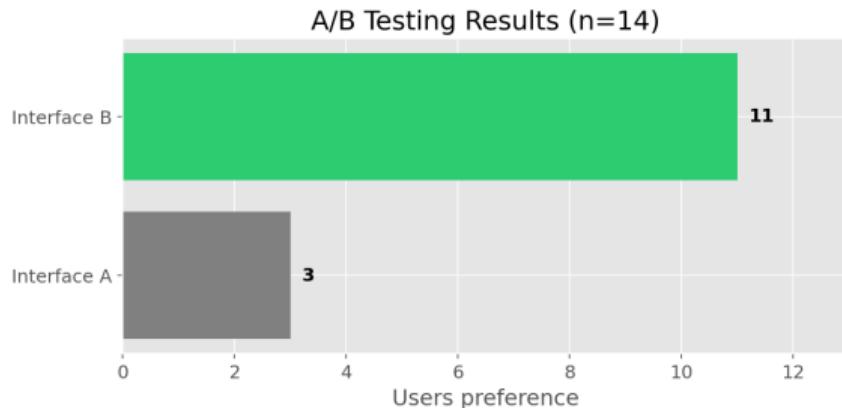
⑦ Câu hỏi & Trả lời

## 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)
- $p\text{-value} = \mathbf{0.059}$  (two-tailed)

## Quyết định:

- Với  $\alpha = 0.05$ : **Chưa đủ bằng chứng** (biên giới!)
- Với  $\alpha = 0.10$ : **Có bằng chứng** B tốt hơn

# Mục lục

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

Ví dụ Y học

Ví dụ Marketing

Ví dụ Tâm lý học

④ Demo Python/Jupyter

⑤ FAQs

⑥ Kết luận

⑦ Câu hỏi & Trả lời

## 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:

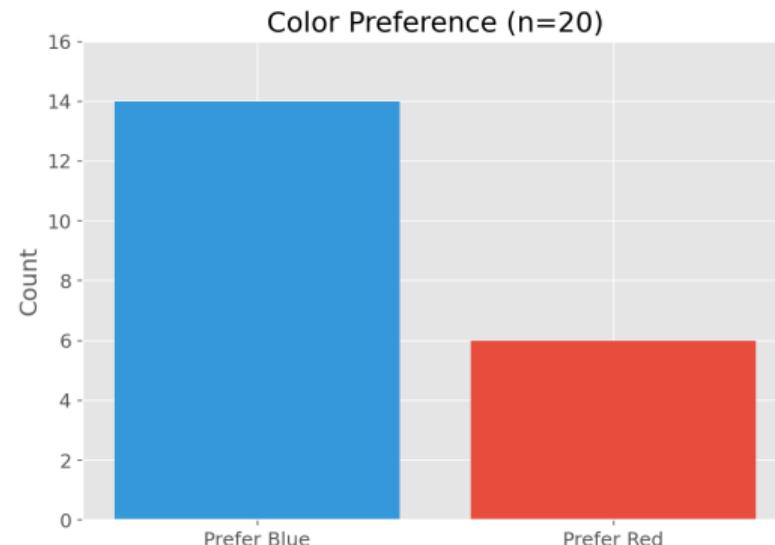
- Người ta có thích màu xanh hơn màu đỏ không?
- (So với ngẫu nhiên 50-50)

## 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 ( $p = 0.5$ )

# Mục lục

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

④ Demo Python/Jupyter

Setup Demo

Tính hiệu số và dấu

Thực hiện kiểm định

Trực quan hóa

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)
```

# Mục lục

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

④ Demo Python/Jupyter

Setup Demo

Tính hiệu số và dấu

Thực hiện kiểm định

Trực quan hóa

⑤ FAQs

⑥ Kết luận

⑦ Câu hỏi & Trả lời

# 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	

# Mục lục

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

④ Demo Python/Jupyter

Setup Demo

Tính hiệu số và dấu

Thực hiện kiểm định

Trực quan hóa

⑤ FAQs

⑥ Kết luận

⑦ Câu hỏi & Trả lời

## 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

# Mục lục

① Lịch sử & Động lực

② Lý thuyết Cốt lõi

③ Ví dụ Thực tế

④ Demo Python/Jupyter

Setup Demo

Tính hiệu số và dấu

Thực hiện kiểm định

Trực quan hóa

⑤ FAQs

⑥ Kết luận

⑦ Câu hỏi & Trả lời

# Demo Step 3: Visualization

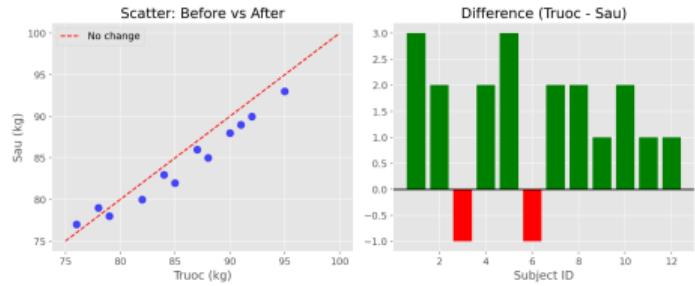
Listing 4: Cell 4: Visualize results

```
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

Câu hỏi thường gặp

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 gắp (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
- **Khuyên 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

Tóm tắt

Tài liệu tham khảo

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

# 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

Tóm tắt

Tài liệu tham khảo

⑦ Câu hỏi & Trả lời

# 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:

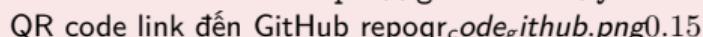
- SciPy Documentation: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.binomtest.html>
- Statistics How To: <https://www.statisticshowto.com/sign-test/>
- Penn State STAT 415: <https://online.stat.psu.edu/stat415/>

## 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\_code\_github.png0.15

# 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

Câu hỏi dự kiến

# Câu hỏi 1: So sánh với các test khác I

## 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 2: Xử lý ties I

### 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 3: Áp dụng cho data thứ bậc I

## 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) → 5 (rất hài lòng)
- So sánh: Trước vs Sau

Phân tích:

- Sau > Trước → dấu +
- Sau < Trước → dấu -
- Sau = Trước → loại

Tại sao phù hợp?

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

Ưu điểm

Không cần giả định data là interval/ratio!

# Câu hỏi 4: Sample size nhỏ I

## 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:

- ❶ **Exact test:** Dùng binomial exact (luôn khuyến nghị)
- ❷ **Report exact p:** Không làm tròn
- ❸ **Consider power:**  $n$  nhỏ = low power
- ❹ **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 5: Báo cáo kết quả I

## 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

# Câu hỏi 5: Báo cáo kết quả II

- 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 6: Multiple comparisons I

## 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 7: Computational implementation I

## 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 8: Assumptions checking I

## 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 ( $>$ ,  $<$ ,  $=$ )

# Câu hỏi 8: Assumptions checking II

- Nominal data → không phù hợp

**Điều KHÔNG cần:**

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

# Câu hỏi 9: Power analysis I

## 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 ( $\alpha$ )

# Câu hỏi 9: Power analysis II

## 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  $p = 0.8$  → 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 10: Extensions và variants I

## 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)
- $2 \times 2$  contingency table
- Paired data

#### ② Cochran's Q:

- Extension cho  $>2$  time points
- Binary repeated measures

#### ③ Friedman Test:

- $>2$  related groups
- Ranked data

# Câu hỏi 10: Extensions và variants II

- 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 Thầy 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>