

# Hypothesis Testing

October 23, 2023

```
[ ]: from IPython.display import Image
```

## 1 Hypothesis Tests for Finance

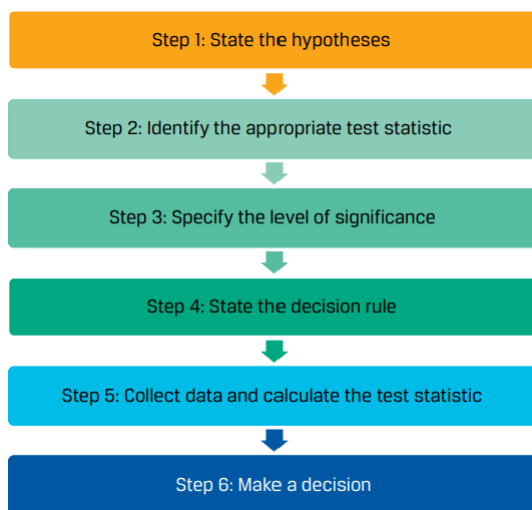
Chúng ta thực hiện kiểm định giả thuyết để đưa ra quyết định thông qua dữ liệu. Cụ thể hơn, kiểm định giả thuyết là một phần của suy luận thống kê, trong đó phân tích viên đưa ra kết luận về một tổng thể dựa trên một mẫu

Các khái niệm và công cụ kiểm định giả thuyết cung cấp một phương tiện khách quan để đánh giá xem liệu có bằng chứng có giá trị nào hỗ trợ cho giả thuyết hay không. Sau khi áp dụng các kiểm định thống kê, chúng sẽ có ý tưởng rõ ràng hơn cho xác suất một giả thuyết nào đó đúng hay không, mặc dù kết luận của chúng ta luôn tồn tại sự thiếu chắc chắn

Quy trình kiểm định giả thuyết được mô tả trong Hình 1

```
[ ]: Image(filename = "Pictures/01.png")
```

```
[ ]:
```



## 2 Tests of Return and Risk in Finance

Kiểm định giả thuyết về lợi tức và rủi ro là một trong những kiểm định thường gặp nhất trong thế giới tài chính. Phân phối mẫu của trung bình, trong trường hợp chưa biết độ lệch chuẩn, là phân

phối  $t$ ; và sẽ là phân phối  $z$  nếu đã biết trước sai số chuẩn

## 2.1 Test of a Single Mean

Kiểm định của một trung bình đơn thường được sử dụng để đưa ra kết luận về đặc điểm sinh lợi cũng như rủi ro của khoản đầu tư (ví dụ như đưa ra kết luận về việc liệu tỷ suất sinh lợi có ở mức dưới chuẩn hay không)

Giả thuyết không của kiểm định này có thể là: (1)  $\mu = \mu_0$ ; (2)  $\mu \leq \mu_0$ ; (3)  $\mu \geq \mu_0$

Kiểm định được sử dụng trong trường hợp này thường là kiểm định  $t$ . Giá trị thống kê  $t$  được tính bởi công thức:

$$t_{n-1} = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

## 2.2 Test of Single Variance

Kiểm định của một phương sai đơn thường được sử dụng để đưa ra kết luận về đặc điểm rủi ro của khoản đầu tư (vốn được mô tả thông qua độ lệch chuẩn)

Giả thuyết không của kiểm định này có thể là: (1)  $\sigma^2 = \sigma_0^2$ ; (2)  $\sigma^2 \leq \sigma_0^2$ ; (3)  $\sigma^2 \geq \sigma_0^2$

Kiểm định được sử dụng trong trường hợp này thường là kiểm định  $\chi^2$ . Giá trị thống kê  $\chi^2$  được tính bởi công thức:

$$\chi^2(n-1) = \frac{(n-1)s^2}{\sigma_0^2}$$

## 2.3 Test Concerning Differences between Means with Independent Samples

Trong nhiều trường hợp, chúng ta cần biết liệu giá trị trung bình của hai tổng thể có khác nhau hay không. Khi có cơ sở hợp lý để tin rằng các mẫu từ các tổng thể có phân bố tiệm cận chuẩn và các mẫu cũng độc lập với nhau, chúng ta sử dụng kiểm định sự chênh lệch giá trị trung bình.

Giả thuyết không của kiểm định này là  $\mu_1 - \mu_2 = 0$

Trong trường hợp chưa biết trước độ lệch chuẩn ở cả hai tổng thể, có hai kịch bản xác định giá trị kiểm định  $t$ :

1. Nếu kích thước của cả hai mẫu đủ lớn

$$t_{n_1+n_2-2} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

2. Nếu kích thước mẫu nhỏ, phân tích viên có thể sử dụng phương sai gộp (pooled variance), với điều kiện phương sai hai tổng thể như nhau

$$t_{n_1+n_2-2} = \frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$$

trong đó

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

## 2.4 Test Concerning Differences between Means with Dependent Samples

Khi chúng ta cần kiểm định về sự chênh lệch giữa hai trung bình từ hai mẫu mà chúng ta tin rằng có sự phụ thuộc, chúng ta sử dụng kiểm định trung bình sự chênh lệch (kiểm định so sánh theo cặp)

Kiểm định so sánh theo cặp mạnh hơn so với kiểm định chênh lệch trung bình, vì bằng cách sử dụng yếu tố chung (chẳng hạn như cùng thời kỳ hoặc công ty), phân tích viên hoàn toàn loại bỏ sự khác biệt giữa các mẫu gây ra bởi lý do nào đó khác với những gì phân tích viên đang kiểm tra

Giả thuyết không của kiểm định này có thể là: (1)  $d = d_0$ ; (2)  $d \leq d_0$ ; (3)  $d \geq d_0$

Phân phối của trung bình chênh lệch là phân phối  $t$ , với giá trị kiểm định được tính bằng công thức:

$$t_d(n-1) = \frac{\bar{d} - d_0}{s_d/\sqrt{n}}$$

## 3 Parametric versus Nonparametric Tests

**Kiểm định tham số (parametric test)** là một phương pháp thống kê để kiểm tra một giả thuyết về một hoặc nhiều tham số của một quần thể, như trung bình, phương sai, tỷ lệ, hệ số tương quan, hệ số hồi quy... Kiểm định tham số yêu cầu các giả định về phân phối của dữ liệu, thường là phân phối chuẩn. Kiểm định tham số có thể cho kết quả chính xác và tin cậy hơn khi các giả định được thoả mãn.

**Kiểm định phi tham số (nonparametric test)** là một phương pháp kiểm tra giả thuyết thống kê mà không cần giả định về phân phối của dữ liệu

Kiểm định phi tham số thường được sử dụng trong 4 trường hợp: (1) Dữ liệu không đáp ứng giả định về phân phối; (2) Tồn tại quan sát ngoại lai; (3) Dữ liệu sử dụng thang đo thứ tự; (4) Các giả thuyết liên quan không liên quan đến một tham số tổng thể (ví dụ như kiểm định mẫu có ngẫu nhiên hay không)

Một số kiểm định phi tham số thay thế cho kiểm định tham số được trình bày trong Hình 2

```
[ ]: Image(filename = "Pictures/02.png")
```

```
[ ]:
```

	Parametric	Nonparametric
Tests concerning a single mean	$t$ -distributed test z-distributed test	Wilcoxon signed-rank test
Tests concerning differences between means	$t$ -distributed test	Mann–Whitney $U$ test (Wilcoxon rank sum test)
Tests concerning mean differences (paired comparisons tests)	$t$ -distributed test	Wilcoxon signed-rank test Sign test