Xác suất thống kê - Kiểm định thống kê

Ngày 19 tháng 12 năm 2024

Ví dụ dẫn nhập

Hai người nhận hai đồng xu. Biết rằng một đồng xu là cân đối, một đồng xu không cân đối.

Ví dụ dẫn nhập

Hai người nhận hai đồng xu. Biết rằng một đồng xu là cân đối, một đồng xu không cân đối.
Hai người tung đồng xu này 100 lần.

Ví dụ dẫn nhập

Hai người nhận hai đồng xu. Biết rằng một đồng xu là cân đối, một đồng xu không cân đối.
Hai người tung đồng xu này 100 lần.

- Người thứ nhất thấy đồng xu nhận mặt ngửa 49 lần và 51 lần mặt úp.

Ví dụ dẫn nhập

Ví dụ dẫn nhập

Hai người nhận hai đồng xu. Biết rằng một đồng xu là cân đối, một đồng xu không cân đối.
Hai người tung đồng xu này 100 lần.

- Người thứ nhất thấy đồng xu nhận mặt ngửa 49 lần và 51 lần mặt úp.
- Người thứ hai thấy đồng xu nhận mặt ngửa 53 lần và 47 lần mặt úp.

Hỏi đồng xu của người nào là cân đối ?

Ví dụ dẫn nhập

Nhận xét

Cùng câu hỏi, tuy nhiên nếu biết

Ví dụ dẫn nhập

Nhận xét

Cùng câu hỏi, tuy nhiên nếu biết

- Người thứ nhất thấy đồng xu nhận mặt ngửa 49 lần và 51 lần mặt úp.
- Người thứ hai thấy đồng xu nhận mặt ngửa 40 lần và 60 lần mặt úp.

Ví dụ dẫn nhập

Nhận xét

Cùng câu hỏi, tuy nhiên nếu biết

- Người thứ nhất thấy đồng xu nhận mặt ngửa 49 lần và 51 lần mặt úp.
- Người thứ hai thấy đồng xu nhận mặt ngửa 40 lần và 60 lần mặt úp.

Trong trường hợp này, rõ ràng người thứ hai có đồng xu không cân đối.

Giả thuyết thống kê

Giả thuyết thống kê là một nhận định hay một phỏng đoán liên quan đến một hay nhiều tổng thể.

Xác suất thống kê - Kiểm định thống kê

Dẫn nhập

Kiểm định giả thuyết thống kê

Kiểm định kì vọng một mẫu

Kiểm định tỉ lệ một mẫu

So sánh trung bình hai mẫu

Dữ liệu cặp

So sánh tỉ lệ hai mẫu

Giả thuyết thống kê

Giả thuyết thống kê là một nhận định hay một phỏng đoán liên quan đến một hay nhiều tổng thể.

Giả thuyết không và đối thuyết

Trong lý thuyết kiểm định, ta quan tâm đến hai loại giả thuyết thống kê sau:

- Giả thuyết không: Là nhận định, hay phỏng đoán mà ta muốn kiểm định. Đây là giả thuyết mà ta ***mong muốn*** đúng.
- Đối thuyết: Là nhận định, hay phỏng đoán ngược lại với giả thuyết không.

Ta kí hiệu H_0 cho giả thuyết không, và H_1 cho đối thuyết.

Ví dụ

Trong ví dụ dẫn nhập, một người tung đồng xu 100 lần và được 49 lần mặt ngửa. Vậy đồng xu có cân đối hay không ?

Ví dụ

Trong ví dụ dẫn nhập, một người tung đồng xu 100 lần và được 49 lần mặt ngửa. Vậy đồng xu có cân đối hay không ?
Trong trường hợp này, ta có cặp giả thuyết không - đối thuyết:

$$\begin{cases} H_0 : p = 1/2 \\ H_1 : p \neq 1/2 \end{cases}$$

trong đó p là xác suất đồng xu nhận mặt ngửa.

Kết luận của bài toán kiểm định

Khi kết thúc kiểm định, ta cần quyết định giả thuyết H_0 hay H_1 nào có nhiều "bằng chứng" để chấp nhận hơn. Như vậy, ta có hai loại kết luận sau:

- *Bác bỏ* H_0: Dữ liệu cho thấy đối thuyết H_1 phù hợp hơn H_0.
- *Không đủ cơ sở để bác bỏ* H_0: Dữ liệu cho thấy giả thuyết không H_0 phù hợp hơn H_1.

Dẫn nhập về kiểm định giả thuyết thống kê

Kết luận của bài toán kiểm định

Khi kết thúc kiểm định, ta cần quyết định giả thuyết H_0 hay H_1 nào có nhiều "bằng chứng" để chấp nhận hơn. Như vậy, ta có hai loại kết luận sau:

- *Bác bỏ* H_0: Dữ liệu cho thấy đối thuyết H_1 phù hợp hơn H_0.
- *Không đủ cơ sở để bác bỏ* H_0: Dữ liệu cho thấy giả thuyết không H_0 phù hợp hơn H_1.

Thống kê kiểm định

Một kiểm định sẽ luôn đi kèm với một thống kê để thực hiện việc kết luận của bài toán kiểm định. Thống kê này sẽ được gọi là ***thống kê kiểm định***.

Giá trị tới hạn

Cho kiểm định giả thuyết thống kê với giả thuyết H_0 và đối thuyết H_1.

• Nếu thống kê kiểm định T bác bỏ H_0 nếu $T > c$ (hay $T < c$, $|T| > c$) thì giá trị hằng số c được gọi là ***giá trị tới hạn***.

• Tập hợp các giá trị của mẫu $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ thỏa điều kiện bác bỏ H_0 được gọi là ***miền bác bỏ*** của kiểm định, và thường kí hiệu là R.

Ví dụ

Trong bài toán tung đồng xu, ta có

$$\begin{cases} H_0 : p = 1/2 \\ H_1 : p \neq 1/2 \end{cases}$$

Một người tung đồng xu 100 lần và ghi lại kết quả.

Ví dụ

Trong bài toán tung đồng xu, ta có

$$\begin{cases} H_0 : p = 1/2 \\ H_1 : p \neq 1/2 \end{cases}$$

Một người tung đồng xu 100 lần và ghi lại kết quả.

(a) Hãy đưa ra một "quy định" để bác bỏ H_0. Biểu diễn quy định này bởi một thống kê kiểm định và miền bác bỏ tương ứng.

Ví dụ

Trong bài toán tung đồng xu, ta có

$$\begin{cases} H_0 : p = 1/2 \\ H_1 : p \neq 1/2 \end{cases}$$

Một người tung đồng xu 100 lần và ghi lại kết quả.

(a) Hãy đưa ra một "quy định" để bác bỏ H_0. Biểu diễn quy định này bởi một thống kê kiểm định và miền bác bỏ tương ứng.

(b) Trong các trường hợp sau, trường hợp nào ta bác bỏ H_0 ?
- Tung 100 được 50 lần mặt ngửa, 50 lần mặt úp.
- Tung 100 được 49 lần mặt ngửa, 51 lần mặt úp.
- Tung 100 được 52 lần mặt ngửa, 48 lần mặt úp.
- Tung 100 được 70 lần mặt ngửa, 30 lần mặt úp.

Dẫn nhập về kiểm định giả thuyết thống kê

Sai lầm của kiểm định

	H_0 đúng	H_0 sai
Bác bỏ H_0	***Sai lầm loại I***	✓
Không đủ cơ sở để bác bỏ H_0	✓	***Sai lầm loại II***

Sai lầm của kiểm định

	H_0 đúng	H_0 sai
Bác bỏ H_0	***Sai lầm loại I***	✓
Không đủ cơ sở để bác bỏ H_0	✓	***Sai lầm loại II***

Mức ý nghĩa của kiểm định

Số $\alpha \in (0, 1)$ thể hiện xác suất xảy ra sai lầm loại I được gọi là ***mức ý nghĩa*** của kiểm định.

$$\alpha = \mathbb{P}(\text{Phép kiểm định bác bỏ } H_0 | \text{ Khi } H_0 \text{ đúng})$$

Sai lầm của kiểm định

	H_0 đúng	H_0 sai
Bác bỏ H_0	***Sai lầm loại I***	✓
Không đủ cơ sở để bác bỏ H_0	✓	***Sai lầm loại II***

Mức ý nghĩa của kiểm định

Số $\alpha \in (0,1)$ thể hiện xác suất xảy ra sai lầm loại I được gọi là ***mức ý nghĩa*** của kiểm định.

$$\alpha = \mathbb{P}(\text{Phép kiểm định bác bỏ } H_0 | \text{ Khi } H_0 \text{ đúng})$$

Mức ý nghĩa α thể hiện "độ nhạy" của phép kiểm định. Số α càng nhỏ thì phép kiểm định càng "nhạy".

Độ mạnh của kiểm định

Số $\beta \in (0, 1)$ thể hiện xác suất xảy ra sai lầm loại II, tức

$$\beta = \mathbb{P}(\text{Phép kiểm định chấp nhận } H_0 | \text{ Khi } H_0 \text{ sai})$$

Khi đó, giá trị của $1 - \beta$ được gọi là ***độ mạnh*** của kiểm định.

Ví dụ

Cho (X_1, X_2) là mẫu ngẫu nhiên được lấy từ tổng thể có phân phối chuẩn $\mathcal{N}(\mu, 1)$. Xét bài toán kiểm định sau:

$$\begin{cases} H_0 : \mu = 0 \\ H_1 : \mu = 1 \end{cases}$$

Sử dụng thống kê kiểm định $T = X_1 + X_2$ và ta bác bỏ H_0 nếu $|T| > c$.

(a) Xác định miền bác bỏ R của kiểm định.

(b) Với $c = 3$, tìm mức ý nghĩa và độ mạnh của phép kiểm định này.

(c) Tìm giá trị của c để mức ý nghĩa của phép kiểm định này là $\alpha = 0.05$

p-giá trị

p-giá trị

Cho bài toán kiểm định giả thuyết thống kê với giả thuyết không H_0, và đối thuyết H_1. Khi đó, p-giá trị là mức ý nghĩa nhỏ nhất mà ta có thể bác bỏ H_0.

Khi ta có mức ý nghĩa α và p-giá trị, ta sẽ bác bỏ H_0 (với mức ý nghĩa α) nếu p-giá trị $\leq \alpha$

Kiểm định về kì vọng (một mẫu)

Trường hợp phương sai σ^2 đã biết

Cho mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$ được lấy từ tổng thể có kì vọng μ chưa biết, và phương sai σ^2 đã biết, cỡ mẫu $n \geq 30$ hoặc $n < 30$ và mẫu có phân phối chuẩn. Với giá trị μ_0 cho trước, xét bài toán kiểm định sau:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

(tương ứng, $H_1 : \mu < \mu_0$, và $H_1 : \mu > \mu_0$).

Trường hợp phương sai σ^2 đã biết

Cho mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$ được lấy từ tổng thể có kì vọng μ chưa biết, và phương sai σ^2 đã biết, cỡ mẫu $n \geq 30$ hoặc $n < 30$ và mẫu có phân phối chuẩn. Với giá trị μ_0 cho trước, xét bài toán kiểm định sau:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

(tương ứng, $H_1 : \mu < \mu_0,$ và $H_1 : \mu > \mu_0$).

- **Thống kê kiểm định:** $T = \sqrt{n} \cdot \frac{\overline{X} - \mu_0}{\sigma}$. Khi H_0 đúng, ta có $T \approx \mathcal{N}(0, 1)$.
- **Miền bác bỏ:**

(TH1) $H_1 : \mu \neq \mu_0$, bác bỏ H_0 nếu $|T| > z_{1-\alpha/2}$

(TH2) $H_1 : \mu < \mu_0$, bác bỏ H_0 nếu $T < z_{\alpha}$

(TH3) $H_1 : \mu > \mu_0$, bác bỏ H_0 nếu $T > z_{1-\alpha}$

Kiểm định về kì vọng (một mẫu)

Trường hợp phương sai σ^2 đã biết (dùng p-giá trị)

Cho mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$ được lấy từ tổng thể có kì vọng μ chưa biết, và phương sai σ^2 đã biết, cỡ mẫu $n \geq 30$ hoặc $n < 30$ và mẫu có phân phối chuẩn. Với giá trị μ_0 cho trước, xét bài toán kiểm định sau:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

(tương ứng, $H_1 : \mu < \mu_0,$ và $H_1 : \mu > \mu_0$).

Kiểm định về kì vọng (một mẫu)

Trường hợp phương sai σ^2 đã biết (dùng p-giá trị)

Cho mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$ được lấy từ tổng thể có kì vọng μ chưa biết, và phương sai σ^2 đã biết, cỡ mẫu $n \geq 30$ hoặc $n < 30$ và mẫu có phân phối chuẩn. Với giá trị μ_0 cho trước, xét bài toán kiểm định sau:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

(tương ứng, $H_1 : \mu < \mu_0,$ và $H_1 : \mu > \mu_0$).

- **Thống kê kiểm định:** $T = \sqrt{n} \cdot \frac{\overline{X} - \mu_0}{\sigma}$. Đặt $t(\mathbf{x})$ là giá trị tính được trên mẫu.
- ***p*-giá trị**:

$$\begin{array}{ll} \text{(TH1)} & H_1 : \mu \neq \mu_0,\ p\text{-giá trị} = 2 - 2 \cdot \Phi(|t(\mathbf{x})|) \\ \text{(TH2)} & H_1 : \mu < \mu_0,\ p\text{-giá trị} = \Phi(t(\mathbf{x})) \\ \text{(TH3)} & H_1 : \mu > \mu_0,\ p\text{-giá trị} = 1 - \Phi(t(\mathbf{x})) \end{array}$$

- **Miền bác bỏ:** Bác bỏ H_0 khi p-giá trị $\leq \alpha$.

Trường hợp phương sai σ^2 chưa biết

Cho mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$ được lấy từ tổng thể có kì vọng μ chưa biết, và phương sai σ^2 chưa biết, cỡ mẫu $n \geq 30$ hoặc $n < 30$ và mẫu có phân phối chuẩn. Với giá trị μ_0 cho trước, xét bài toán kiểm định sau:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

(tương ứng, $H_1 : \mu < \mu_0$, và $H_1 : \mu > \mu_0$).

Kiểm định về kì vọng (một mẫu)

Trường hợp phương sai σ^2 chưa biết

Cho mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$ được lấy từ tổng thể có kì vọng μ chưa biết, và phương sai σ^2 chưa biết, cỡ mẫu $n \geq 30$ hoặc $n < 30$ và mẫu có phân phối chuẩn. Với giá trị μ_0 cho trước, xét bài toán kiểm định sau:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

(tương ứng, $H_1 : \mu < \mu_0$, và $H_1 : \mu > \mu_0$).

- **Thống kê kiểm định:** $T = \sqrt{n} \cdot \frac{\overline{X} - \mu_0}{S}$. Khi H_0 đúng, $T \sim \text{Student}(n-1)$.
- **Miền bác bỏ:**

(TH1) $H_1 : \mu \neq \mu_0$, bác bỏ H_0 nếu $|T| > t^{n-1}_{1-\alpha/2}$

(TH2) $H_1 : \mu < \mu_0$, bác bỏ H_0 nếu $T < t^{n-1}_{\alpha}$

(TH3) $H_1 : \mu > \mu_0$, bác bỏ H_0 nếu $T > t^{n-1}_{1-\alpha}$

Trường hợp phương sai σ^2 chưa biết

Cho mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$ được lấy từ tổng thể có kì vọng μ chưa biết, và phương sai σ^2 chưa biết, cỡ mẫu $n \geq 30$ hoặc $n < 30$ và mẫu có phân phối chuẩn. Với giá trị μ_0 cho trước, xét bài toán kiểm định sau:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

(tương ứng, $H_1 : \mu < \mu_0$, và $H_1 : \mu > \mu_0$).

Kiểm định về kì vọng (một mẫu)

Trường hợp phương sai σ^2 chưa biết

Cho mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$ được lấy từ tổng thể có kì vọng μ chưa biết, và phương sai σ^2 chưa biết, cỡ mẫu $n \geq 30$ hoặc $n < 30$ và mẫu có phân phối chuẩn. Với giá trị μ_0 cho trước, xét bài toán kiểm định sau:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

(tương ứng, $H_1 : \mu < \mu_0$, và $H_1 : \mu > \mu_0$).

- **Thống kê kiểm định:** $T = \sqrt{n} \cdot \frac{\overline{X} - \mu_0}{S}$. Đặt $t(\boldsymbol{x})$ là giá trị tính được trên mẫu.
- ***p*-giá trị:**

(TH1) $H_1 : \mu \neq \mu_0$, p-giá trị $= 2 - 2 \cdot \mathbb{P}(T \leq |t(\boldsymbol{x})|)$.
(TH2) $H_1 : \mu < \mu_0$, p-giá trị $= \mathbb{P}(T < t(\boldsymbol{x}))$.
(TH3) $H_1 : \mu > \mu_0$, p-giá trị $= 1 - \mathbb{P}(T \leq t(\boldsymbol{x}))$

  với $T \sim \text{Student}(n-1)$.

- **Miền bác bỏ:** Bác bỏ H_0 nếu p-giá trị $\leq \alpha$.

Kiểm định về kì vọng (một mẫu)

Ví dụ 1

Một mẫu ngẫu nhiên 100 hồ sơ tử tuất trong năm qua cho thấy rằng tuổi thọ trung bình là 71.8 năm. Giả sử độ lệch chuẩn của tổng thể là 8.9 năm, điều này có cho thấy rằng tuổi thọ trung bình ngày nay cao hơn 70 năm hay không ? Sử dụng mức ý nghĩa 0.05.

Ví dụ 1

Một mẫu ngẫu nhiên 100 hồ sơ tử tuất trong năm qua cho thấy rằng tuổi thọ trung bình là 71.8 năm. Giả sử độ lệch chuẩn của tổng thể là 8.9 năm, điều này có cho thấy rằng tuổi thọ trung bình ngày nay cao hơn 70 năm hay không ? Sử dụng mức ý nghĩa 0.05.

Ví dụ 2: Bài tập 5.32

Một thử nghiệm điểm nóng chảy của $n = 10$ mẫu của một chất kết dính được sử dụng trong sản xuất một chất đẩy nhiên liệu tên lửa dẫn với trung bình $\bar{x} = 154.2°F$. Giả sử điểm nóng chảy có phân phối chuẩn với $\sigma = 1.5°F$.

(a) Kiểm định $H_0 : \mu = 155$ với đối thuyết $H_1 : \mu \neq 155$ với $\alpha = 0.01$.

(b) Tính p-giá trị.

(c) Tính β nếu biết $\mu = 150$.

Ví dụ 3: Bài tập 5.51

Một máy đóng gói các sản phẩm có khối lượng 1 kg. Nghi ngờ máy hoạt động không bình thường, người ta chọn ra một mẫu ngẫu nhiên gồm 100 sản phẩm thì thấy như sau:

Khối lượng	0.95	0.97	0.99	1.01	1.03	1.05
Số gói	9	31	40	15	3	2

Với mức ý nghĩa 0.05, hãy kết luận về nghi ngờ trên.

So sánh tỉ lệ với một số

Cho $(X_1, X_2, \ldots, X_n) \sim \text{Ber}(p)$ với tham số p chưa biết. Với giá trị p_0 cho trước, giả sử $np_0 \geq 5$, và $n(1-p_0) \geq 5$, xét bài toán kiểm định giả thuyết thống kê:

$$\begin{cases} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{cases}$$

(tương ứng $H_1 : p < p_0$, $H_1 : p > p_0$).

- **Thống kê kiểm định:** $T = \sqrt{n}\frac{\bar{p}-p_0}{\sqrt{p_0(1-p_0)}}$, với $\bar{p} = \frac{1}{n}\sum_{k=1}^{n} X_k$. Khi H_0 đúng, ta có $T \approx \mathcal{N}(0,1)$.
- **Miền bác bỏ:**

(TH1) $H_1 : p \neq p_0$, bác bỏ H_0 nếu $|T| > z_{1-\alpha/2}$.
(TH2) $H_1 : p < p_0$, bác bỏ H_0 nếu $T < z_{\alpha}$.
(TH3) $H_1 : p > p_0$, bác bỏ H_0 nếu $T > z_{1-\alpha}$.

So sánh tỉ lệ với một số (p-giá trị)

Cho $(X_1, X_2, \ldots, X_n) \sim \text{Ber}(p)$ với tham số p chưa biết. Với giá trị p_0 cho trước, giả sử $np_0 \geq 5$, và $n(1-p_0) \geq 5$, xét bài toán kiểm định giả thuyết thống kê:

$$\begin{cases} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{cases}$$

(tương ứng $H_1 : p < p_0$, $H_1 : p > p_0$).

- **Thống kê kiểm định:** $T = \frac{\bar{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$, với $\bar{p} = \frac{1}{n}\sum_{k=1}^{n} X_k$.

  Khi H_0 đúng, ta có $T \approx \mathcal{N}(0,1)$. Gọi $t(\boldsymbol{x})$ là giá trị của thống kê T tính trên mẫu $\boldsymbol{x}$.
- ***p*-giá trị**:

(TH1) $H_1 : p \neq p_0$, p-giá trị $= 2 - 2 \cdot \Phi(|t(\boldsymbol{x})|)$

(TH2) $H_1 : p < p_0$, p-giá trị $= \Phi(t(\boldsymbol{x}))$

(TH3) $H_1 : p > p_0$, p-giá trị $= 1 - \Phi(t(\boldsymbol{x}))$

- **Miền bác bỏ:** Bác bỏ H_0 khi p-giá trị $\leq \alpha$.

Ví dụ: Bài tập 5.73

Giả sử rằng, 1000 khách hàng được khảo sát và 850 người hài lòng hoặc rất hài lòng với các sản phẩm và dịch vụ của công ty. Gọi p là tỉ lệ khách hàng hài lòng và rất hài lòng với các sản phẩm và dịch vụ của công ty.

(a) Kiểm định giả thuyết $H_0 : p = 0.9$ với đối thuyết $H_1 : p \neq 0.9$ với mức ý nghĩa $\alpha = 0.05$.

(b) Tính p-giá trị cho bài toán ở câu (a).

So sánh tỉ lệ với một số

Ví dụ: Bài tập 5.73

Giả sử rằng, 1000 khách hàng được khảo sát và 850 người hài lòng hoặc rất hài lòng với các sản phẩm và dịch vụ của công ty. Gọi p là tỉ lệ khách hàng hài lòng và rất hài lòng với các sản phẩm và dịch vụ của công ty.

(a) Kiểm định giả thuyết $H_0 : p = 0.9$ với đối thuyết $H_1 : p \neq 0.9$ với mức ý nghĩa $\alpha = 0.05$.

(b) Tính p-giá trị cho bài toán ở câu (a).

Ví dụ: Bài tập 5.74

Giả sử người ta kiểm tra 500 thành phần máy móc do một nhà máy sản xuất và thấy có 10 thành phần bị loại bỏ. Kiểm định giả thuyết $H_0 : p = 0.03$ với đối thuyết $H_1 : p < 0.03$ với mức ý nghĩa $\alpha = 0.05$. Tìm p-giá trị.

So sánh trung bình hai mẫu

Cho ***hai mẫu ngẫu nhiên độc lập*** $(X_1, X_2, \ldots, X_n)$ và $(Y_1, Y_2, \ldots, Y_m)$. Giả sử hai mẫu này lần lượt được lấy từ các phân phối có kì vọng μ_X, μ_Y và phương sai σ_X^2, σ_Y^2, tức

$$\mathbb{E}(X_1) = \mu_X, \mathbb{E}(Y_1) = \mu_Y$$

$$\mathrm{Var}(X_1) = \sigma_X^2, \mathrm{Var}(Y_1) = \sigma_Y^2$$

Ta xét bài toán kiểm định sau:

$$\begin{cases} H_0 : \mu_X = \mu_Y \\ H_1 : \mu_X \neq \mu_Y \end{cases}$$

(tương ứng: $H_1 : \mu_X < \mu_Y$ và $H_1 : \mu_X > \mu_Y$)

So sánh trung bình hai mẫu - Trường hợp 1

Trường hợp 1: σ_X^2, σ_Y^2 đã biết

- **Giả định:**
  - Phương sai σ_X^2, σ_Y^2 đã biết.
  - Cỡ mẫu $n, m \geq 30$.
  - Nếu $n < 30$ hoặc $m < 30$ thì hai mẫu phải có phân phối chuẩn.
- **Thống kê kiểm định**: $T = \dfrac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$. Khi H_0 đúng thì $T \approx \mathcal{N}(0, 1)$.
- **Miền bác bỏ:**

(TH1) $H_1 : \mu_X \neq \mu_Y$, bác bỏ H_0 nếu $|T| > z_{1-\alpha/2}$.
(TH2) $H_1 : \mu_X < \mu_Y$, bác bỏ H_0 nếu $T < z_{\alpha}$.
(TH3) $H_1 : \mu_X > \mu_Y$, bác bỏ H_0 nếu $T > z_{1-\alpha}$.

Trường hợp 1: σ_X^2, σ_Y^2 đã biết

- **Giả định:**
  - Phương sai σ_X^2, σ_Y^2 đã biết.
  - Cỡ mẫu $n, m \geq 30$.
  - Nếu $n < 30$ hoặc $m < 30$ thì hai mẫu phải có phân phối chuẩn.
- **Thống kê kiểm định**: $T = \dfrac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$. Khi H_0 đúng thì $T \approx \mathcal{N}(0,1)$.
- **p-giá trị:** Gọi t là giá trị của thống kê kiểm định T tính trên mẫu.

(TH1) $H_1 : \mu_X \neq \mu_Y$, p-giá trị$= 2(1 - \Phi(|t|))$
(TH2) $H_1 : \mu_X < \mu_Y$, p-giá trị$= \Phi(t)$.
(TH3) $H_1 : \mu_X > \mu_Y$, p-giá trị$= 1 - \Phi(t)$.

- Bác bỏ H_0 nếu p-giá trị$\leq \alpha$

Ví dụ (Bài tập 5.92)

Hai máy được sử dụng để làm đầy các chai nhựa với khối lượng tịnh 16 ounce. Khối lượng làm đầy có thể được giả định có phân phối chuẩn với độ lệch chuẩn $\sigma_1 = 0.02$ ounce và $\sigma_2 = 0.025$ ounce. Một thành viên của đội ngũ nhân viên kỹ thuật chất lượng nghi ngờ rằng cả hai máy đều có cùng khối lượng trung bình. Một mẫu ngẫu nhiên gồm 10 chai được lấy ra từ đầu ra của mỗi máy:

$$\begin{aligned}
\text{Máy 1: } & 16.03, 16.01, 16.04, 15.96, 16.05,\\
& 15.98, 16.05, 16.02, 16.02, 15.99\\
\text{Máy 2: } & 16.02, 16.03, 15.97, 16.04, 15.96,\\
& 16.02, 16.01, 16.01, 15.99, 16.00
\end{aligned}$$

(a) Sự nghi ngờ của thành viên đội ngũ kỹ thuật chất lượng có hợp lý hay không ? Sử dụng $\alpha = 0.05$.

(b) Tính p-giá trị của bài toán kiểm định trên.

Xác suất thống kê - Kiểm định thống kê

Dẫn nhập

Kiểm định giả thuyết thống kê

Kiểm định kì vọng một mẫu

Kiểm định tỉ lệ một mẫu

So sánh trung bình hai mẫu

Dữ liệu cặp

So sánh tỉ lệ hai mẫu

Trường hợp 2: σ_X^2, σ_Y^2 không biết, mẫu lớn

- **Giả định:**
  - Phương sai σ_X^2, σ_Y^2 không biết.
  - Cỡ mẫu $n, m \geq 30$.
- **Thống kê kiểm định**: $T = \dfrac{\bar{X} - \bar{Y}}{\sqrt{\dfrac{S_X^2}{n} + \dfrac{S_Y^2}{m}}}$. Khi H_0 đúng thì

  $T \approx \mathcal{N}(0, 1)$.
- **Miền bác bỏ:**

(TH1) $H_1 : \mu_X \neq \mu_Y$, bác bỏ H_0 nếu $|T| > z_{1-\alpha/2}$.
(TH2) $H_1 : \mu_X < \mu_Y$, bác bỏ H_0 nếu $T < z_\alpha$.
(TH3) $H_1 : \mu_X > \mu_Y$, bác bỏ H_0 nếu $T > z_{1-\alpha}$.

Trường hợp 2: σ_X^2, σ_Y^2 không biết, mẫu lớn

- **Giả định:**
  - Phương sai σ_X^2, σ_Y^2 không biết.
  - Cỡ mẫu $n, m \geq 30$.
- **Thống kê kiểm định**: $T = \dfrac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}}$. Khi H_0 đúng thì $T \approx \mathcal{N}(0, 1)$.
- **p-giá trị:** Gọi t là giá trị của thống kê kiểm định T tính trên mẫu.

(TH1) $H_1 : \mu_X \neq \mu_Y$, p-giá trị $= 2(1 - \Phi(|t|))$.
(TH2) $H_1 : \mu_X < \mu_Y$, p-giá trị $= \Phi(t)$.
(TH3) $H_1 : \mu_X > \mu_Y$, p-giá trị $= 1 - \Phi(t)$.

- Bác bỏ H_0 nếu p-giá trị $\leq \alpha$.

Ví dụ: (Bài tập 5.96)

Xét kiểm định giả thuyết $H_0 : \mu_1 = \mu_2$ với đối thuyết $H_1 : \mu_1 \neq \mu_2$. Với cỡ mẫu $n_1 = n_2 = 15$ có $\bar{x}_1 = 4.7, \bar{x}_2 = 7.8$ và $s_1^2 = 4, s_2^2 = 6.25$. Giả sử $\sigma_1^2 = \sigma_2^2$, và mẫu lấy từ phân phối chuẩn. Sử dụng $\alpha = 0.05$, hãy:

(a) Kiểm định giả thuyết thống kê trên.

(b) Tính p-giá trị cho bài toán kiểm định trên.

Trường hợp 3: σ_X^2, σ_Y^2 không biết, $\sigma_X^2 = \sigma_Y^2$, mẫu nhỏ

- **Giả định:**
  - Phương sai σ_X^2, σ_Y^2 không biết, nhưng biết $\sigma_X^2 = \sigma_Y^2$.
  - Cỡ mẫu $n, m < 30$.
  - Hai mẫu có phân phối chuẩn, tức $(X_1, \ldots, X_n) \sim \mathcal{N}(\mu_X, \sigma_X^2), (Y_1, \ldots, Y_m) \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$.
- **Thống kê kiểm định**: $T = \frac{\bar{X} - \bar{Y}}{\sqrt{S^2\left(\frac{1}{n} + \frac{1}{m}\right)}}$, trong đó $S^2 = \frac{n-1}{n+m-2} S_X^2 + \frac{m-1}{n+m-2} S_Y^2$. Khi H_0 đúng thì $T \sim \text{Student}(n+m-2)$.
- **Miền bác bỏ:**

(TH1) $H_1 : \mu_X \neq \mu_Y$, bác bỏ H_0 nếu $|T| > t_{1-\alpha/2}^{n+m-2}$.

(TH2) $H_1 : \mu_X < \mu_Y$, bác bỏ H_0 nếu $T < t_{\alpha}^{n+m-2} = -t_{1-\alpha}^{n+m-2}$.

(TH3) $H_1 : \mu_X > \mu_Y$, bác bỏ H_0 nếu $T > t_{1-\alpha}^{n+m-2}$.

Trường hợp 3: σ_X^2, σ_Y^2 không biết, $\sigma_X^2 = \sigma_Y^2$, mẫu nhỏ

- **Giả định:**
  - Phương sai σ_X^2, σ_Y^2 không biết, nhưng biết $\sigma_X^2 = \sigma_Y^2$.
  - Cỡ mẫu $n, m < 30$.
  - Hai mẫu có phân phối chuẩn, tức $(X_1, \ldots, X_n) \sim \mathcal{N}(\mu_X, \sigma_X^2), (Y_1, \ldots, Y_m) \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$.
- **Thống kê kiểm định**: $T = \frac{\bar{X}-\bar{Y}}{\sqrt{S^2\left(\frac{1}{n}+\frac{1}{m}\right)}}$, trong đó $S^2 = \frac{n-1}{n+m-2}S_X^2 + \frac{m-1}{n+m-2}S_Y^2$. Khi H_0 đúng thì $T \sim \text{Student}(n+m-2)$.
- **p-giá trị:** Gọi t là giá trị của thống kê T tính được trên mẫu.

(TH1) $H_1 : \mu_X \neq \mu_Y$, p-giá trị $= 2(1 - \mathbb{P}(T \leq |t|))$

(TH2) $H_1 : \mu_X < \mu_Y$, p-giá trị $= \mathbb{P}(T \leq t)$.

(TH3) $H_1 : \mu_X > \mu_Y$, p-giá trị $= 1 - \mathbb{P}(T > t)$.

- Bác bỏ H_0 nếu p-giá trị $\leq \alpha$.

Trường hợp 4: σ_X^2, σ_Y^2 không biết, $\sigma_X^2 \neq \sigma_Y^2$, mẫu nhỏ

- **Giả định:**
  - Phương sai σ_X^2, σ_Y^2 không biết, nhưng biết $\sigma_X^2 \neq \sigma_Y^2$.
  - Cỡ mẫu $n, m < 30$.
  - Hai mẫu có phân phối chuẩn, tức $(X_1, \dots, X_n) \sim \mathcal{N}(\mu_X, \sigma_X^2), (Y_1, \dots, Y_m) \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$.
- **Thống kê kiểm định**: $T = \dfrac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}}$. Khi H_0 đúng thì

$$T \sim \text{Student(df), với df} = \frac{(s_X^2/n + s_Y^2/m)^2}{\frac{(s_X^2/n)^2}{n-1} + \frac{(s_Y^2/m)^2}{m-1}}$$

- **Miền bác bỏ:**

(TH1) $H_1 : \mu_X \neq \mu_Y$, bác bỏ H_0 nếu $|T| > t^{\text{df}}_{1-\alpha/2}$.

(TH2) $H_1 : \mu_X < \mu_Y$, bác bỏ H_0 nếu $T < t^{\text{df}}_{\alpha} = -t^{\text{df}}_{1-\alpha}$.

(TH3) $H_1 : \mu_X > \mu_Y$, bác bỏ H_0 nếu $T > t^{\text{df}}_{1-\alpha}$.

So sánh trung bình hai mẫu - Trường hợp 4

Trường hợp 4: σ_X^2, σ_Y^2 không biết, $\sigma_X^2 \neq \sigma_Y^2$, mẫu nhỏ

- **Giả định:**
  - Phương sai σ_X^2, σ_Y^2 không biết, nhưng biết $\sigma_X^2 \neq \sigma_Y^2$.
  - Cỡ mẫu $n, m < 30$.
  - Hai mẫu có phân phối chuẩn, tức $(X_1, \ldots, X_n) \sim \mathcal{N}(\mu_X, \sigma_X^2), (Y_1, \ldots, Y_m) \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$.
- **Thống kê kiểm định**: $T = \dfrac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}}$. Khi H_0 đúng thì

$$T \sim \text{Student}(df), \text{ với } df = \frac{(s_X^2/n + s_Y^2/m)^2}{\frac{(s_X^2/n)^2}{n-1} + \frac{(s_Y^2/m)^2}{m-1}}$$

- **p-giá trị:** Gọi t là giá trị của thống kê T tính trên mẫu.

(TH1) $H_1 : \mu_X \neq \mu_Y$, p-giá trị $2(1 - \mathbb{P}(T \leq |t|))$.
(TH2) $H_1 : \mu_X < \mu_Y$, p-giá trị $\mathbb{P}(T \leq t)$.
(TH3) $H_1 : \mu_X > \mu_Y$, p-giá trị $1 - \mathbb{P}(T \leq t)$.

- Bác bỏ H_0 khi p-giá trị $\leq \alpha$.

Dữ liệu cặp

Dữ liệu cặp (Paired datasets)

Cho hai mẫu ngẫu nhiên có cùng cỡ mẫu n, $(X_1, X_2, \ldots, X_n)$ và $(Y_1, Y_2, \ldots, Y_n)$. Ta nói hai mẫu này tạo thành dữ liệu cặp nếu:

- Cặp biến ngẫu nhiên (X_i, Y_i) và (X_j, Y_j) là hai vecto ngẫu nhiên độc lập, với $i \neq j$.
- Hai biến ngẫu nhiên X_i và Y_i ***không độc lập.***

Một số trường hợp mà ta có thể gặp dữ liệu dạng cặp:

- (X_i, Y_i) là chỉ số huyết áp của bệnh nhân thứ i trước và sau khi sử dụng thuốc điều trị.
- (X_i, Y_i) là cân nặng của người thứ i trước và sau khi thực hiện chế độ ăn kiêng.
- (X_i, Y_i) là điểm của học sinh thứ i theo 2 phương pháp kiểm tra: Kiểm tra viết và kiểm tra nói.

Kiểm định với dữ liệu dạng cặp

Cho bộ dữ liệu dạng cặp $(X_1, X_2, \ldots, X_n)$ và $(Y_1, \ldots, Y_n)$.

- **Các giả định:** $(X_1, X_2, \ldots, X_n) \sim \mathcal{N}(\mu_1, \sigma_1^2)$ và $(Y_1, Y_2, \ldots, Y_n) \sim \mathcal{N}(\mu_2, \sigma_2^2)$.
- **Bài toán kiểm định**:
$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$
(hoặc $H_1 : \mu_1 < \mu_2, H_1 : \mu_1 > \mu_2$)
- **Thống kê kiểm định:** $T = \sqrt{n} \cdot \dfrac{\bar{D}}{S_D}$, trong đó $\bar{D}$ và S_D lần lượt là trung bình mẫu, độ lệch chuẩn mẫu của $(D_1, D_2, \ldots, D_n)$ với $D_k = X_k - Y_k$. Khi H_0 đúng, ta có $T \sim \text{Student}(n-1)$.

Miền bác bỏ và p-giá trị của kiểm định

Gọi t là giá trị của thống kê kiểm định tính được trên mẫu. Ta có bảng sau:

	Miền bác bỏ	p-giá trị
$H_1 : \mu_D \neq 0$	$\|T\| > t^{n-1}_{1-\alpha/2}$	$2(1 - \mathbb{P}(T \leq \|t\|))$
$H_1 : \mu_D < 0$	$T < t^{n-1}_{\alpha}$	$\mathbb{P}(T \leq t)$
$H_1 : \mu_D > 0$	$T > t^{n-1}_{1-\alpha}$	$1 - \mathbb{P}(T \leq t)$

trong đó, t^{n-1}_{β} là phân vị thứ β của phân phối Student, bậc tự do $n-1$, và $\mu_D = \mu_1 - \mu_2$.

Ví dụ: Bài tập 5.124

Một viện nghiên cứu được thực hiện bởi Khoa Thực phẩm và Dinh Dưỡng Con Người ở Viện Kỹ Thuât Virginia, dữ liệu sau đây được ghi lại về các thặng dư chất bảo quản chống nấm mốc axit sorbic, tính theo một phần triệu trong giăm bông sau khi ngâm trong dung dịch sorbate và sau 60 ngày bảo quản:

Miếng	Trước bảo quản	Sau bảo quản
1	224	116
2	270	96
3	400	239
4	444	329
5	590	437
6	660	597
7	1400	689
8	680	576

Ví dụ: Bài tập 5.124 (tt)

Giả sử các tổng thể có phân phối chuẩn, hỏi có đủ bằng chứng, tại mức ý nghĩa 0.05, để nói rằng thời gian lưu trữ ảnh hưởng đến nồng độ thặng dư axit sorbic hay không ?

Đáp án.

Gọi $(X_1, X_2, \ldots, X_8)$ và $(Y_1, Y_2, \ldots, Y_8)$ lần lượt là nồng độ thặng dư axit sorbic trước và sau khi bảo quản. Giả sử nồng độ trung bình của hai mẫu là μ_1 và μ_2. Ta có bài toán kiểm định

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$$

với mức ý nghĩa $\alpha = 0.05$.

Đáp án. (tt)

Đặt $D_i = Y_i - X_i$, thể hiện sự sai khác về nồng độ thặng dư axit sorbic trước và sau khi bảo quản. Khi đó, bài toán kiểm định được viết lại thành

$$\begin{cases} H_0 : \mu_D = 0 \\ H_1 : \mu_D < 0 \end{cases}$$

với $\mu_D = \mu_2 - \mu_1$.

- Thống kê kiểm định: $T = \sqrt{n}\dfrac{\bar{D}}{S_D}$. Khi H_0 đúng, $T \sim \text{Student}(n-1)$.
- Trên mẫu, ta có: $n = 8, \bar{d} = -198.625, s_d \approx 210.165$ và giá trị của thống kê kiểm định $t \approx -2.673$.
- p-giá trị $= \mathbb{P}(T \leq t) = \mathbb{P}(T \leq -2.673) = 1 - \mathbb{P}(T \leq 2.673)$ $\approx 0.0159 \leq \alpha$. Vậy ta bác bỏ H_0. tức nồng độ thặng dư axit sorbic sau bảo quản thấp hơn nồng độ thặng dư axit sorbic trước khi bảo quản, với mức ý nghĩa $\alpha = 0.05$

Đáp án. (tt)

Hoặc ta có thể sử dụng miền bác bỏ:

- Miền bác bỏ: Ta bác bỏ H_0 khi $T < t_{\alpha}^{n-1} = -t_{1-\alpha}^{n-1} = -t_{0.95}^{7} \approx -1.8946$.
- Do $t \approx -2.673 < -t_{0.95}^{7}$ nên ta bác bỏ H_0. tức nồng độ thặng dư axit sorbic sau bảo quản thấp hơn nồng độ thặng dư axit sorbic trước khi bảo quản, với mức ý nghĩa $\alpha = 0.05$.

So sánh tỉ lệ hai mẫu

Cho ***hai mẫu ngẫu nhiên độc lập*** $(X_1, X_2, \ldots, X_n) \sim \text{Ber}(p_X)$, và $(Y_1, Y_2, \ldots, Y_m) \sim \text{Ber}(p_Y)$. Ta xét bài toán kiểm định sau:

$$\begin{cases} H_0 : p_X = p_Y \\ H_1 : p_X \neq p_Y \end{cases}$$

(tương ứng $H_1 : p_X < p_Y, H_1 : p_X > p_Y$).

Thống kê kiểm định

Thống kê kiểm định cho bài toán này là

$$Z = \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n} + \frac{1}{m}\right)}}, \text{ với } \hat{p} = \frac{n}{n+m}\hat{p}_X + \frac{m}{n+m}\hat{p}_Y$$

Khi H_0 đúng ($p_X = p_Y$), và cỡ mẫu lớn ($n, m \geq 30$), ta có $Z \approx \mathcal{N}(0, 1)$

Miền bác bỏ và p-giá trị

Gọi z là giá trị tính được trên mẫu. Ta có bảng sau:

	Miền bác bỏ	p-giá trị
$H_1 : p_X \neq p_Y$	$\lvert z \rvert > z_{1-\alpha/2}$	$2(1 - \Phi(\lvert z \rvert))$
$H_1 : p_X < p_Y$	$z < z_\alpha$	$\Phi(z)$
$H_1 : p_X > p_Y$	$z > z_{1-\alpha}$	$1 - \Phi(z)$

Ví dụ (Bài tập 5.126)

Hai loại máy ép phun khác nhau được sử dụng để tạo thành các bộ phận bằng nhựa. Một phần được coi là khiếm khuyết nếu nó bị co rút quá mức hoặc bị đổi màu. Hai mẫu ngẫu nhiên , mỗi mẫu có kích thước 300, được chọn, và 15 bộ phận bị lỗi được tìm thấy trong mẫu từ máy 1, và 8 bộ phận bị lỗi được tìm thấy trong mẫu từ máy 2. Có hợp lý để kết luận rằng cả hai máy sản xuất cùng một tỉ lệ các bộ phận lỗi, sử dụng $\alpha = 0.05$? Tìm p-giá trị cho kiểm định này.