Ước lượng và Kiểm định trong Thống kê

Ngô Hoàng Long
Trường Đại học Sư phạm Hà Nội

VIASM, 17-21/8/2020

1. Khởi động
2. Ước lượng điểm
3. Một số phân phối quan trọng
4. Ước lượng khoảng
5. Kiểm định giả thuyết

Khởi động

- Trong hộp có n quả bóng đánh số từ 1 đến n.
- Số quả bóng n là chưa biết.
- Bạn Ngô Nga lấy ra ngẫu nhiên m quả bóng từ hộp và xem số của nó.
- Hãy giúp bạn Ngô Nga ước lượng số bóng trong hộp.

Phương pháp ước lượng hợp lý cực đại

Cho $\mathbf{X} = (X_1, \ldots, X_n)$ là một mẫu ngẫu nhiên từ phân phối với hàm mật độ $f(x; \theta)$. Hàm hợp lý xác định bởi

$$L(\mathbf{x}; \theta) = \prod_{i=1}^{n} f(x_i; \theta).$$

Định nghĩa

Với mỗi điểm mẫu $\mathbf{x}$, đặt $\hat{\theta}(\mathbf{x})$ là một giá trị tham số mà tại đó $L(\mathbf{x}; \theta)$ đạt cực đại như là một hàm số của θ, với $\mathbf{x}$ cố định. Một ước lượng hợp lý cực đại của tham số θ dựa trên một mẫu $\mathbf{X}$ là $\hat{\theta}(\mathbf{X})$.

Kí hiệu $X_1, \ldots, X_n$ là một mẫu ngẫu nhiên từ phân phối với hàm mật độ $f(x;\theta),\ \theta \in \Theta$. Kí hiệu θ_0 là giá trị đúng của θ.

Định lý

Giả sử rằng

(R0) $f(.;\theta) \neq f(.;\theta')$ *với mọi* $\theta \neq \theta'$;

(R1) *mọi hàm* $f(.;\theta), \theta \in \Theta$ *đều có giá chung với mọi* θ.

Khi đó

$$\lim_{n\to\infty} \mathbb{P}_{\theta_0}[L(\mathbf{X};\theta_0) > L(\mathbf{X};\theta)] = 1, \qquad \text{với mọi} \ \ \theta \neq \theta_0.$$

Phương pháp moment

Cho $(X_1, \ldots, X_n)$ là một mẫu ngẫu nhiên từ một phân phối với hàm mật độ $f(x;\theta)$ trong đó $\theta = (\theta_1, \ldots, \theta_k) \in \Theta \subset \mathbb{R}^k$. Các ước lượng bằng phương pháp moment được tìm ra bằng cách lập k phương trình của k moment mẫu đầu tiên với k moment quần thể tương ứng, và giải hệ phương trình. Cụ thể hơn, ta định nghĩa

$$\mu_j = \mathbb{E}[X^j] = g_j(\theta_1, \ldots, \theta_k), \quad j = 1, \ldots, k.$$

và

$$m_j = \frac{1}{n}\sum_{i=1}^{n} X_i^j.$$

Ước lượng moment $(\hat{\theta}_1, \ldots, \hat{\theta}_k)$ thu được bằng cách giải hệ phương trình

$$m_j = g_j(\theta_1, \ldots, \theta_k),\ j = 1, \ldots, k.$$

Phân phối Gamma

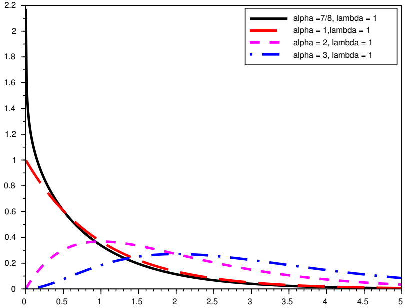
Định nghĩa

Một biến ngẫu nhiên X được gọi là có phân phối Gamma $\mathcal{G}(\alpha, \lambda)$ nếu hàm mật độ xác suất của nó được xác định bởi

$$f_X(x) = \frac{x^{\alpha-1} e^{-x/\lambda}}{\Gamma(\alpha)\lambda^{\alpha}} I_{\{x>0\}},$$

trong đó, $\Gamma(\alpha) = \int\limits_0^{+\infty} x^{\alpha-1} e^{-x} dx$ được gọi là hàm số Gamma.

Chú ý rằng $\mathcal{G}(1, \lambda) = Exp(\lambda)$.



Hình: Mật độ của phân phối Gamma

Tính chất của phân phối Gamma

Mệnh đề

Nếu X có phân phối $\mathcal{G}(\alpha, \lambda)$ thì

$$\mathbb{E}[X] = \alpha\lambda, \quad DX = \alpha\lambda^2.$$

Hơn nữa, hàm đặc trưng của X được xác định bởi

$$\varphi_X(t) = \int_0^\infty e^{itx} \frac{x^{\alpha-1} e^{-x/\lambda}}{\Gamma(\alpha)\lambda^\alpha} dx = \left(\frac{1}{1 - i\lambda t}\right)^\alpha.$$

Hệ quả

Cho $(X_i)_{1 \le i \le n}$ là dãy các biến ngẫu nhiên độc lập. Giả sử với mỗi i, X_i có phân phối $\mathcal{G}(\alpha_i, \lambda)$. Khi đó, $S = X_1 + \cdots + X_n$ có phân phối $\mathcal{G}(\alpha_1 + \cdots + \alpha_i, \lambda)$.

Phân phối khi bình phương

Định nghĩa

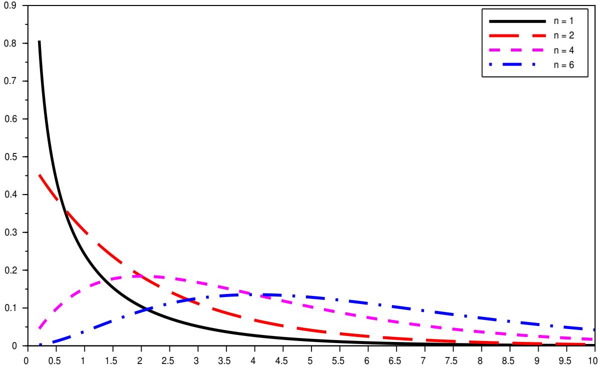
Cho $(Z_i)_{1\leq i\leq n}$ là một dãy các biến ngẫu nhiên độc lập cùng phân phối chuẩn tắc. Khi đó, phân phối của $V = Z_1^2 + \ldots + Z_n^2$ được gọi là *phân phối khi bình phương với bậc tự do n* và được kí hiệu là χ_n^2.

Chú ý: vì Z_i^2 có phân phối $\mathcal{G}(\frac{1}{2}, 2)$ nên χ_n^2 có phân phối $\mathcal{G}(\frac{n}{2}, 2)$. Hơn nữa,

$$\mathbb{E}[\chi_n^2] = n, \quad D\chi_n^2 = 2n.$$

Tính chất của phân phối khi bình phương

Một hệ quả cần chú ý từ định nghĩa của phân phối khi bình phương là nếu hai biến ngẫu nhiên U và V độc lập với $U \sim \chi^2_n$ và $V \sim \chi^2_m$ thì $U + V \sim \chi^2_{m+n}$.



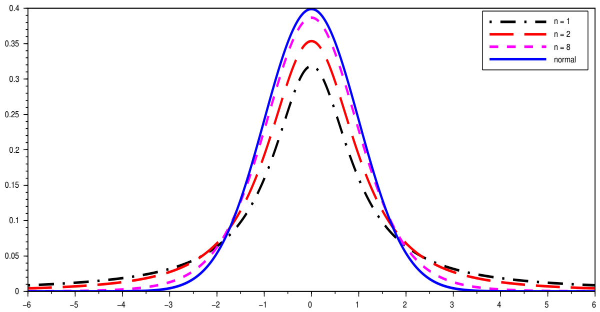
Hình: Mật độ của phân phối χ^2

Phân phối Student

Định nghĩa

Nếu Z và U là hai biến ngẫu nhiên độc lập với $Z \sim N(0;1)$ và $U \sim \chi^2_n$ thì phân phối của $\dfrac{Z}{\sqrt{U/n}}$ được gọi là *phân phối Student với bậc tự do n*.

Phân phối Student còn được gọi là phân phối t.



Tính chất của phân phối Student

Mệnh đề

Hàm mật độ xác suất của phân phối Student với bậc tự do n là

$$f_n(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-(n+1)/2}.$$

Ngoài ra,

$$f_n(t) \xrightarrow{n\to\infty} \frac{1}{\sqrt{2\pi}}e^{-t^2/2}.$$

Phân phối F

Định nghĩa

Cho U và V là hai biến ngẫu nhiên độc lập có cùng phân phối khi bình phương với bậc tự do lần lượt là m và n. Khi đó, phân phối của biến ngẫu nhiên

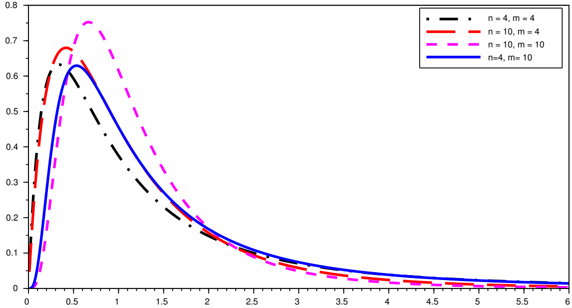
$$W = \frac{U/m}{V/n}$$

được gọi là *phân phối F với bậc tự do m và n* và được kí hiệu là $F_{m,n}$.

Hàm mật độ của phân phối F

Hàm mật độ xác suất của W là

$$f(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{m}{n}\right)^{m/2} x^{m/2-1} \left(1 + \frac{m}{n}x\right)^{-(m+n)/2}, \quad x \geq 0.$$



Hình: Mật độ của phân phối F

Phân phối của trung bình và phương sai mẫu

Cho (X_n) là dãy các biến ngẫu nhiên độc lập có cùng phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$, và đặt

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n-1}(X_i - \overline{X}_n)^2$$

- Biến ngẫu nhiên $\overline{X}_n$ và vectơ của các biến ngẫu nhiên $(X_1 - \overline{X}_n, X_2 - \overline{X}_n, \ldots, X_n - \overline{X}_n)$ độc lập với nhau.
- $\overline{X}_n$ và s_n^2 độc lập với nhau.
- Phân phối của $(n-1)s_n^2/\sigma^2$ là phân phối khi bình phương với $n-1$ bậc tự do.
- Đại lượng

$$\frac{\overline{X}_n - \mu}{s_n/\sqrt{n}} \sim t_{n-1}.$$

Ước lượng khoảng

- Ước lượng khoảng là gì?
- Tại sao lại cần ước lượng khoảng?
- Xây dựng ước lượng khoảng như thế nào?
- Độ tin cậy trong ước lượng khoảng có nghĩa là gì?

Định nghĩa ước lượng khoảng

Định nghĩa

Cho $(X_1, \dots, X_n)$ là một mẫu từ một phân phối $F(x, \theta)$, $\theta \in \Theta$. Một khoảng ngẫu nhiên (φ_1, φ_2), trong đó φ_1 và φ_2 là hai ước lượng của θ, được gọi là một *khoảng tin cậy* $(1 - \alpha)$ cho θ nếu

$$\mathbb{P}(\varphi_1 < \theta < \varphi_2) = 1 - \alpha,$$

với $\alpha \in [0, 1]$.

Câu hỏi: Ước lượng khoảng cho tham số có duy nhất không?

Định nghĩa ước lượng khoảng

Định nghĩa

Cho $(X_1, \ldots, X_n)$ là một mẫu từ một phân phối $F(x, \theta)$, $\theta \in \Theta$. Một khoảng ngẫu nhiên (φ_1, φ_2), trong đó φ_1 và φ_2 là hai ước lượng của θ, được gọi là một *khoảng tin cậy* $(1 - \alpha)$ cho θ nếu

$$\mathbb{P}(\varphi_1 < \theta < \varphi_2) = 1 - \alpha,$$

với $\alpha \in [0, 1]$.

Câu hỏi: Ước lượng khoảng cho tham số có duy nhất không?

Xây dựng ước lượng khoảng

- $(X_1, \ldots, X_n) \sim \mathcal{N}(a; \sigma^2)$ trong đó σ^2 đã biết.
- Ta biết rằng $\overline{x}_n$ là một ước lượng vững, không chệch của a. Nhưng $\overline{x}_n$ xấp xỉ a bao nhiêu?
- Bởi vì $\overline{x}_n \sim \mathcal{N}(a; \sigma^2/n)$, ta có $(\overline{x}_n - a)/(\sigma/\sqrt{n})$ có phân phối chuẩn tắc $\mathcal{N}(0; 1)$. Do đó,

$$0.954 = \mathbb{P}\Big[-2 < \frac{\overline{x}_n - a}{\sigma/\sqrt{n}} < 2\Big] = \mathbb{P}\Big[\overline{x}_n - 2\frac{\sigma}{\sqrt{n}} < a < \overline{x}_n + 2\frac{\sigma}{\sqrt{n}}\Big]. \qquad (1)$$

Xây dựng ước lượng khoảng

- Biểu thức

$$0.954 = \mathbb{P}\left[-2 < \frac{\overline{x}_n - a}{\sigma/\sqrt{n}} < 2\right] = \mathbb{P}\left[\overline{x}_n - 2\frac{\sigma}{\sqrt{n}} < a < \overline{x}_n + 2\frac{\sigma}{\sqrt{n}}\right]. \quad (2)$$

chỉ ra rằng trước khi lấy mẫu, xác suất để a thuộc vào khoảng ngẫu nhiên $\left(\overline{x}_n - 2\frac{\sigma}{\sqrt{n}} < a < \overline{x}_n + 2\frac{\sigma}{\sqrt{n}}\right)$ là 0.954.

- Sau khi lấy mẫu, khoảng thu được $\left(\overline{x}_n - 2\frac{\sigma}{\sqrt{n}} < a < \overline{x}_n + 2\frac{\sigma}{\sqrt{n}}\right)$ hoặc chứa a hoặc không.
- Nhưng vì xác suất thành công trước khi lấy mẫu rất cao nên ta có thể gọi khoảng $\left(\overline{x}_n - 2\frac{\sigma}{\sqrt{n}} < a < \overline{x}_n + 2\frac{\sigma}{\sqrt{n}}\right)$ là một *khoảng ước lượng với độ tin cậy* 95.4% cho a.

Xây dựng ước lượng khoảng

- Ta có thể nói, với sự tin cậy nhất định, $\overline{x}$ cách a một khoảng $2\frac{\sigma}{\sqrt{n}}$. Hằng số $0.954 = 95.4\%$ được gọi là *độ tin cậy*.
- Thay vì sử dụng 2, ta có thể dùng 1.645, 1.96 hoặc 2.576 để thu được các khoảng tin cậy $90\%, 95\%$ hoặc 99% cho a.
- Chú ý rằng độ dài các khoảng tin cậy tăng khi độ tin cậy tăng; nghĩa là, việc tăng độ tin cậy kéo theo việc giảm độ chính xác.
- Mặt khác, với hệ số tin cậy bất kì, việc tăng cỡ mẫu sẽ làm thu hẹp khoảng tin cậy.

Ước lượng trung bình của mẫu chuẩn với phương sai chưa biết

- Cho $X_1, \ldots, X_n$ là một mẫu từ một phân phối chuẩn $\mathcal{N}(a, \sigma^2)$.
- Biến ngẫu nhiên $T = (\overline{x} - a)/(s/\sqrt{n})$ có phân phối t với bậc tự do $n-1$.
- Với mỗi $\alpha \in (0,1)$, kí hiệu $t_{\alpha/2,n-1}$ thỏa mãn

$$\frac{\alpha}{2} = \mathbb{P}\Big(T > t_{\alpha/2,n-1}\Big).$$

- Vì tính đối xứng của phân phối t, ta có

$$\begin{aligned}1-\alpha &= \mathbb{P}\Big(-t_{\alpha/2,n-1} < T < t_{\alpha/2,n-1}\Big) = \mathbb{P}\Big(-t_{\alpha/2,n-1} < \frac{\overline{x}-a}{S/\sqrt{n}} < t_{\alpha/2,n-1}\Big) \\ &= \mathbb{P}\Big(\overline{x} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} < a < \overline{x} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}\Big).\end{aligned}$$

- Do đó, một khoảng tin cậy $(1-\alpha)$ cho a được xác định bởi

$$\Big(\overline{x} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}, \overline{x} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\Big). \tag{3}$$

Ước lượng tham số của mẫu cỡ lớn

- Giả sử $(X_1, \ldots, X_n) \sim F(x; \theta)$.
- θ_0 là giá trị đúng chưa biết của tham số θ.
- Giả sử φ là một ước lượng của θ_0 sao cho

$$\sqrt{n}(\varphi - \theta_0) \xrightarrow{w} \mathcal{N}(0, \sigma_\varphi^2). \tag{4}$$

- Tham số σ_φ^2 là phương sai tiệm cận của $\sqrt{n}\varphi$.
- Đặt $Z = \sqrt{n}(\varphi - \theta_0)/\sigma_\varphi$ là biến ngẫu nhiên chuẩn tắc hóa. Khi đó Z tiệm cận $\mathcal{N}(0, 1)$. Do đó, $\mathbb{P}[-1.96 < Z < 1.96] = 0.95$. Điều này suy ra

$$0.95 = \mathbb{P}\left[\varphi - 1.96\frac{\sigma_\varphi}{\sqrt{n}} < \theta_0 < \varphi + 1.96\frac{\sigma_\varphi}{\sqrt{n}}\right] \tag{5}$$

- Bởi vì khoảng $\left(\varphi - 1.96\frac{\sigma_\varphi}{\sqrt{n}} < \theta_0 < \varphi + 1.96\frac{\sigma_\varphi}{\sqrt{n}}\right)$ là một hàm của biến ngẫu nhiên φ nên ta sẽ gọi nó là một *khoảng ngẫu nhiên*. Xác suất để khoảng ngẫu nhiên này chứa θ_0 xấp xỉ 0.95.
- Vì trong thực tế ta thường chưa biết σ_φ nên ta giả sử tồn tại một ước lượng vững của σ_φ, kí hiệu là S_φ. Theo định lý Slutsky ta có

$$\frac{\sqrt{n}(\varphi - \theta_0)}{S_\varphi} \xrightarrow{w} N(0, 1).$$

Do đó, khoảng $\left(\varphi - 1.96 S_\varphi/\sqrt{n}, \varphi - 1.96 S_\varphi/\sqrt{n}\right)$ là một khoảng ngẫu nhiên với xác suất xấp xỉ 95% chứa θ_0.

Kiểm định giả thuyết

- Kiểm định giả thuyết là gì?
- Giả thuyết và đối thuyết được chọn như thế nào?
- Nguyên lý xây dựng tiêu chuẩn là gì?
- Phương pháp xây dựng tiêu chuẩn kiểm định?

Bài toán

Giả sử mối quan tâm của chúng ta tập trung ở một biến ngẫu nhiên X mà có hàm mật độ $f(x;\theta)$ trong đó $\theta \in \Theta$. Giả sử ta nghi ngờ, dựa trên lý thuyết hoặc một thí nghiệm ban đầu, rằng $\theta \in \Theta_0$ hoặc $\theta \in \Theta_1$ trong đó Θ_0 và Θ_1 là các tập con của Θ và $\Theta_0 \cup \Theta_1 = \Theta$. Ta đặt giả thuyết như sau

$$H_0 : \theta \in \Theta_0 \quad \text{và} \quad H_1 : \theta \in \Theta_1. \tag{6}$$

Ta gọi H_0 là *giả thuyết* và H_1 là *đối thuyết*.

- Giả thuyết H_0 biểu diễn sự không thay đổi hoặc sự không phân biệt từ quá khứ.
- Đối thuyết H_a1 biểu diễn sự thay đổi hoặc sự phân biệt. Đối thuyết thường là giả thuyết của những nhà nghiên cứu.
- Quy tắc quyết định lấy H_0 hay H_1 dựa trên một mẫu $X_1, \ldots, X_n$ từ phân phối của X và do đó, quyết định đấy có thể đúng hoặc sai.
- Có hai loại sai lầm ta thường gặp phải:
    - Bác bỏ H_0 khi H_0 đúng (được gọi là sai lầm loại 1)
    - Chấp nhận H_0 khi H_0 sai (được gọi là sai lầm loại 2).

Bảng: Bảng quyết định cho một kiểm định giả thuyết

Quyết định	H_0 đúng	H_1 đúng
Bác bỏ H_0	Sai lầm loại 1	Quyết định đúng
Chấp nhận H_0	Quyết định đúng	Sai lầm loại 2

Kí hiệu $\mathcal{D}$ là không gian mẫu. Một kiểm định của H_0 đối lập với H_1 dựa trên một tập con C của $\mathcal{D}$. Tập C được gọi là *miền tiêu chuẩn* và nguyên tắc quyết định tương ứng của nó là:

- Bác bỏ H_0 (Chấp nhận H_1) nếu $(X_1, \ldots, X_n) \in C$;
- Giữ lại H_0 (Bác bỏ H_1) nếu $(X_1, \ldots, X_n) \notin C$.

Nguyên lý chọn miền tiêu chuẩn: Miền tiêu chuẩn được chọn sao cho, một mặt, giữ xác suất sai lầm loại 1 tại mức α nào đó, mặt khác cực tiểu xác suất xảy ra sai lầm loại 2.

Định nghĩa

P-value là mức ý nghĩa nhỏ nhất làm bác bỏ giả thuyết H_0 với dữ liệu cho trước.

Điều này có nghĩa là nếu $\alpha \geq P$-value, ta sẽ bác bỏ H_0 trong khi nếu $\alpha < P$-value, ta sẽ không bác bỏ H_0.

Câu hỏi: Tại sao nên dùng *p*-value?

Định nghĩa

P-value là mức ý nghĩa nhỏ nhất làm bác bỏ giả thuyết H_0 với dữ liệu cho trước.

Điều này có nghĩa là nếu $\alpha \geq P$-value, ta sẽ bác bỏ H_0 trong khi nếu $\alpha < P$-value, ta sẽ không bác bỏ H_0.

Câu hỏi: Tại sao nên dùng p-value?

Phương pháp tỉ số hợp lý

Cho $L(\mathbf{x};\theta)$ là hàm hợp lý của mẫu $(X_1,\ldots,X_n)$ từ một phân phối có hàm mật độ $p(x;\theta)$.

Định nghĩa

Thống kê kiểm định hợp lý cho bài toán kiểm định giả thuyết
$H_0 : \theta \in \Theta_0$ đối lập với $H_1 : \theta \in \Theta_1$ là

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta\in\Theta_0} L(\mathbf{x};\theta)}{\sup_{\theta\in\Theta} L(\mathbf{x};\theta)}.$$

Một kiểm định tỉ số hợp lý là một kiểm định mà có một miền bác bỏ có dạng $C = \{\mathbf{x} : \lambda(\mathbf{x}) \le c\}$ với $c \in [0,1]$.

Ý tưởng của kiểm định tỉ số hợp lý từ thực tế là nếu θ_0 là giá trị đúng của θ thì $L(\theta_0)$ xấp xỉ giá trị cực đại của $L(\theta)$. Do đó, nếu H_0 đúng thì λ sẽ gần bằng 1; trong khi nếu H_1 đúng thì λ sẽ nhỏ hơn.

Đánh giá tiêu chuẩn kiểm định

Bây giờ ta xét một kiểm định giả thuyết đơn H_0 và đối thiết đơn H_1. Kí hiệu $f(x;\theta)$ là hàm mật độ của một biến ngẫu nhiên X trong đó $\theta \in \Theta = \{\theta_0, \theta_1\}$. Cho $\mathbf{X} = (X_1, \ldots, X_n)$ là một mẫu ngẫu nhiên từ phân phối của X.

Định nghĩa

Một tập con C của không gian mẫu được gọi là một *miền bác bỏ tốt nhất* với mức ý nghĩa α cho kiểm định giả thuyết đơn

$$H_0 : \theta = \theta_0 \quad \text{và} \quad H_1 : \theta = \theta_1,$$

nếu $\mathbb{P}_{\theta_0}[X \in C] = \alpha$ và với mọi tập con A của không gian mẫu

$$\mathbb{P}_{\theta_0}[X \in A] = \alpha \ \text{ implies } \ \mathbb{P}_{\theta_1}[X \in C] \geq \mathbb{P}_{\theta_1}[X \in A].$$

Định lý Neyman và Pearson

Định lý

Cho $(X_1, \ldots, X_n)$ *là một mẫu từ một phân phối có hàm mật độ* $f(x;\theta)$. *Khi đó, hàm hợp lý của* $X_1, X_2, \ldots, X_n$ *là*

$$L(\mathbf{x};\theta) = \prod_{i=1}^{n} f(x_i;\theta), \quad \text{for } \mathbf{x} = (x_1, \ldots, x_n).$$

Cho θ_0 *và* θ_1 *là các giá trị cố định phân biệt của* θ *để* $\Theta = \{\theta_0, \theta_1\}$, *và cho* k *là một số dương. Đặt* C *là một tập con của không gian mẫu sao cho*

(a) $\frac{L(\mathbf{x};\theta_0)}{L(\mathbf{x};\theta_1)} \leq k$ *với mỗi* $\mathbf{x} \in C$;

(b) $\frac{L(\mathbf{x};\theta_0)}{L(\mathbf{x};\theta_1)} \geq k$ *với mỗi* $\mathbf{x} \in \mathcal{D} \backslash C$;

(c) $\alpha = \mathbb{P}_{\theta_0}[\mathbf{X} \in C]$.

Khi đó C *là miền bác bỏ tốt nhất với mức ý nghĩa* α *cho kiểm định giả thuyết đơn*

$$H_0 : \theta = \theta_0 \quad \text{đối lập với} \quad H_1 : \theta = \theta_1.$$

Bây giờ ta định nghĩa một miền tiêu chuẩn khi nó tồn tại, mà là miền tiêu chuẩn tốt nhất cho kiểm định một giả thuyết đơn H_0 đối lập với một đối thiết hợp H_1.

Định nghĩa

Miền tiêu chuẩn C là một miền tiêu chuẩn *mạnh đều nhất* (UMP) cỡ α cho kiểm định giả thuyết đơn H_0 đối lập với đối thiết hợp H_1 nếu tập C là một miền tiêu chuẩn tốt nhất cỡ α cho kiểm định H_0 đối lập với mỗi đối thiết đơn trong H_1. Một kiểm định được định nghĩa bởi miền tiêu chuẩn C được gọi là một kiểm định *mạnh đều nhất*, với mức ý nghĩa α, cho kiểm định giả thuyết đơn H_0 đối lập với đối thiết H_1.

Ta đều biết rằng các tiêu chuẩn mạnh đều nhất thường không tồn tại. Tuy nhiên, khi chúng tồn tại, định lý Neyman-Pearson đưa ra một kĩ thuật để tìm chúng.

Trình bày lời giải bài toán KĐGT

Cho dữ liệu điểm kiểm tra của 10 sinh viên như sau:

$$75 \quad 64 \quad 75 \quad 65 \quad 72 \quad 80 \quad 71 \quad 68 \quad 78 \quad 62.$$

Giả sử điểm kiểm tra có phân phối chuẩn với kì vọng μ và phương sai $\sigma^2 = 36$ đã biết, hãy kiểm định các giả thuyết sau với mức ý nghĩa 0.05 và tìm P-value cho mỗi bài toán.

$$H_0 : \mu = 70 \text{ và } H_1 : \mu \neq 70.$$

Trình bày lời giải bài toán KĐGT

Ta sẽ giải mỗi bài toán theo quy trình gồm 6 bước như sau:

Bước 1. Xác định tham số cần xác định là μ.

Bước 2. Xây dựng bài toán kiểm định: $H_0 : \mu = 70$, và $H_1 : \mu \neq 70$.

Bước 3. Xác định cỡ mẫu $n = 10$ và kì vọng mẫu $\overline{X} = 71$.

Bước 4. Mức ý nghĩa $\alpha = 0.05$ suy ra $z_{\alpha/2} = 1.96$.

Bước 5. Xác định thống kê kiểm định

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{71 - 70}{6/\sqrt{10}} = 0.5270.$$

Bước 6. Vì $|Z_0| < z_{\alpha/2}$ nên ta không bác bỏ $H_0 : \mu = 70$ ủng hộ $H_1 : \mu \neq 70$ với mức ý nghĩa $0,05$. Chính xác hơn, ta kết luận điểm trung bình là 70 dựa trên mẫu gồm 10 điểm sinh viên.

Suy ra P-value của kiểm định là $2(1 - \Phi(|Z_0|)) = 2(1 - \Phi(0.5270)) = 0.598$.