TIỂU LUẬN

NHẬP MÔN SUY DIỄN THỐNG KÊ
Kiểm định giả thiết một biến số

Đại học Quốc gia Hà Nội
Đại học Khoa học Tự nhiên
Khoa Toán cơ tin

Giảng viên: Hoàng Phương Thảo

Thành viên nhóm: NGUYỄN MẠNH LINH

Ngày: 1st January 2022

Tóm tắt

Một kiểm định giả thiết trong mô hình Neyman-Pearson là một tiêu chí quyết định cho phép chúng ta lựa chọn giữa 2 giả thiết. Trước khi thực hiện test thống kê, ta định nghĩa giả thiết **null** H_0, được giả định là đúng. Giả thiết được so sánh với đối thiết H_1. Đối thiết H_1 thường được gọi là giả thiết nghiên cứu vì thường về lí thuyết các tham số được chỉ định trong giải thiết thay thế này.
Bài này sẽ nghiên cứu những kiến thức cơ bản về kiểm định giả thiết và các phương pháp kiểm định cho trung bình, độ lệch và tỉ lệ.

Mục lục

1 Giới thiệu

Một kiểm định giả thiết trong mô hình Neyman-Pearson là một tiêu chí quyết định cho phép chúng ta lựa chọn giữa 2 giả thiết. Trước khi thực hiện test thống kê, ta định nghĩa giả thiết **null** H_0, được giả định là đúng. Giả thiết được so sánh với đối thiết H_1. Đối thiết H_1 thường được gọi là giả thiết nghiên cứu vì thường về lí thuyết các tham số được chỉ định trong giải thiết thay thế này.

Các giả thiết có 1 miền xác định tham số trong không gian tham số Θ của các tham số θ. Giả thiết null H_0 được định nghĩa trong miền $[\theta \in \Theta_0]$ và đối thiết H_1 được định nghĩa trong miền $[\theta \in \Theta_1]$ và $\Theta_0 \cup \Theta_1 = \Theta$

1.1 Sai lầm loại I và sai lầm loại II

Khi sử dụng phương pháp kiểm định giả thiết chúng ta luôn có xác suất mắc sai lầm. Ví dụ chúng ta có thể bác bỏ giả thiết null trong khi giả thiết trên thực tế là đúng hoặc ngược lại, giả thiết null trên thực tế là sai mà chúng ta lại chấp nhận nó.
Ta định nghĩa 2 loại sai lầm và sẽ tìm hiểu xác suất mắc phải sai lầm đó trong kiểm định giả thiết thống kê.

Sai lầm loại I	Giả thiết null là đúng nhưng bị bác bỏ
Sai lầm loại II	Giả thiết null là sai nhưng không bị bác bỏ

Cả 2 loại này đều dẫn đến quyết định sai lầm trong kết luận.
Xác suất mắc sai lầm loại I còn được gọi là **mức ý nghĩa**

$$\begin{aligned} \alpha = P(\text{sai lầm loại I}) &= P(\text{bác bỏ } H_0 | H_0 \text{ đúng}) \\ &= P(\text{chấp nhận } H_1 | H_0 \text{ đúng}) \end{aligned} \tag{1.1}$$

Xác suất mắc sai lầm loại II β:

$$\begin{aligned} \beta = P(\text{sai lầm loại II}) &= P(\text{không bác bỏ } H_0 | H_0 \text{ sai}) \\ &= P(\text{chấp nhận } H_0 | H_1 \text{ đúng}) \end{aligned} \tag{1.2}$$

1.2 Power function

Đối thiết $H_1 : \theta \in \Theta_1$, hàm power của kiểm định được định nghĩa như sau:

$$\begin{aligned} Power(\theta) = P(\text{bác bỏ } H_0 | H_0 \text{ sai}) &= P(\text{chấp nhận } H_1 | H_1 \text{ đúng}) \\ &= 1 - \beta(\theta) \end{aligned} \tag{1.3}$$

Với $\beta(\theta)$ là xác suất mắc sai lầm loại II với 1 ước lượng θ cho trước.

1.3 Kiểm định đồng nhất tốt nhất

Trước hết chúng ta cùng xem xét ví dụ sau đây

Ví dụ: Cho một tổng thể với phân bố chuẩn $N(\mu, 1)$, thực hiện 1 phép thử. Kiểm định giả thiết $H_0 : \mu = 1$ với đối thiết $H_1 : \mu = 2$. Tìm mức ý nghĩa và power của kiểm định trong các miền bác bỏ sau:

(a) $(2.036, \infty)$

(b) $(1.100, 1.300) \cup (2.461, \infty)$

(a) $R = (2.036, \infty)$,

$$\alpha = P(X > 2.036|N(1,1)) = P(\frac{X-1}{1} > \frac{2.036-1}{1}) = P(Z > 1.036) = 0.150$$

$$\beta = P(X > 2.036|N(2,1)) = P(\frac{X-2}{1} > \frac{2.036-2}{1}) = P(Z \leqslant 1.036) = 0.514$$

hàm power của test $1 - \beta = 1 - 0.514 = 0.486$

(b) miền bác bỏ $(1.100, 1.300) \cup (2.461, \infty)$, ta có xác suất mắc sai lầm loại I:

$$\begin{aligned}\alpha &= P(1.100 < X < 1.300|N(1,1)) + P(X > 2.461|N(1,1) \\ &= P(0.100 < Z < 0.300) + P(Z > 1.461) = 0.150\end{aligned}$$

xác suất mắc sai lầm loại II:

$$\begin{aligned}\beta &= P(X \leqslant 1.100\ N(2,1)) + P(1.300 \leqslant X \leqslant 2.461|N(2,1)) \\ &= P(Z \leqslant -0.900) + P(-0.700 \leqslant Z \leqslant 0.461) = 0.620\end{aligned}$$

hàm power $1 - \beta = 0.380$ Ta có thể thực hiện tính toán trong R như sau

```
alpha = pnorm(1.3, 1, 1) - pnorm(1.1, 1, 1) + (1 - pnorm
  (2.461, 1, 1))
round(alpha, 3)

beta = pnorm(1.1, 2, 1) + pnorm(2.416, 2, 1) - pnorm(1.3, 2,
  1)
round(beta, 3)
```

Ta nhận thấy rằng với cùng 1 mức ý nghĩa $\alpha = 0.150$ thì hàm power có giá trị khác nhau với 2 miền bác bỏ trong phần a và phần b. Nhìn chung, sẽ tồn tại 1 test được coi là "tốt hơn" (theo nghĩa có giá trị power lớn hơn hay là xác xuất mắc sai lầm loại II nhỏ hơn). Thế nên chúng ta sẽ cùng đi tìm 1 test "tốt nhất" (**uniformly most powerful**).

1.4 p-value hay mức tới hạn

p-value hay mức tới hạn được định nghĩa là một xác suất nhỏ nhất mắc sai lầm loại I với giả định giả thiết null là đúng. Cách tính p-value được cho trong bảng sau:

$H_1 : \theta < \theta_0$	$P(T \leqslant t_{obs} \vert H_0)$
$H_1 : \theta > \theta_0$	$P(T \geqslant t_{obs} \vert H_0)$
$H_1 : \theta \neq \theta_0$	$2min\{P(T \leqslant t_{obs} \vert H_0), P(T \geqslant t_{obs} \vert H_0)\}$

Lưu ý rằng p-value không phải một giá trị cố định mà nó được tính dựa trên mẫu quan sát và với điệu kiện giả thiết H_0 được coi là đúng. Một giá trị p-value nhỏ thường có ý nghĩa ủng hộ giả thuyết H_1. Cho nên, với 1 mức ý nghĩa α cho trước, chúng ta sẽ bác bỏ H_0 nếu giá trị p-value $< \alpha$

1.5 Kiểm định với mức ý nghĩa

Để giải bài toán kiểm định giả thuyết thống kê, ta theo các bước sau:

Bước 1: **Giả thuyết**: Thiết lập giả thuyết null, $H_0 : \theta = \theta_0$. Xác định đối thiết H_1 và với tham số θ, xác định $\theta < \theta_0$ hay $\theta > \theta_0$ hay $\theta \neq \theta_0$.

Bước 2: **Test thống kê**: Lựa chọn test thống kê phù hợp với phân phối của mẫu, hoặc test thống kê chuẩn hóa với giả định H_0 là đúng.

Lựa chọn test thống kê, tham số $\hat{\theta}$ là giá trị kì vọng của tham số ứng với H_0. Ví dụ đối với kiểm định cho trung bình μ thì $\hat{\theta} = \bar{X}$ hoặc với kiểm định cho tỉ lệ π thì $\hat{\theta} = P$.

Đại lượng thống kê chuẩn hóa:

$$T = t(X) = \frac{\hat{\theta}(X) - \theta_0}{\sqrt{Var[\hat{\theta}(X)]}} \tag{1.4}$$

Bước 3: **Tính toán miền bác bỏ**: Ta có thể tính toán miền bác bỏ bằng cách sử dụng mức ý nghĩa α hoặc tính p-value. Sau đó tính toán giá trị $t(X)$ với giả định H_0 đúng. Kí hiệu $t(X) = t_{obs}$

Bước 4: **Kết luận về mặt thống kê**: Sử dụng miền bác bỏ hoặc p-value để đưa ra kết luận về giả thiết null. Nếu t_{obs} rơi vào miền bác bỏ, ta bác bỏ H_0, nếu không, ta không ta không đủ cơ sở bác bỏ H_0. Nếu p-value nhỏ hơn mức ý nghĩa α, bác bỏ H_0, nếu không, ta không ta không đủ cơ sở bác bỏ H_0.

Bước 4: **Kết luận**: Đưa ra kết luận bằng lời.

Ước lượng khoảng và kiểm định thống kê

Khi ước lượng khoảng cho một đại lượng thống kê $\hat{\theta}(X)$ đã biết phân phối với trung bình θ và phương sai $\sigma_{\hat{\theta}(X)}$. Tham số chuẩn hóa có dạng $\frac{\hat{\theta}(X)-\theta}{\sigma_{\hat{\theta}(X)}}$ (với phân phối đã biết, kí hiệu là T). Với khoảng tin cậy $(1-\alpha) * 100\%$, ta có ước lượng khoảng cho $\hat{\theta}(X)$

$$CI_{1-\alpha}(\theta) = [\hat{\theta}(X) + t_{\alpha/2}\sigma_{\hat{\theta}(X)}, \hat{\theta}(X) + t_{1-\alpha/2}\sigma_{\hat{\theta}(X)}] \quad (1.5)$$

Trong kiểm định thống kê, khi chuẩn hóa test thống kê có dạng tương đương với dạng chuẩn hóa của ước lượng khoảng $t_{obs} = \frac{\hat{\theta}(X)-\theta_0}{\sigma_{\hat{\theta}(X)}}$, ước lượng khoảng và vùng chấp nhận của giả thiết null tuân theo cùng một phân phối thì với 1 độ tin cậy $(1-\alpha)*100\%$, giả thiết null cho θ là $H_0 : \theta = \theta_0$ không bị bác bỏ.

Ta có tổng quát qua bảng:

Đối thiết	Miền chấp nhận H_0	Ước lượng khoảng $(1-\alpha)$
$H_1 : \theta < \theta_0$	$t_{obs} \geqslant t_\alpha$	$(-\infty, \hat{\theta}(X) - t_\alpha\sigma_{\hat{\theta}(X)}]$
$H_1 : \theta > \theta_0$	$t_{obs} \leqslant t_\alpha$	$[\hat{\theta}(X) - t_\alpha\sigma_{\hat{\theta}(X)}, \infty)$
$H_1 : \theta \neq \theta_0$	$t_{\alpha/2} \leqslant t_{obs} \leqslant t_{1-\alpha/2}$	$[\hat{\theta}(X) + t_{\alpha/2}\sigma_{\hat{\theta}(X)}, \hat{\theta}(X) + t_{1-\alpha/2}\sigma_{\hat{\theta}(X)}]$

2 Kiểm định giả thiết cho trung bình

2.1 Kiểm định giả thiết cho trung bình với mẫu có phân bố chuẩn và phương sai đã biết

Giả thiết null $H_0 : \mu = \mu_0$
Thực hiện test với mẫu ngẫu nhiên kích thước n, phân bố của $\bar{X}$ là phân bố chuẩn $N(\mu_0, \sigma/\sqrt{n})$ (Giải sử H_0 là đúng).
Định lý giới hạn trung tâm cũng chỉ ra rằng $\bar{X}$ có phân bố chuẩn với cỡ mẫu lớn.
Tham số chuẩn hóa

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

Ta có

Đối thiết	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$	$H_1 : \mu \neq \mu_0$
Miền bác bỏ	$z < z_\alpha$	$z > z_{1-\alpha}$	$\lvert z\rvert > z_{1-\alpha/2}$

Trong đó $\Phi(z_\alpha) = \alpha$
Nhắc lại, trong ngôn ngữ R

```
z.x = qnorm(x)
```

demo code:

```
sigma <- 6
mu <- 40
pnorm(-2)
pnorm(2)
```

2.2 Kiểm định cho trung bình với mẫu có phân bố chuẩn và phương sai chưa biết

Giả thiết null $H_0 : \mu = \mu_0$
Tham số chuẩn hóa được xem xét trong trường hợp này khác với trường hợp đã biết phương sai của tổng thể :

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$$

Ta có:

Đối thiết	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$	$H_1 : \mu \neq \mu_0$
Miền bác bỏ	$t < t_{\alpha;n-1}$	$t > t_{1-\alpha;n-1}$	$\lvert t\rvert > t_{1-\alpha/2;n-1}$

Để tính $t_{x;y}$ trong R ta dùng hàm:

```
t.x.y = qt(x, y)
```

3 Compiling the document

To generate a PDF-file from your TEX-file on your own Latex distribution you need to run the following commands. We assume you have a master file `main.tex` that you want to typeset.

```
pdflatex main
pdflatex main
makeglossaries main
bibtex main
pdflatex main
pdflatex main
```

Listing 1 Commands to compile this document

3.1 Known Issues

Under some configurations on Windows machines, the `makeglossaries` command silently fails, which results in empty lists of accronyms and symbols. Same goes for the implicitly called `makeindex` command.

Phụ lục

A Some Appendix Section

Appendices provide only two structural levels, viz., `\section`, and `\subsection`.

The numbering of figures, listings, tables, and footnotes is not reset. Thus, it continues as usual in the appendix.

A.1 Some Appendix Subsection

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi.

Declaration of Academic Integrity

I hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly attributed. Where I have quoted from the work of others, the source is always given. I am aware that the thesis in digital form can be examined for the use of unauthorised aid and in order to determine whether the thesis as a whole or in parts may amount to plagiarism. I am aware that a false assurance fulfils the elements of fraud in accord with § 10 and § 13 ABMPO/TechFak and will result in the consequences proclaimed there. This paper was not previously presented to another examination board and has not been published.

, Ngày 30 tháng 4 năm 2022

NGUYỄN MẠNH LINH