Kiểm định giả thuyết

Trò chơi tung đồng xu: Giả thuyết của tôi là đồng xu tôi dùng là đồng nhất.

Giả thuyết

Giả thuyết không là một phát biểu về giá trị của thông số. Giả thuyết đối nghịch phát biểu giá trị của thông số khi *giả thuyết không* không đúng.

Giả thuyết không và giả thuyết ngượ c lạ i tạ o nên giớ i hạ n của không gian thông số (đó là nhữ ng giá trị có thể chấp nhận của thông số).

Các ký hiệu đượ c minh họ a bằng nhữ ng giả thuyết về trung bình tổng thể

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_1$$

Kiểm định hai phía là phù hợ p

$$H_0 : \mu \leq \mu_0$$
$$H_1 : \mu > \mu_1$$

Kiểm định một phía là phù hợ p.

Trong mỗi trườ ng hợ p, "chuẩn" đượ c xác định bằng dấu bằng và dấu bằng luôn luôn xuất hiện trong *giả thuyết không*

Chiến Lượ c Kiểm Định Cho Trung Bình Tổng Thể

Chúng ta bắt đầu vớ i giả thiết là tổng thể tuân theo phân phối chuẩn và biết trướ c phương sai của tổng thể.

Để kiểm định giả thuyết không, chúng ta lấy một mẫu ngẫu nhiên từ tổng thể quan tâm và tính giá trị thống kê mẫu, giá trị này là một ướ c lượ ng tốt cho trung bình tổng thể mà chúng ta chưa biết: đó chính là chúng ta cần tính trung bình mẫu.

Nếu giá trị thự c $(\overline{X}^*)$ của trung bình mẫu gần bằng vớ i giá trị giả định của trung bình tổng thể, chúng ta không bác bỏ giả thuyết. Nếu giá trị thự c của trung bình mẫu không gần vớ i giá trị giả định của trung bình tổng thể, chúng ta bác bỏ giả thuyết.

Chúng ta định nghĩa thế nào là "gần" ?

Nếu chúng ta giả sử là giá trị thự c của trung bình tổng thể bằng vớ i chuẩn, chúng ta có thể viết phân phối mẫu của trung bình mẫu: chúng ta gọ i là phân phối mẫu *theo giả thuyết không*

Phân phối mẫu này cho chúng ta biết các giá trị mà trung bình mẫu có thể nhận và nhữ ng giá trị mà trung bình mẫu không thể nhận nếu trung bình tổng thể bằng với giá trị chuẩn đượ c xác định trong *giả thuyết không.* Nếu giá trị thự c của trung bình mẫu là một giá trị "không thể" nhận thì chúng ta có thể xem xét để bác bỏ *giả thuyết không*.

Có ba cách khác nhau để tiến hành kiểm định. Chúng ta có thể gọ i các phương pháp này là:

- truyền thống
- Kiểm định Z
- trị thống kê p-value

Ví dụ 1: Truyền thống

Giả sử tổng thể tuân theo phân phối chuẩn với trung bình chưa biết và phương sai là $\sigma^2 = 36$:

$$X \sim N\left(\mu, \sigma^2 = 36\right)$$

Giả thuyết không và giả thuyết ngượ c lạ i là

$$H_0 : \mu \le 100$$
$$H_1 : \mu > 100$$

Lấy một mẫu ngẫu nhiên có kích thước là $n = 9$ và tính trung bình mẫu $\overline{X}$.

Theo *giả thuyết không,* phân phối mẫu của $\overline{X}$ là $\overline{X} \sim N\left(100, \sigma_{\overline{X}}^2 = 4\right)$.

Nếu giá trị thự c của trung bình mẫu khá xa về phía bên phải của phân phối mẫu thì kết quả này có thể dẫn đến việc chúng ta bác bỏ *giả thuyết không*

Xác suất mà một nhà điều tra xem là "không thể" có lẽ tùy sở thích, nhưng thường các giá trị này là 1%, 5% và 10%.

Xác suất đượ c dùng cho mụ c đích này đượ c gọ i là "mức ý nghĩa" của kiểm định; α là ký hiệu qui ước cho mức ý nghĩa.

Phần đầu của phân phối mẫu của trung bình mẫu đượ c gọ i là "giá trị ngưỡ ng" của kiểm định; $\overline{X}_{(1-\alpha)}$ là ký hiệu qui ước của giá trị ngưỡ ng trong kiểm định phía bên phải.

Nếu giá trị thự c của trung bình mẫu vượ t qua giá trị ngưỡ ng, chúng ta bác bỏ *giả thuyết không*

Tiếp tụ c ví dụ này, chúng ta chọ n một mức ý nghĩa $\alpha = 0.05$, sau đó chúng ta trình bày dưới dạ ng xác suất

$$P\left(\overline{X}_{(1-\alpha)} \le \overline{X} \mid \mu = 100\right) = 0.05$$

Chuẩn hóa chúng ta có

$$P\left(\frac{\overline{X}_{(1-\alpha)}-\mu}{\sigma/\sqrt{n}} \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \mid \mu = 100\right) = P\left(Z_{(1-\alpha)} < Z\right) = 0.05$$

Sử dụng máy tính hoặc các bảng thống kê thích hợp, chúng ta tìm được

$$Z_{(1-\alpha)} = Z_{(0.95)} = 1.96 = \frac{\overline{X}_{(1-\alpha)}-\mu}{\sigma/\sqrt{n}}$$

Giải phương trình này chúng ta tìm được giá trị ngưỡng

$$\overline{X}_{(1-\alpha)} = \mu_0 + Z_{(1-\alpha)}\left(\sigma/\sqrt{n}\right) = 103.92$$

Nguyên tắc quyết định của chúng ta là: Nếu giá trị thực của trị thống kê mẫu $\overline{X}^*$ lớn hơn 103.02 chúng ta bác bỏ *giả thuyết không*. Ngược lại chúng ta không thể bác bỏ giả thuyết này.

Ví dụ 1: trị thống kê Z

Xem lại phát biểu dưới dạng xác suất chúng ta đã vận dụng để tìm giá trị ngưỡng của trung bình mẫu $\overline{X}$:

$$P\left(\frac{\overline{X}_{(1-\alpha)}-\mu}{\sigma/\sqrt{n}} \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \mid \mu = 100\right) = P\left(Z_{(1-\alpha)} \le Z\right) = 0.05$$

Giá trị thực của trị thống kê Z (Z^*) có thể được so sánh với $Z_{(1-\alpha)}$ để kiểm định.

$$Z^* = \frac{\overline{X}^*-\mu}{\sigma/\sqrt{n}}$$

và nguyên tắc quyết định được phát biểu lại như sau:

Nếu Z^* vượt quá giá trị ngưỡng $Z_{(1-\alpha)} = 1.96$ thì bác bỏ *giả thuyết không*. Ngược lại, không thể bác bỏ giả thuyết này.

Ví dụ 1: giá trị p

Đánh giá phát biểu dạng xác suất sau:

$$P\left(\overline{X}^* < \overline{X}\right) = P\left(Z^* < Z\right) = p\text{-value}$$

Nếu $P\left(\overline{X}_{(1-\alpha)} < \overline{X}^*\right)$ thì $P\left(Z_{(1-\alpha)} < Z^*\right)$ và $p\text{-value} < \alpha$

Chúng ta có thể phát biểu lại qui luật quyết định như sau:

Nếu giá trị p nhỏ hơn mức ý nghĩa (α) chọn trước cho kiểm định thì chúng ta bác bỏ *giả thuyết không*. Ngược lại chúng ta không bác bỏ được giả thuyết này.

Ba cách kiểm định này luôn luôn đem lại kết quả như nhau, vì thật ra đây chỉ đơn giản là phát biểu cùng một thông tin theo những cách khác nhau. Phát biểu qui luật quyết định theo giá trị p thường thuận tiện nhất và Eviews báo cáo giá trị p cho hầu hết các kiểm định thống kê chúng ta quan tâm.

Khi chúng ta bác bỏ *giả thuyết không,* chúng ta thường nói là"trung bình mẫu không khác với giá trị giả định một cách có ý nghĩa về mặt thống kê" Ở đây xuất hiện hai thuật ngữ "có ý nghĩa về mặt thống kê" và "mức ý nghĩa"

Ví dụ 2: Kiểm định hai phía

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_1$$

Trong trường hợp này chúng ta muốn bác bỏ *giả thuyết không* nếu trung bình mẫu nằm quá xa về phía hai đầu. Điều này có nghĩa là chúng ta chia phân nửa mức ý nghĩa cho mỗi phía và có phát biểu dạng xác suất như sau:

$$P\left(\frac{\overline{X}_{(\alpha/2)} - \mu}{\sigma/\sqrt{n}} \leq \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{\overline{X}_{(1-\alpha/2)} - \mu}{\sigma/\sqrt{n}} \quad | \mu = 100\right) = P\left(Z_{(\alpha/2)} \leq Z \leq Z_{(1-\alpha/2)}\right) = 1 - \alpha$$

Qui luật quyết định được phát biểu dưới một trong ba cách sau:

Nếu $\overline{X}^* < \overline{X}_{(\alpha/2)}$ hay $\overline{X}^* > \overline{X}_{(1-\alpha/2)}$ thì bác bỏ *giả thuyết không*.

Nếu $Z^* < Z_{(\alpha/2)}$ hay $Z^* > Z_{(1-\alpha/2)}$ thì bác bỏ *giả thuyết không*.

Nếu p-value < α thì bác bỏ *giả thuyết không*.

Trong trường hợp hai phía, chúng ta tính giá trị p như sau: $2 \times P\left(|Z^*| < Z\right)$

Chúng ta nhận thấy giả định là phương sai của tổng thể đã biết có thể là không thực tế. Trong trường hợp này chúng ta giữ nguyên giả định tổng thể phân phối chuẩn và dùng ước lượng phương sai của mẫu.

Ví dụ 3: kiểm định t

Giả thuyết cần có là tổng thể tuân theo phân phối chuẩn. Chiến lược kiểm định hoàn toàn giống như cách thực hiện trên đây. Chỉ có một khác biệt là phát biểu dạng xác suất căn cứ trên biến ngẫu nhiên t với (n – 1) bậc tự do.

$$\frac{\overline{X} - \mu}{s/\sqrt{n}} = t\text{-Stat} \sim t_{(n-1)}$$

Xem bài tập phần 6 trong ví dụ

Tổng thể có phân phối không chuẩn

Nếu tổng thể không tuân theo phân phối chuẩn, chúng ta phải áp dụng CLT. Khi kích thước mẫu đủ lớn thì trị thống kê t định nghĩa như trên có phân phối gần với phân phối chuẩn và giá trị thực t* có thể so sánh được với giá trị ngưỡng của Z được lấy từ bảng phân phối chuẩn.

Kiểm Định Giả Thuyết Về Phương Sai

Trong bài tập cuối cùng của Bài tập phần 5, chúng ta đã đưa ra một giả định là phương sai của tổng thể là một giá trị cụ thể. Sau đó chúng ta tìm một ước lượng của phương sai tổng thể. Giá trị ước lượng khác với giá trị giả định. Một thắc mắc thú vị là sự khác biệt này có ý nghĩa thống kê hay không.

Kiểm định sau đây yêu cầu phân phối của tổng thể phải là phân phối chuẩn.

Hãy xem xét các *giả thuyết không* và giả thuyết ngược lại sau:

$$H_0 : \sigma^2 = \sigma_0^2$$
$$H_1 : \sigma^2 \neq \sigma_0^2$$

Bây giờ xem xét phân phối mẫu của kiểm định thống kê sau theo *giả thuyết không*:

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

Nếu giả thuyết là đúng, thì kỳ vọng của trị thống kê là (n-1) và nếu giả thuyết là sai thì chúng ta kỳ vọng trị thống kê rơi vào một trong hai đầu của phân phối.

Phát biểu dạng xác suất thích hợp là:

$$P\left(\chi^2_{\left(n-1, \alpha/2\right)} \leq (n-1)\frac{s^2}{\sigma_0^2} \leq \chi^2_{\left(n-1, 1-\alpha/2\right)} \right) = 1-\alpha$$

Qui luật quyết định rất rõ ràng, và cách thực hiện xuất hiện trong bài tập phần 6

Kiểm Định Hai Phương Sai Có Bằng Nhau Không

Hãy tưởng tượng chúng ta quan tâm đến việc so sánh hai tổng thể. Cách chúng ta so sánh giá trị trung bình của hai tổng thể này phụ thuộc vào phương sai của hai tổng thể có bằng nhau hay không.

Giả thuyết cần là cả hai tổng thể đều tuân theo phân phối chuẩn.

Xem xét *giả thuyết không* và giả thuyết ngược lại sau:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

Hãy tưởng tượng chúng ta có các mẫu từ cả hai tổng thể và xem xét phân phối mẫu của kiểm định thống kê sau theo *giả thuyết không*:

$$\frac{(n_1-1)\frac{s_1^2}{\sigma^2} \Big/ (n_1-1)}{(n_2-1)\frac{s_2^2}{\sigma^2} \Big/ (n_2-1)} \sim F_{((n_1-1),(n_2-1))}$$

Theo *giả thuyết không,* đơn giản phân số ta có

$$\frac{s_1^2}{s_2^2} \sim F_{((n_1-1),(n_2-1))}$$

Phát biểu dạ ng xác suất tương ứng là:

$$P\left(F_{\left(n_1-1, n_2-1, \alpha/2\right)} \le \frac{s_1^2}{s_2^2} \le F_{\left(n_1-1, n_2-1, 1-\alpha/2\right)} \right) = 1-\alpha$$

Và một lần nữ a qui luật ra quyết định cũ ng rất rõ ràng.

Kiểm Định Hai Trung Bình Bằng Nhau Hay Không

Xét các giả thuyết sau:

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

Viết lạ i cho thuận tiện hơn:

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 \neq 0$$

Vớ i giả thuyết là các phương sai tổng thể bằng nhau và tổng thể có phân phối chuẩn, kiểm định thống kê tương ứng là:

$$\frac{\overline{X}_1 - \overline{X}_2}{s_{\left(\overline{X}_1 - \overline{X}_2\right)}} = T\text{-Stat}$$

Khó khăn là phải ướ c lượ ng mẫu số.

Nếu phương sai như nhau thì mẫu số có thể đượ c ướ c lượ ng dự a trên phương sai "gộp":

$$s_p^2 = \frac{\sum\left(X_{1i} - \overline{X}_1\right)^2 + \sum\left(X_{2i} - \overline{X}_2\right)^2}{n_1 + n_2 - 2}$$

Phương sai gộp là một ướ c lượ ng của phương sai chung cho cả hai tổng thể.

Phương sai của hiệu $\overline{X}_1 - \overline{X}_2$ là tổng của các phương sai (vì đây là các mẫu ngẫu nhiên độc lập) vậy phương sai bằng:

$$s^2_{\left(\overline{X}_1 - \overline{X}_2\right)} = \frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}$$

Thống kê T trên đây tuân theo phân phối Chi bình phương với $\left(n_1 + n_2 - 2\right)$ bậc tự do

Nếu các phương sai khác nhau, bước tính phương sai chung có thể bỏ qua, và ước lượng phương sai của hiệu số là:

$$s^2_{\left(\bar{x}_1 - \bar{x}_2\right)} = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$$

Tuy nhiên trong trường hợp này, chúng ta phải có cỡ mẫu lớn và kiểm định thống kê là gần với phân phối chuẩn:

$$\frac{\bar{X}_1 - \bar{X}_2}{s_{\left(\bar{x}_1 - \bar{x}_2\right)}} = Z\text{-stat} \sim \text{approx } N(0,1)$$