x_i	195	200	205	
n_i	5	18	2	
$x_i n_i$	975	3600	410	$\overline{x} = \frac{4985}{25} = 199,4$
$(x_i - 200)^2$	25	0	25	
$(x_i - \overline{x})^2$	19,36	0,36	31,36	

độ tin cậy 95%, qua mẫu cụ thể này, khoảng tin cậy của σ^2 là

$$\left(\frac{175}{40,646}; \frac{175}{13,12}\right) \approx (4,305; 13,338).$$

b) Ta có $(n-1)s^2 = \sum_{i=1}^{3}(x_i - 199,4)^2 n_i = 166$. Tra Bảng III phần Phụ lục ta có $\chi^2_{0,025}(24) = 39,364$ và $\chi^2_{0,975}(24) = 12,401$. Khoảng tin cậy cho σ^2 là

$$\left(\frac{166}{39,364}; \frac{166}{12,401}\right) \approx (4,217; 13,386).$$

3.5. KIỂM ĐỊNH GIẢ THUYẾT THỐNG KÊ

Trong phần này, chúng ta sẽ bàn đến những phương pháp thống kê dùng để trả lời những câu hỏi dạng "có hay không một hiện tượng hay hiệu ứng nào đó". Ví dụ: dữ liệu của trang web A có được bảo mật tốt không?, loại thuốc này có hiệu nghiệm không?, chất thải của nhà máy này có làm hại sức khỏe của nhân dân xung quanh không?, đạo đức của thanh niên có thực sự phụ thuộc vào hoàn cảnh gia đình?, học viên của Tiểu đoàn A học môn Thống kê tốt hơn học viên của Tiểu đoàn B?... Mỗi tình huống "có hay không" như vậy thường được coi là các *giả thuyết*, chúng có thể đúng và cũng có thể sai. Vấn đề xác định đúng sai của một giả thuyết sẽ được gọi là *kiểm định*.

Trong thống kê chúng ta xuất phát từ một mẫu $X_1, X_2, \cdots, X_n$ chọn từ một tổng thể chưa biết phân phối hoặc có phân phối nhưng chưa biết tham số θ. Ta có thể phát biểu nhiều nhận xét khác nhau về các yếu tố chưa biết, đó là các giả thuyết thống kê. Giả thuyết được đưa ra kiểm định được gọi là *giả thuyết gốc*, kí hiệu là H_0. Các giả thuyết khác với giả thuyết gốc được gọi là *giả thuyết đối* hay *đối thuyết*, kí hiệu là H_1. Ta thừa nhận khi đã chọn cặp H_0, H_1 thì việc chấp nhận H_0 sẽ chính là bác bỏ H_1 và ngược lại. Việc kiểm định một giả thuyết là đúng hay sai dựa trên thông tin mẫu sẽ được gọi là *kiểm định thống kê*.

Cần chú ý rằng đối thuyết H_1 không nhất thiết là phủ định của giả thuyết H_0. Chẳng hạn giả thuyết H_0: "thu nhập bình quân của cư dân một thành phố nào đó trong 1 năm là 80 triệu đồng", nếu ta nghi ngờ rằng thu nhập này không đúng thì đối thuyết H_1 là khác 80 triệu đồng, > 80 triệu đồng hoặc < 80 triệu đồng, ...

Lí thuyết kiểm định giả thuyết thống kê có nhiều ứng dụng trong thực tế, giúp các nhà quản lí kiểm tra tính đúng đắn của các quyết định.

Có một điều mà người học cần hết sức chú ý. Đó là, mỗi phương pháp kiểm định bằng thống kê chỉ thích hợp trong những tình huống nhất định, khi các giả

sử nhất định được thoả mãn. Với một vấn đề kiểm định thống kê trong thực tế cần thực hiện, thì phải chọn lựa phương pháp đúng đắn, và rất có thể là phương pháp mà người học cần đến không nằm trong giáo trình này, do đó bạn đọc sẽ phải tìm hiểu sâu thêm về thống kê để chọn lựa được phương pháp thích hợp cho vấn đề của mình.

3.5.1. Các nguyên tắc chung của kiểm định giả thuyết thống kê

Nguyên tắc chung của kiểm định giả thuyết thống kê dựa trên hai nguyên lí sau:

- *Nguyên lí xác suất nhỏ* : nếu một sự kiện có xác suất xuất hiện rất nhỏ thì có thể coi rằng nó không xảy ra khi thực hiện một phép thử có liên quan đến sự kiện đó.

- *Phương pháp phản chứng*: nếu từ giả thuyết H_0 đúng dẫn đến một điều vô lí thì ta bác bỏ H_0 (chấp nhận đối thuyết H_1).

Dựa vào hai nguyên lí này ta đưa ra phương pháp chung để kiểm định một giả thuyết thống kê như sau: Để kiểm định H_0, trước hết giả sử H_0 đúng, từ đó ta tìm được sự kiện A mà xác suất xuất hiện sự kiện A là rất nhỏ và ta có thể xem A không thể xảy ra trong một phép thử về sự kiện này. Lúc đó nếu trên một mẫu cụ thể quan sát được mà sự kiện A xuất hiện thì điều này trái với nguyên lí xác suất nhỏ. Vậy H_0 sai và bác bỏ nó; ngược lại, nếu A không xảy ra thì ta chưa có cơ sở để bác bỏ H_0.

Ta thực hiện phương pháp trên bằng các bước cụ thể sau:

a) Tiêu chuẩn kiểm định giả thuyết thống kê

Từ BNN gốc X của tổng thể lập mẫu ngẫu nhiên $X_1, \cdots, X_n$, chọn thống kê $T = T(X_1, \cdots, X_n)$ có thể phụ thuộc vào tham số đã biết trong giả thuyết H_0. Nếu giả thuyết H_0 đúng thì luật phân phối của T phải hoàn toàn xác định. Một thống kê như vậy được gọi là *tiêu chuẩn kiểm định.*

b) Quy tắc kiểm định

Nếu ta thành công trong việc chia miền xác định của tiêu chuẩn kiểm định T thành hai phần R_α và $\overline{R_\alpha}$ trong đó R_α là *miền bác bỏ* H_0, còn $\overline{R_\alpha}$ là *miền chấp nhận* H_0, thì quy tắc kiểm định khá đơn giản: với α rất nhỏ cho trước (thường $\alpha \in [0,01; 0,05]$) và với điều kiện H_0 đúng, nếu

$$P(T \in R_\alpha | H_0) = \alpha \qquad (3.13)$$

thì ta bác bỏ H_0; nếu ngược lại, ta chấp nhận H_0. Giá trị α được gọi là *mức ý nghĩa* của kiểm định.

So sánh giá trị của tiêu chuẩn kiểm định T với miền bác bỏ R_α và kết luận theo quy tắc sau:

(i) Nếu $T \in R_\alpha$, theo nguyên tắc kiểm định thì H_0 sai, do đó ta bác bỏ H_0, thừa nhận H_1.

(ii) Nếu $T \notin R_\alpha$ thì điều này chưa khẳng định rằng H_0 đúng mà chỉ có nghĩa là qua mẫu cụ thể này chưa khẳng định được H_0 sai. Do đó ta chỉ có thể nói rằng qua mẫu cụ thể này chưa có cơ sở để bác bỏ H_0 (trên thực tế là thừa nhận H_0).

c) Sai lầm loại I và sai lầm loại II

Với quy tắc kiểm định như trên có thể mắc hai loại sai lầm sau đây:

(i) Sai lầm loại I: bác bỏ một giả thuyết đúng. Từ *(3.13)*, ta thấy xác suất mắc sai lầm loại I đúng bằng mức ý nghĩa α. Sai lầm loại I sinh ra do kích thước mẫu quá nhỏ, do phương pháp lấy mẫu...

(ii) Sai lầm loại II: chấp nhận một giả thuyết sai. Xác suất sai lầm loại II là β xác định như sau:

$$P(T \notin R_\alpha | H_1) = \beta. \qquad (3.14)$$

Ta gọi $P(T \in R_\alpha | H_1) = 1 - \beta$ là lực lượng kiểm định của tiêu chuẩn T.

Cả hai loại sai lầm đều có thể gây ra những hậu quả không tốt, chúng ta mong muốn cả hai xác suất *(3.13)* và *(3.14)* càng nhỏ càng tốt. Trong thực tế ta không thể đồng thời làm giảm cả hai xác suất đó, tức là muốn α giảm thì β tăng và ngược lại. Vì muốn hạn chế sai lầm loại I, ta có xu hướng dè dặt trong việc bác bỏ và sẽ có khuynh hướng dễ dãi trong việc chấp nhận, khi đó lại dễ phạm sai lầm loại II. Còn muốn giảm sai lầm loại II, ta dè dặt trong việc chấp nhận và dẫn đến dễ dãi trong việc bác bỏ. Điều này làm cho nguy cơ phạm sai lầm loại I tăng lên. (Tất nhiên có một cách làm giảm cả hai xác suất sai lầm nếu tăng kích thước mẫu n lên. Nhưng khi đó chi phí cũng tăng lên và đôi khi ta không trực tiếp làm ra được số liệu).

Thực tế / Quyết định	H_0 đúng	H_0 sai
Bác bỏ H_0	Sai lầm loại I, α	Quyết định đúng, $1 - \beta$
Không bác bỏ H_0	Quyết định đúng, $1 - \alpha$	Sai lầm loại II, β

Giải quyết mâu thuẫn này bằng cách nào?

Thực ra sai lầm loại I và loại II rất tương đối, nó không có sẵn từ đầu, mà chỉ xác định khi ta đã đặt giả thuyết. Chẳng hạn đối với một bác sĩ khám bệnh, ông ta có thể phạm phải một trong hai tình huống sai lầm sau:

(A) Người có bệnh, sau khi xét nghiệm, ông kết luận không có bệnh.

(B) Người không bệnh, sau khi xét nghiệm, ông kết luận: nhập viện!

- *Đâu là sai lầm loại I?* Tất nhiên là chưa thể trả lời được. Nếu bác sĩ đặt giả thuyết H_0: "người này có bệnh" thì trường hợp (A) là sai lầm loại I và (B) là sai lầm loại II. Còn nếu bác sĩ đặt giả thuyết H_0: "người này không bệnh" thì trường hợp (A) là sai lầm loại II còn (B) là sai lầm loại I.

- *Nên đặt giả thuyết thế nào?* Muốn vậy, ta phải xem xét sai lầm nào quan trọng hơn, tức là khi phạm phải sẽ chịu tổn thất lớn hơn, thì ta sẽ đặt bài toán để sai lầm đó là loại I. Với trường hợp trên, ta chọn đặt giả thuyết H_0: "người này có bệnh".

Sau khi đã đặt bài toán và xác định được sai lầm loại I, các nhà thống kê đưa ra phương pháp sau: cố định xác suất sai lầm loại I không vượt quá mức ý nghĩa α nhỏ (không đáng kể) nào đó, với mẫu kích thước n xác định, ta chọn ra miền bác bỏ R_α sao cho xác suất sai lầm loại II β là nhỏ nhất hoặc có thể chấp nhận được.

Việc chọn mức ý nghĩa α bằng bao nhiêu tùy thuộc vào từng trường hợp cụ thể, tùy thuộc vào ý nghĩa của bài toán.Với mỗi cặp giả thuyết và đối thuyết đã cho, không phải lúc nào cũng tồn tại hoặc tìm được một quy tắc sao cho lực lượng của phép kiểm định $1-\beta$ là lớn nhất. Những quy tắc đưa ra trong giáo trình này là những quy tắc thông dụng.

d) Thủ tục kiểm định giả thuyết thống kê

Qua nội dung trình bày ở trên ta có thể xây dựng một thủ tục kiểm định giả thuyết thống kê bao gồm các bước sau:

(i) Phát biểu giả thuyết H_0 và đối thuyết H_1.

(ii) Từ tổng thể nghiên cứu lập mẫu ngẫu nhiên kích thước n.

(iii) Chọn tiêu chuẩn kiểm định T và xác định quy luật phân phối xác suất của T với điều kiện giả thuyết H_0 đúng.

(iv) Dựa vào luật phân phối xác suất của T, tìm miền bác bỏ R_α sao cho

$$P(T \in R_\alpha | H_0) = \alpha.$$

(v) Dựa vào mẫu cụ thể kích thước n, tính các thông số của mẫu cần thiết, thay thế vào thống kê T tính được giá trị T_0 và gọi là giá trị quan sát thực tế hay giá trị thực nghiệm của thống kê T tương ứng với mẫu.

(vi) So sánh giá trị quan sát T_0 của tiêu chuẩn kiểm định T với miền bác bỏ R_α và kết luận.

Ví dụ 3.14. Tung một đồng tiền 20 lần, ra 2 lần mặt sấp và 18 lần mặt ngửa. Có thể coi đồng tiền là cân bằng (hai mặt sấp và ngửa đều có xác suất 50%) không? Giả thuyết H_0 là "đồng tiền cân bằng".

Lời giải. Gọi BNN X là "số lần hiện mặt sấp trong 20 lần tung". Khi đó

$$P(X \leq 2 | H_0) = \frac{C_{20}^0 + C_{20}^1 + C_{20}^2}{2^{20}} \approx 0,02\%.$$

Giá trị này quá nhỏ (thông thường $\alpha \in [0,01; 0,05]$) để chấp nhận giả thuyết H_0. Vậy không thể coi đồng tiền là cân bằng.

3.5.2. Các kiểm định dùng một mẫu

a) Kiểm định về kì vọng của tổng thể có phân phối chuẩn

Giả sử BNN của tổng thể là $X \sim \mathcal{N}(\mu; \sigma^2)$ và mẫu ngẫu nhiên $(X_1, \cdots, X_n)$. Bài toán đặt ra là với mức ý nghĩa α cho trước hãy kiểm định giả thuyết $H_0 : \mu = \mu_0$ (với μ_0 cho trước).

(i) Trường hợp phương sai σ^2 đã biết

Chọn tiêu chuẩn kiểm định $T = \dfrac{\overline{X} - \mu_0}{\sigma}\sqrt{n}$. Từ giả thiết phân phối chuẩn của X (hoặc theo Định lí 2.6) ta có $T \sim \mathcal{N}(0; 1)$. Ta xây dựng các miền bác bỏ dựa vào đối thuyết H_1 như sau:

- *Kiểm định hai phía*: khi $H_1 : \mu \neq \mu_0$, từ *(2.9)* ta có $P(|T| > U_{\alpha/2} | H_0) = \alpha$, do đó miền bác bỏ là

$$R_\alpha = (-\infty; -U_{\alpha/2}) \cup (U_{\alpha/2}; +\infty).$$

- *Kiểm định một phía*:

• Nếu $H_1 : \mu > \mu_0$, từ *(2.9)* ta có $P(T > U_\alpha | H_0) = \alpha$, do đó miền bác bỏ là $R_\alpha = (U_\alpha; +\infty)$;

• Nếu $H_1 : \mu < \mu_0$, từ khái niệm hàm phân phối và Định nghĩa 2.19 ta có

$$P(T < -U_\alpha | H_0) = \Phi(-U_\alpha) = 1 - \Phi(U_\alpha) = \alpha,$$

do đó miền bác bỏ là $(-\infty; -U_\alpha)$.

Với mẫu cụ thể, ta tính được giá trị quan sát của tiêu chuẩn kiểm định là $T_0 = \dfrac{\overline{x} - \mu_0}{\sigma}\sqrt{n}$ và so sánh với miền bác bỏ R_α để kết luận.

Ví dụ 3.15. Một hãng bảo hiểm thông báo "số tiền trung bình hãng chi trả cho khách hàng bị tai nạn ôtô là 170 triệu đồng". Để kiểm tra lại, người ta kiểm tra ngẫu nhiên hồ sơ chi trả của 25 trường hợp thì thấy trung bình mẫu là 180 triệu đồng. Giả sử số tiền chi trả tuân theo luật phân phối chuẩn với $\sigma = 50$ triệu đồng, hãy kiểm định lại thông báo của hãng bảo hiểm trên với $\alpha = 0,05$.

Lời giải. Gọi μ là số tiền trung bình hãng bảo hiểm chi trả cho khách hàng bị tai nạn ôtô.

Ta kiểm định: Giả thuyết $H_0 : \mu = 170$, đối thuyết $H_1 : \mu \neq 170$.

Tiêu chuẩn kiểm định $T = \dfrac{\overline{X} - 170}{\sigma}\sqrt{n}$. Với $\alpha = 0,05$ thì $U_{0,025} = 1,96$.

Với mẫu đã cho: $\overline{x} = 180$, $n = 25$, ta có giá trị quan sát thực tế là

$$T_0 = \frac{180 - 170}{50}\sqrt{25} = 1 < 1,96 = U_{0,025}$$

nên ta không có cơ sở để bác bỏ thông báo của hãng bảo hiểm.

(ii) Trường hợp phương sai σ^2 chưa biết và kích thước mẫu $n \geq 30$

Khi kích thước mẫu đủ lớn ($n \geq 30$) thì độ lệch chuẩn σ có thể được xấp xỉ bởi độ lệch chuẩn mẫu hiệu chỉnh S. Hơn nữa, theo Định lí giới hạn trung tâm 2.6 thì thống kê $\dfrac{\overline{X} - \mu}{\sigma}\sqrt{n} \sim \mathcal{N}(0;1)$. Ta xây dựng các miền bác bỏ dựa vào đối thuyết H_1 tương tự như trong trường hợp *(i)*, tức là:

- *Kiểm định hai phía*: khi $H_1 : \mu \neq \mu_0$ thì miền bác bỏ là

$$R_\alpha = (-\infty; -U_{\alpha/2}) \cup (U_{\alpha/2}; +\infty).$$

- *Kiểm định một phía*:

• Nếu $H_1 : \mu > \mu_0$ thì miền bác bỏ là $R_\alpha = (U_\alpha; +\infty)$;

• Nếu $H_1 : \mu < \mu_0$ thì miền bác bỏ là $R_\alpha = (-\infty; -U_\alpha)$.

Với mẫu cụ thể, ta tính được giá trị quan sát của tiêu chuẩn kiểm định là $T_0 = \dfrac{\overline{x} - \mu_0}{s}\sqrt{n}$ và so sánh với miền bác bỏ R_α để kết luận.

(iii) Trường hợp phương sai σ^2 chưa biết và kích thước mẫu $n < 30$

Từ *(3.10)*, thống kê $T = \dfrac{\overline{X} - \mu}{S}\sqrt{n} \sim T(n-1)$. Do đó, sử dụng tính chất *(iii)* của phân phối Student, ta xây dựng các miền bác bỏ theo đối thuyết H_1 như sau:

- *Kiểm định hai phía*: khi $H_1 : \mu \neq \mu_0$, ta có $P(|T| > t_{\alpha/2}(n-1)|H_0) = \alpha$, do đó miền bác bỏ là

$$R_\alpha = \big(-\infty; -t_{\alpha/2}(n-1)\big) \cup \big(t_{\alpha/2}(n-1); +\infty\big).$$

- *Kiểm định một phía*:

• Nếu $H_1 : \mu > \mu_0$, ta có $P\big(T > t_\alpha(n-1)|H_0\big) = \alpha$, do đó miền bác bỏ là $R_\alpha = \big(t_\alpha(n-1); +\infty\big)$;

• Nếu $H_1 : \mu < \mu_0$, ta có $P\big(T < -t_\alpha(n-1)|H_0\big) = \alpha$, do đó miền bác bỏ là $R_\alpha = \big(-\infty; -t_\alpha(n-1)\big)$.

Với mẫu cụ thể, ta tính được giá trị quan sát của tiêu chuẩn kiểm định là $T_0 = \dfrac{\overline{x} - \mu_0}{s}\sqrt{n}$ và so sánh với miền bác bỏ R_α để kết luận.

Ví dụ 3.16. Độ dài chi tiết một máy thông tin là BNN X có luật phân phối chuẩn. Kiểm tra 28 sản phẩm thu được số liệu như sau: (đơn vị tính cm)

20,04	20,04	19,99	20,01	20,00	20,02	20,01	20,02	20,04	20,02
20,00	20,02	19,99	19,97	20,02	19,99	19,99	19,97	20,00	20,04
19,97	20,00	20,00	20,02	20,01	19,97	20,00	20,01	20,04	19,99

Với độ tin cậy 95%, có thể cho rằng trung bình độ dài chi tiết máy này bằng 20cm hay không?

Lời giải. Từ những số liệu trên ta có Bảng tính

x_i	n_i	$x_i n_i$	$(x_i - \overline{x})^2 n_i$
19,97	4	79,98	528.10^{-5}
19,99	5	99.95	133.10^{-5}
20,00	6	120,00	24.10^{-5}
20,01	4	80,04	5.10^{-5}
20,02	6	120,12	112.10^{-5}
20,04	5	100,20	567.10^{-5}
		$\overline{x} = \frac{600,19}{30} \approx 20,01$	$s^2 = \frac{0,0137}{29} \approx 47.10^{-5}$

Gọi μ là độ dài trung bình của chi tiết máy thông tin. Ta kiểm định: Giả thuyết $H_0 : \mu = 20$, đối thuyết $H_1 : \mu \neq 20$.

Tiêu chuẩn kiểm định $T = \dfrac{\overline{X} - 20}{S}\sqrt{n}$. Với độ tin cậy 95% thì $\alpha = 0,05$, tra Bảng IV phần Phụ lục ta có $t_{0,025}(29) = 2,045$.

Với mẫu đã cho: $\overline{x} = 20,01$; $n = 30$; $s = 0,022$ ta có giá trị quan sát thực tế là

$$T_0 = \frac{20,01 - 20}{0,022}\sqrt{30} = 1,596 < 2,045 = t_{0,025}(29)$$

nên ta chấp nhận độ dài trung bình của chi tiết máy thông tin này là 20cm.

b) Kiểm định về tỉ lệ của tổng thể có phân phối Bernoulli

Bài toán kiểm định về tỉ lệ: với mức ý nghĩa α, hãy xác định giả thuyết $H_0 : p = p_0$ biết p là tham số của phân phối Bernoulli $\mathcal{B}(1; p)$.

Trong bài toán khoảng tin cậy cho tỉ lệ, ta đã biết nếu kích thước mẫu n đủ lớn (tức là $n.p \geq 5$ và $n.(1-p) \geq 5$) thì phân phối chuẩn có thể được dùng xấp xỉ phân phối nhị thức $\mathcal{B}(n; p)$, khi đó tần suất mẫu $f \sim \mathcal{N}\left(p; \dfrac{p(1-p)}{n}\right)$ và tiêu chuẩn

kiểm định $T = \dfrac{f-p}{\sqrt{p(1-p)}}\sqrt{n} \sim \mathcal{N}(0;1)$. Từ đó bài toán kiểm định về tỉ lệ không có khác biệt căn bản so với kiểm định về kì vọng.

- *Kiểm định hai phía*: khi $H_1 : p \neq p_0$ thì miền bác bỏ là

$$R_\alpha = (-\infty; -U_{\alpha/2}) \cup (U_{\alpha/2}; +\infty).$$

- *Kiểm định một phía*:

• Nếu $H_1 : \mu > \mu_0$ thì miền bác bỏ là $R_\alpha = (U_\alpha; +\infty)$;

• Nếu $H_1 : \mu < \mu_0$ thì miền bác bỏ là $(-\infty; -U_\alpha)$.

Ví dụ 3.17. Một tạp chí công nghệ thông tin thông báo có 25% học sinh phổ thông trung học là độc giả thường xuyên. Một mẫu ngẫu nhiên gồm 200 học sinh được chọn cho thấy có 45 em đọc tạp chí đó thường xuyên. Kiểm định tính chính xác của thông báo trên với mức ý nghĩa 0,05.

Lời giải. Gọi p là tỉ lệ học sinh phổ thông trung học đọc tạp chí công nghệ thông tin.

Ta kiểm định: Giả thuyết $H_0 : p = 0,25$, đối thuyết $H_1 : p \neq 0,25$.

Tiêu chuẩn kiểm định $T = \dfrac{f-0,25}{\sqrt{0,25.(1-0,25)}}\sqrt{45}$. Với mức ý nghĩa $\alpha = 0,05$ thì $U_{0,025} = 1,96$.

Mẫu đã cho: $f = \dfrac{45}{200} = 0,225$; $n = 45$, do đó giá trị quan sát thực tế là

$$|T_0| = \left|\frac{0,225-0,25}{0,25.0,75}\sqrt{45}\right| = 0,806 < 1,96 = U_{0,025}$$

nên ta không có cơ sở để bác bỏ thông báo của tạp chí đó.