

BÀI GIẢI
XÁC SUẤT THỐNG KÊ
(GV: Trần Ngọc Hội – 2009)

CHƯƠNG 4

KIỂM ĐỊNH GIẢ THIẾT

Bài 4.1. Trọng lượng của một sản phẩm theo qui định là 6kg. Sau một thời gian sản xuất, người ta tiến hành kiểm tra 121 sản phẩm và tính được trung bình mẫu là 5,975kg và phương sai mẫu hiệu chỉnh 5,7596kg². Sản xuất được xem là bình thường nếu các sản phẩm có trọng lượng trung bình bằng trọng lượng qui định. Với mức ý nghĩa 5%, hãy kết luận về tình hình sản xuất.

Lời giải

Gọi X là trọng lượng của một sản phẩm. Giả thiết cho ta:

- Cỡ mẫu n = 121.
- Kỳ vọng mẫu của X là $\bar{X} = 5,975$ (kg).
- Phương sai mẫu hiệu chỉnh của X là $S^2 = 5,7596(\text{kg}^2)$.
- Độ lệch mẫu hiệu chỉnh của X là $S = 2,3999(\text{kg})$.

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 5\% = 0,05$:

$H_0: \mu = 6$ với giả thiết đối $H_1: \mu \neq 6$.

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} = \frac{(5,975 - 6)\sqrt{121}}{2,3999} = -0,1146.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thỏa

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0,95/2 = 0,475$$

ta được $z_\alpha = 1,96$.

Bước 3: Kiểm định. Vì $|z| = 0,1146 < 1,96 = z_\alpha$ nên ta chấp nhận giả thiết $H_0: \mu = 6$.

Kết luận: Với mức ý nghĩa 5%, tình hình sản xuất được xem là bình thường.

Bài 4.2. Trọng lượng của một sản phẩm có phân phối chuẩn với trọng lượng trung bình là 500g. Sau một thời gian sản xuất, người ta nghi ngờ trọng lượng trung bình của loại sản phẩm này có xu hướng giảm nên tiến hành kiểm tra 25 sản phẩm và thu được kết quả sau:

Trọng lượng (g)	480	485	490	495	500	510
Số sản phẩm	2	3	8	5	3	4

Với mức ý nghĩa 3%, hãy kết luận điều nghi ngờ trên có đúng hay không.

Lời giải

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 3\% = 0,03$:

$H_0: \mu = 500$ với giả thiết đối $H_1: \mu < 500$.

Ta có:

X_i	480	485	490	495	500	510
n_i	2	3	8	5	3	4

$$n = 25; \sum X_i n_i = 12350; \sum X_i^2 n_i = 6102800.$$

- Kỳ vọng mẫu của X là

$$\bar{X} = \frac{1}{n} \sum X_i n_i = 494(\text{g}).$$

- Phương sai mẫu của X là:

$$\hat{S}^2 = \frac{1}{n} \sum X_i^2 n_i - \bar{X}^2 = (8,7178)^2(\text{g}^2).$$

- Phương sai mẫu hiệu chỉnh của X là:

$$S^2 = \frac{n}{n-1} \hat{S}^2 = (8,8976)^2(\text{g}^2).$$

Vì $n < 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} = \frac{(494 - 500)\sqrt{25}}{8,8976} = -3,3717.$$

Bước 2: Đặt $k = n - 1 = 24$. Tra bảng phân phối Student ứng với $k = 24$ và $2\alpha = 0,06$ ta được $t_{2\alpha} = 1,974$.

Bước 3: Kiểm định. Vì $-z = 3,3717 > 1,974 = t_{2\alpha}$ nên ta bác bỏ giả thiết $H_0: \mu = 500$, nghĩa là chấp nhận $H_1: \mu < 500$.

Kết luận: Với mức ý nghĩa 3%, điều nghi ngờ trọng lượng trung bình của loại sản phẩm này có xu hướng giảm là đúng.

Bài 4.3. Năng suất lúa trung bình của những vụ trước là 5,5tấn/ha. Vụ lúa năm nay người ta áp dụng một phương pháp kỹ thuật mới cho toàn bộ

diện tích trồng lúa trong vùng. Điều tra năng suất 100ha lúa, ta có bảng số liệu sau:

Năngsuất (tạ/ha)	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80
Diện tích (ha)	7	12	18	27	20	8	5	3

Với mức ý nghĩa 1%, hãy kết luận xem phương pháp kỹ thuật mới có làm tăng năng suất lúa trung bình của vùng này hay không?

Lời giải

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 1\% = 0,01$:

$$H_0: \mu = 55 \text{ với giả thiết đối } H_1: \mu > 55.$$

(5,5tấn = 55tạ).

Ta có:

X_i	42,5	47,5	52,5	57,5	62,5	67,5	72,5	77,5
n_i	7	12	18	27	20	8	5	3

$$n = 100; \sum X_i n_i = 5750; \sum X_i^2 n_i = 337475.$$

- Kỳ vọng mẫu của X là

$$\bar{X} = \frac{1}{n} \sum X_i n_i = 57,5 (\text{tạ}).$$

- Phương sai mẫu của X là:

$$\hat{S}^2 = \frac{1}{n} \sum X_i^2 n_i - \bar{X}^2 = (8,2765)^2 (\text{tạ}^2).$$

- Phương sai mẫu hiệu chỉnh của X là:

$$S^2 = \frac{n}{n-1} \hat{S}^2 = (8,3182)^2 (\text{tạ}^2).$$

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} = \frac{(57,5 - 55)\sqrt{100}}{8,3182} = 3,0055.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thỏa

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,98/2 = 0,49$$

ta được $z_{2\alpha} = 2,33$.

Bước 3: Kiểm định. Vì $z = 3,0055 > 2,33 = z_{2\alpha}$ nên ta bác bỏ giả thiết $H_0: \mu = 55$, nghĩa là chấp nhận $H_1: \mu > 55$.

Kết luận: Với mức ý nghĩa 1%, phương pháp kỹ thuật mới làm tăng năng suất lúa trung bình của vùng này.

Bài 4.4. Một công ty dự định mở một siêu thị tại một khu dân cư. Để đánh giá khả năng mua hàng của dân cư trong khu vực, người ta tiến

hành điều tra về thu nhập của 100 hộ trong khu vực và có bảng số liệu sau:

Thu nhập bình quân (ngàn/người/tháng)	150	200	250	300	350
Số hộ	8	15	38	22	17

Theo bộ phận tiếp thị thì siêu thị chỉ hoạt động có hiệu quả tại khu vực này khi thu nhập bình quân hàng tháng của các hộ tối thiểu là vào khoảng 250ngàn/người/tháng. Vậy theo kết quả điều tra trên, công ty có nên quyết định mở siêu thị tại khu vực này hay không với mức ý nghĩa 5%?

Lời giải

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 5\% = 0,05$:

$$H_0: \mu = 250 \text{ với giả thiết đối } H_1: \mu > 250.$$

Ta có:

X_i	150	200	250	300	350
n_i	8	15	38	22	17

$$n = 100; \sum X_i n_i = 26250; \sum X_i^2 n_i = 7217500.$$

- Kỳ vọng mẫu của X là

$$\bar{X} = \frac{1}{n} \sum X_i n_i = 262,5.$$

- Phương sai mẫu của X là:

$$\hat{S}^2 = \frac{1}{n} \sum X_i^2 n_i - \bar{X}^2 = (57,1730)^2.$$

- Phương sai mẫu hiệu chỉnh của X là:

$$S^2 = \frac{n}{n-1} \hat{S}^2 = (57,4610)^2.$$

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} = \frac{(262,5 - 250)\sqrt{100}}{57,4610} = 2,1754.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thỏa

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,90/2 = 0,45$$

ta được $z_{2\alpha} = 1,65$.

Bước 3: Kiểm định. Vì $z = 2,1754 > 1,65 = z_{2\alpha}$ nên ta bác bỏ giả thiết $H_0: \mu = 250$, chấp nhận giả thiết $H_1: \mu > 250$, nghĩa là thu nhập bình quân của các hộ cao hơn 250ngàn/người/tháng.

Kết luận: Với mức ý nghĩa 5%, công ty nên quyết định mở siêu thị tại khu vực này.

Bài 4.5. Để nghiên cứu nhu cầu của một loại hàng, người ta tiến hành khảo sát nhu cầu của mặt hàng này ở 400 hộ. Kết quả như sau:

Nhu cầu (kg/tháng)	0	0-1	1-2	2-3	3-4	4-5	5-6	6-7
Số hộ	10	35	86	132	78	31	18	10

Giả sử khu vực đó có 4000 hộ. Nếu cho rằng nhu cầu trung bình về mặt hàng này của toàn khu vực là 14tấn/tháng thì có chấp nhận được không với mức ý nghĩa 2%?

Lời giải

Khi cho rằng nhu cầu trung bình về mặt hàng này của toàn khu vực là 14tấn/tháng, nghĩa là nhu cầu trung bình về mặt hàng này của một hộ trong một tháng là

$$\frac{14\text{tấn}}{4000} = \frac{14000\text{kg}}{4000} = 3,5\text{kg.}$$

Do đó đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 2\% = 0,02$:

$$H_0: \mu = 3,5 \text{ với giả thiết đối } H_1: \mu \neq 3,5.$$

Ta có:

X_i	0	0,5	1,5	2,5	3,5	4,5	5,5	6,5
n_i	10	35	86	132	78	31	18	10

$$n = 400; \sum X_i n_i = 1053; \sum X_i^2 n_i = 3577,5.$$

- Kỳ vọng mẫu của X là

$$\bar{X} = \frac{1}{n} \sum X_i n_i = 2,6325.$$

- Phương sai mẫu của X là:

$$\hat{S}^2 = \frac{1}{n} \sum X_i^2 n_i - \bar{X}^2 = (1,4190)^2.$$

- Phương sai mẫu hiệu chỉnh của X là:

$$S^2 = \frac{n}{n-1} \hat{S}^2 = (1,4208)^2.$$

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} = \frac{(2,6325 - 3,5)\sqrt{400}}{1,4208} = -12,2114.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0,98/2 = 0,49$$

ta được $z_\alpha = 2,33$.

Bước 3: Kiểm định. Vì $|z| = 12,2114 > 2,33 = z_\alpha$ nên ta bác bỏ giả thiết $H_0: \mu = 3,5$, chấp nhận giả thiết $H_1: \mu \neq 3,5$.

Kết luận: Với mức ý nghĩa 5%, không thể cho rằng nhu cầu trung bình về mặt hàng này của toàn khu vực là 14tấn/tháng.

Bài 4.6. Trọng lượng của một loại gà công nghiệp ở một trại chăn nuôi có phân phối chuẩn. Trọng lượng trung bình khi xuất chuồng năm trước là 2,8kg/con. Năm nay, người ta sử dụng một loại thức ăn mới. Cân thử 25 con khi xuất chuồng người ta tính được trung bình mẫu là 3,2kg và phương sai mẫu hiệu chỉnh 0,25kg².

- a) Với mức ý nghĩa 5%, hãy kết luận xem loại thức ăn mới có thực sự làm tăng trọng lượng trung bình của đàn gà hay không?
- b) Nếu trại chăn nuôi báo cáo trọng lượng trung bình khi xuất chuồng là 3,3kg/con thì có chấp nhận được không với mức ý nghĩa 5%?

Lời giải

Gọi X là trọng lượng của một con gà sau khi sử dụng loại thức ăn mới.

Giả thiết cho ta:

- X có phân phối chuẩn.
- Cỡ mẫu $n = 25$.
- Kỳ vọng mẫu của X là $\bar{X} = 3,2(\text{kg})$.
- Phương sai mẫu hiệu chỉnh của X là $S^2 = 0,25(\text{kg}^2)$.
- Độ lệch mẫu hiệu chỉnh của X là $S = 0,5(\text{kg})$.

- a) Với mức ý nghĩa 5%, hãy kết luận xem loại thức ăn mới có thực sự làm tăng trọng lượng trung bình của đàn gà hay không?

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 5\% = 0,05$:

$$H_0: \mu = 2,8 \text{ với giả thiết đối } H_1: \mu > 2,8.$$

Vì $n < 30$; X có phân phối chuẩn; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} = \frac{(3,2 - 2,8)\sqrt{25}}{0,5} = 4.$$

Bước 2: Đặt $k = n - 1 = 24$. Tra bảng phân phối Student ứng với $k = 24$ và $2\alpha = 0,1$ ta được $t_{2\alpha} = t_{0,1}^{24} = 1,711$.

Bước 3: Kiểm định. Vì $z = 4 > 1,711 = t_{2\alpha}$ nên ta bác bỏ giả thiết $H_0: \mu = 2,8$, giả là chấp nhận giả thiết $H_1: \mu > 2,8$.

Kết luận: Với mức ý nghĩa 5%, loại thức ăn mới thực sự làm tăng trọng lượng trung bình của đàn gà.

b) Nếu trại chăn nuôi báo cáo trọng lượng trung bình khi xuất chuồng là 3,3kg/con thì có chấp nhận được không với mức ý nghĩa 5%?

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 5\% = 0,05$:

$H_0: \mu = 3,3$ với giả thiết đối $H_1: \mu \neq 3,3$.

Vì $n < 30$; X có phân phối chuẩn; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} = \frac{(3,2 - 3,3)\sqrt{25}}{0,5} = -1.$$

Bước 2: Đặt $k = n - 1 = 24$. Tra bảng phân phối Student ứng với $k = 24$ và $\alpha = 0,05$ ta được $t_\alpha = t_\alpha^k = 2,064$.

Bước 3: Kiểm định. Vì $|z| = 1 < 2,064 = t_\alpha$ nên ta chấp nhận giả thiết $H_0: \mu = 3,3$.

Kết luận: Với mức ý nghĩa 5%, báo cáo trọng lượng trung bình khi xuất chuồng là 3,3kg/con là chấp nhận được.

Bài 4.7. Chiều cao trung bình của 100 nam sinh lớp 12 ở một trường trung học nội thành là 1,68m với độ lệch mẫu hiệu chỉnh 6cm. Trong khi kiểm tra 120 nam sinh lớp 12 ở một huyện ngoại thành thì chiều cao trung bình là 1,64m với độ lệch mẫu hiệu chỉnh 5cm. Với mức ý nghĩa 1%, có thể kết luận rằng nam sinh nội thành thực sự cao hơn nam sinh ngoại thành hay không?

Lời giải

Gọi X, Y (cm) lần lượt là chiều cao của nam sinh nội thành và nam sinh ngoại thành. Bài toán trên chính là bài toán kiểm định so sánh hai kỳ vọng với mức ý nghĩa $\alpha = 1\% = 0,01$:

$H_0: \mu_X = \mu_Y$ với giả thiết đối $H_1: \mu_X > \mu_Y$.

1) Đối với X , giả thiết cho ta:

- Cỡ mẫu $n_X = 100$.
- Kỳ vọng mẫu của X là $\bar{X} = 168(\text{cm})$.
- Độ lệch mẫu hiệu chỉnh của X là $S_X = 6(\text{cm})$.

2) Đối với Y , giả thiết cho ta:

- Cỡ mẫu $n_Y = 120$
- Kỳ vọng mẫu của Y là $\bar{Y} = 164(\text{cm})$.
- Độ lệch mẫu hiệu chỉnh của Y là $S_Y = 5(\text{cm})$.

Vì $n_X > 30$; $n_Y > 30$ nên ta kiểm định như sau:

Bước 1: Ta có:

$$z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} = \frac{168 - 164}{\sqrt{\frac{6^2}{100} + \frac{5^2}{120}}} = 5,3059.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,98/2 = 0,49$$

ta được $z_{2\alpha} = 2,33$.

Bước 3: Kiểm định. Vì $z = 5,3059 > 2,33 = z_{2\alpha}$ nên ta bác bỏ giả thiết $H_0: \mu_X = \mu_Y$, nghĩa là chấp nhận $H_1: \mu_X > \mu_Y$.

Kết luận: Với mức ý nghĩa 1%, có thể kết luận rằng nam sinh nội thành thực sự cao hơn nam sinh ngoại thành.

Bài 4.8. Một hợp tác xã trồng thử hai giống lúa, mỗi giống trên 30 thửa ruộng và được chăm sóc như nhau. Cuối vụ thu hoạch ta được số liệu như sau:

	Năng suất trung bình (tạ/ha)	Độ lệch mẫu hiệu chỉnh
Giống lúa 1	45	2,5
Giống lúa 2	46,5	4,0

a) Với mức ý nghĩa 2%, có thể xem năng suất của hai giống lúa trên là như nhau hay không?

b) Với mức ý nghĩa 2%, có thể xem năng suất của giống lúa 2 cao hơn của giống lúa 1 hay không?

Lời giải

Gọi X, Y (tạ/ha) lần lượt là năng suất của giống lúa 1 và 2. Khi đó:

1) Đối với X , giả thiết cho ta:

- Cỡ mẫu $n_X = 30$.
- Kỳ vọng mẫu của X là $\bar{X} = 45$.
- Độ lệch mẫu hiệu chỉnh của X là $S_X = 2,5$.

2) Đối với Y , giả thiết cho ta:

- Cỡ mẫu $n_Y = 30$.
- Kỳ vọng mẫu của Y là $\bar{Y} = 46,5$.
- Độ lệch mẫu hiệu chỉnh của Y là $S_Y = 4$.

a) Với mức ý nghĩa 2%, có thể xem năng suất của hai giống lúa trên là như nhau hay không?

Đây là bài toán kiểm định so sánh hai kỳ vọng với mức ý nghĩa $2\% = 0,02$:

$H_0: \mu_X = \mu_Y$ với giả thiết đối $H_1: \mu_X \neq \mu_Y$.

Vì $n_X = n_Y = 30$ nên ta kiểm định như sau:

Bước 1: Ta có:

$$z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} = \frac{45 - 46,5}{\sqrt{\frac{2,5^2}{30} + \frac{4^2}{30}}} = -1,7418.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0,98/2 = 0,49$$

ta được $z_\alpha = 2,33$.

Bước 3: Kiểm định. Vì $|z| = 1,7418 < 2,33 = z_\alpha$ nên ta chấp nhận giả thiết $H_0: \mu_X = \mu_Y$.

Kết luận: Với mức ý nghĩa 2%, có thể xem năng suất của hai giống lúa trên là như nhau.

b) Với mức ý nghĩa 2%, có thể xem năng suất của giống lúa 2 cao hơn của giống lúa 1 hay không?

Đây là bài toán kiểm định so sánh hai kỳ vọng với mức ý nghĩa $\alpha = 2\% = 0,02$:

$H_0: \mu_X = \mu_Y$ với giả thiết đối $H_1: \mu_X < \mu_Y$.

Vì $n_X = n_Y = 30$ nên ta kiểm định như sau:

Bước 1: Tương tự câu a), ta có:

$$z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} = -1,7418.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,96/2 = 0,48$$

ta được $z_{2\alpha} = 2,06$.

Bước 3: Kiểm định. Vì $-z = 1,7418 < 2,06 = z_{2\alpha}$ nên ta chấp nhận giả thiết $H_0: \mu_X = \mu_Y$.

Kết luận: Với mức ý nghĩa 2%, chưa thể xem năng suất của giống lúa 2 cao hơn của giống lúa 1.

Bài 4.9. Một máy sản xuất tự động, lúc đầu tỉ lệ sản phẩm loại A là 45%. Sau khi áp dụng một phương pháp sản xuất mới, người ta lấy ra 400 sản phẩm để kiểm tra thì thấy có 215 sản phẩm loại A. Với mức ý nghĩa 5%, hãy kết luận xem phương pháp mới có thực sự làm tăng tỉ lệ sản phẩm loại A hay không?

Lời giải

Từ giả thiết ta suy ra:

- Cỡ mẫu $n = 400$.
- Số sản phẩm loại A có trong mẫu là $m = 215$.
- Tỉ lệ mẫu sản phẩm loại A là $F_n = m/n = 215/400 = 0,5375$.

Ta đưa bài toán về bài toán kiểm định giả thiết về tỉ lệ p các sản phẩm loại A với mức ý nghĩa $\alpha = 5\% = 0,05$:

$H_0: p = 45\% = 0,45$ với giả thiết đối $H_1: p > 0,45$.

Ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(F_n - p_0)\sqrt{n}}{\sqrt{p_0(1 - p_0)}} = \frac{(0,5375 - 0,45)\sqrt{400}}{\sqrt{0,45(1 - 0,45)}} = 3,5176.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,90/2 = 0,45$$

ta được $z_{2\alpha} = 1,65$.

Bước 3: Kiểm định. Vì $z = 3,5176 > 1,65 = z_{2\alpha}$ nên ta bác bỏ giả thiết $H_0: p = 0,45$, nghĩa là chấp nhận $H_1: p > 0,45$.

Kết luận: Với mức ý nghĩa 5%, phương pháp mới thực sự làm tăng tỉ lệ sản phẩm loại A.

Bài 4.10. Thống kê 10650 trẻ sơ sinh ở một địa phương người ta thấy có 5410 bé trai.

- Với mức ý nghĩa 3%, hỏi có sự khác biệt về tỉ lệ sinh bé trai và bé gái hay không?
- Với mức ý nghĩa 1%, hỏi tỉ lệ sinh bé trai có thực sự cao hơn tỉ lệ sinh bé gái hay không?

Lời giải

Từ các giả thiết của bài toán ta suy ra:

1) Khi khảo sát tỉ lệ bé trai p_1 :

- Cỡ mẫu $n_1 = 10650$.
- Số bé trai là $m_1 = 5410$.
- Tỉ lệ bé trai $F_{n_1} = 5410/10650$.

2) Khi khảo sát tỉ lệ bé gái p_2 :

- Cỡ mẫu $n_2 = 10650$.
- Số bé gái là $m_2 = 10650 - 5410 = 5240$.
- Tỉ lệ bé gái $F_{n_2} = 5240/10650$.

3) $p_0 = 0,5$.

a) Với mức ý nghĩa 3%, hỏi có sự khác biệt về tỉ lệ sinh bé trai và bé gái hay không?

Đây là bài toán kiểm định giả thiết so sánh hai tỉ lệ với mức ý nghĩa $\alpha = 3\% = 0,03$:

$H_0: p_1 = p_2 (= p_0)$ với giả thiết đối $H_1: p_1 \neq p_2$

Ta kiểm định như sau:

Bước 1: Ta có:

$$z = \frac{F_{n1} - F_{n2}}{\sqrt{p_0(1-p_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{\frac{5410}{10650} - \frac{5240}{10650}}{\sqrt{0,5(1-0,5)\left(\frac{1}{10650} + \frac{1}{10650}\right)}} = 2,3296.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0,97/2 = 0,485$$

ta được $z_\alpha = 2,17$.

Bước 3: Kiểm định. Vì $|z| = 2,3296 > 2,17 = z_\alpha$ nên ta bác bỏ giả thiết $H_0: p_1 = p_2$, nghĩa là chấp nhận $H_1: p_1 \neq p_2$.

Kết luận: Với mức ý nghĩa 3%, có sự khác biệt về tỉ lệ sinh bé trai và bé gái.

b) Với mức ý nghĩa 1%, hỏi tỉ lệ sinh bé trai có thực sự cao hơn tỉ lệ sinh bé gái hay không?

Đây là bài toán kiểm định giả thiết so sánh hai tỉ lệ với mức ý nghĩa $\alpha = 1\% = 0,01$:

$$H_0: p_1 = p_2 \text{ với giả thiết đối } H_1: p_1 > p_2$$

Ta kiểm định như sau:

Bước 1: Tương tự câu a), ta có:

$$z = \frac{F_{n1} - F_{n2}}{\sqrt{p_0(1-p_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 2,3296.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,98/2 = 0,49$$

ta được $z_{2\alpha} = 2,33$.

Bước 3: Kiểm định. Vì $z = 2,3296 < 2,33 = z_{2\alpha}$ nên ta chấp nhận giả thiết $H_0: p_1 = p_2$.

Kết luận: Với mức ý nghĩa 1%, chưa thể nói tỉ lệ sinh bé trai thực sự cao hơn tỉ lệ sinh bé gái.

Bài 4.11. Bệnh A có thể chữa bằng hai loại thuốc H và K. Công ty sản xuất thuốc H tuyên bố tỉ lệ bệnh nhân khỏi bệnh do dùng thuốc của họ là 85%. Người ta dùng thử thuốc H chữa cho 250 bệnh nhân thì thấy có 210 người khỏi bệnh, dùng thử thuốc K cho 200 bệnh nhân thì thấy có 175 người khỏi bệnh.

a) Với mức ý nghĩa 1% có thể kết luận thuốc K có khả năng chữa bệnh A tốt hơn thuốc H hay không?

b) Xét xem hiệu quả chữa bệnh của thuốc H có đúng như công ty quảng cáo với mức ý nghĩa 5% hay không.

Lời giải

Từ các giả thiết của bài toán ta suy ra:

1) Đối với loại thuốc H:

- Cỡ mẫu $n_1 = 250$.
- Số bệnh nhân khỏi bệnh: 210.
- Tỉ lệ mẫu bệnh nhân khỏi bệnh $F_{n1} = 210/250 = 0,84$.

2) Đối với loại thuốc K:

- Cỡ mẫu $n_2 = 200$.
- Số bệnh nhân khỏi bệnh: 175.
- Tỉ lệ mẫu bệnh nhân khỏi bệnh $F_{n2} = 175/200 = 0,875$.

$$3) p_0 = \frac{n_1 F_{n1} + n_2 F_{n2}}{n_1 + n_2} = \frac{250 \cdot 0,84 + 200 \cdot 0,875}{250 + 200} = \frac{385}{450}$$

a) Với mức ý nghĩa 1% có thể kết luận thuốc K có khả năng chữa bệnh A tốt hơn thuốc H hay không?

Đây là bài toán kiểm định giả thiết so sánh hai tỉ lệ với mức ý nghĩa $\alpha = 1\% = 0,01$:

$$H_0: p_1 = p_2 \text{ với giả thiết đối } H_1: p_1 < p_2$$

Ta kiểm định như sau:

Bước 1: Ta có:

$$z = \frac{F_{n1} - F_{n2}}{\sqrt{p_0(1-p_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0,84 - 0,875}{\sqrt{\frac{385}{450}\left(1 - \frac{385}{450}\right)\left(\frac{1}{250} + \frac{1}{200}\right)}} = -1,0495.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,98/2 = 0,49$$

ta được $z_{2\alpha} = 2,33$.

Bước 3: Kiểm định. Vì $-z = 1,0495 < 2,33 = z_{2\alpha}$ nên ta chấp nhận giả thiết $H_0: p_1 = p_2$.

Kết luận: Với mức ý nghĩa 1%, không thể kết luận thuốc K có khả năng chữa bệnh A tốt hơn thuốc H.

b) Xét xem hiệu quả chữa bệnh của thuốc H có đúng như công ty quảng cáo với mức ý nghĩa 5% hay không.

Đây là bài toán kiểm định giả thiết về tỉ lệ p_1 các bệnh nhân khỏi bệnh A khi được điều trị bằng thuốc H với mức ý nghĩa $\alpha = 5\% = 0,05$:

$$H_0: p_1 = 85\% = 0,85 \text{ với giả thiết đối } H_1: p_1 < 0,85.$$

Ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(F_{n_1} - p_{01})\sqrt{n_1}}{\sqrt{p_{01}q_{01}}} = \frac{(0,84 - 0,85)\sqrt{250}}{\sqrt{0,85(1-0,85)}} = -0,4428.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,90/2 = 0,45$$

ta được $z_{2\alpha} = 1,65$.

Bước 3: Kiểm định. Vì $-z = 0,4428 < 1,65 = z_{2\alpha}$ nên ta chấp nhận giả thiết $H_0: p_1 = 0,85$.

Kết luận: Với mức ý nghĩa 5%, hiệu quả chữa bệnh của thuốc H đúng như công ty quảng cáo.

Bài 4.12. Để khảo sát chỉ tiêu X của một loại sản phẩm, người ta quan sát một mẫu và có kết quả sau:

X(cm)	11-15	15-19	19-23	23-27	27-31	31-35	35-39
Số sản phẩm	8	9	20	16	16	13	18

Những sản phẩm có chỉ tiêu X từ 19cm trở xuống được gọi là những sản phẩm loại B.

- Giả sử trung bình tiêu chuẩn của chỉ tiêu X là 29cm. Hãy cho nhận xét về tình hình sản xuất với mức ý nghĩa 2%.
- Bằng phương pháp sản xuất mới, sau một thời gian người ta thấy giá trị trung bình của chỉ tiêu X của những sản phẩm loại B là 16cm. Hãy cho kết luận về phương pháp sản xuất mới với mức ý nghĩa 1% (GS X có phân phối chuẩn).
- Một tài liệu thống kê cũ cho rằng tỉ lệ những sản phẩm loại B là 12%. Hãy nhận định về tài liệu này với mức ý nghĩa 5%.

Lời giải

Lập bảng:

X_i	13	17	21	25	29	33	37
n_i	8	9	20	16	16	13	18

Ta có:

$$n = 100; \sum X_i n_i = 2636; \sum X_i^2 n_i = 75028.$$

- Kỳ vọng mẫu của X là

$$\bar{X} = \frac{1}{n} \sum X_i n_i = 26,36(\text{cm}).$$

- Phương sai mẫu của X là:

$$\hat{S}^2 = \frac{1}{n} \sum X_i^2 n_i - \bar{X}^2 = (7,4452)^2 (\text{cm}^2).$$

- Phương sai mẫu đã hiệu chỉnh của X là:

$$S^2 = \frac{n}{n-1} \hat{S}^2 = (7,4827)^2 (\text{cm}^2).$$

- Giả sử trung bình tiêu chuẩn của chỉ tiêu X là 29cm. Hãy cho nhận xét về tình hình sản xuất với mức ý nghĩa 2%.

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 2\% = 0,02$:

$H_0: \mu = 29$ với giả thiết đối $H_1: \mu \neq 29$.

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} = \frac{(26,36 - 29)\sqrt{100}}{7,4827} = -3,5281.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0,98/2 = 0,49$$

ta được $z_\alpha = 2,33$.

Bước 3: Kiểm định. Vì $|z| = 3,5281 > 2,33 = z_\alpha$ nên ta bác bỏ giả thiết $H_0: \mu=29$, nghĩa là chấp nhận $H_1: \mu \neq 29$.

Kết luận: Với mức ý nghĩa 1%, tình hình sản xuất không bình thường vì giá trị trung bình của chỉ tiêu X không đúng tiêu chuẩn.

- Bằng phương pháp sản xuất mới, sau một thời gian người ta thấy giá trị trung bình của chỉ tiêu X của những sản phẩm loại B là 16cm. Hãy cho kết luận về phương pháp sản xuất mới với mức ý nghĩa 1% (GS X có phân phối chuẩn).

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu_B = M(X_B)$ của chỉ tiêu $X = X_B$ của những sản phẩm loại B với mức ý nghĩa $\alpha = 1\% = 0,01$:

$H_0: \mu_B = 16$ với giả thiết đối $H_1: \mu_B \neq 16$.

Ta lập bảng số liệu của X_B :

X_{Bi}	13	17
n_{Bi}	8	9

Từ bảng trên ta tính được:

$$n_B = 17; \sum X_{Bi} n_{Bi} = 257; \sum X_{Bi}^2 n_{Bi} = 3,953.$$

- Kỳ vọng mẫu của X_B là

$$\bar{X}_B = \frac{1}{n} \sum X_{Bi} n_{Bi} = 15,1176 (\text{cm}).$$

- Phương sai mẫu của X_B là:

$$\hat{S}_B^2 = \frac{1}{n} \sum X_{Bi}^2 n_{Bi} - \bar{X}_B^2 = (1,9965)^2 (\text{cm}^2).$$

- Phương sai mẫu đã hiệu chỉnh của X_B là:

$$S_B^2 = \frac{n_B}{n_B - 1} \hat{S}_B^2 = (2,0580)^2 (\text{cm}^2).$$

Vì $n_B < 30$, X_B có phân phối chuẩn, $\sigma_B^2 = D(X_B)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X}_B - \mu_0)\sqrt{n_B}}{S_B} = \frac{(15,1176 - 16)\sqrt{17}}{2,0580} = -1,7678.$$

Bước 2: Đặt $k = n_B - 1 = 16$. Tra bảng phân phối Student ứng với $k = 16$ và $\alpha = 0,01$ ta được $t_{\alpha} = 2,921$.

Bước 3: Kiểm định. Vì $|z| = 1,7678 < 2,921 = t_{\alpha}$ nên ta chấp nhận giả thiết $H_0: \mu_B = 16$.

Kết luận: Với mức ý nghĩa 2%, phương pháp mới không có tác dụng làm thay đổi giá trị trung bình của chỉ tiêu X_B của các sản phẩm loại B.

c) Một tài liệu thống kê cũ cho rằng tỉ lệ sản phẩm loại B là 12%. Hãy nhận định về tài liệu này với mức ý nghĩa 5%.

Đây là bài toán kiểm định giả thiết về tỉ lệ p các sản phẩm loại B với mức ý nghĩa $\alpha = 5\% = 0,05$:

$H_0: p = 12\% = 0,12$ với giả thiết đối $H_1: p \neq 0,12$.

Ta có qui tắc kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(F_n - p_0)\sqrt{n}}{\sqrt{p_0 q_0}} = \frac{(0,17 - 0,12)\sqrt{100}}{\sqrt{0,12(1 - 0,12)}} = 1,5386.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_{α} thoả

$$\varphi(z_{\alpha}) = (1 - \alpha)/2 = 0,95/2 = 0,475$$

ta được $z_{\alpha} = 1,96$.

Bước 3: Kiểm định. Vì $|z| = 1,5386 < 1,96 = z_{\alpha}$ nên ta chấp nhận giả thiết $H_0: p = 0,12$.

Kết luận: Với mức ý nghĩa 5%, tài liệu cũ về tỉ lệ sản phẩm loại B còn phù hợp.

Bài 4.13. Để khảo sát chiều cao X của một giống cây trồng, người ta quan sát một mẫu và có kết quả sau:

X(cm)	95-105	105-115	115-125	125-135	135-145	145-155	155-165
Số cây	10	10	15	30	10	10	15

a) Một tài liệu thống kê cũ cho rằng chiều cao trung bình của giống cây trồng trên là 127cm. Hãy cho kết luận về tài liệu đó với mức ý nghĩa 1%.

b) Những cây trồng có chiều cao từ 135cm trở lên được gọi là những cây “cao”. Trước đây, tỉ lệ những cây “cao” của loại cây trồng trên là 40%. Các số liệu trên thu thập được sau khi đã áp dụng một kỹ thuật mới. Hãy cho kết luận về kỹ thuật mới với mức ý nghĩa 5%.

c) Những cây trồng có chiều cao từ 105cm đến 125cm được gọi là những cây loại A. Bằng phương pháp mới, sau một thời gian người ta thấy chiều cao trung bình của những cây loại A là 119,5cm. Hãy cho kết luận về phương pháp mới với mức ý nghĩa 1% (GS X có phân phối chuẩn).

Lời giải

X_i	100	110	120	130	140	150	160
n_i	10	10	15	30	10	10	15

Ta có:

$$n = 100; \sum X_i n_i = 13100; \sum X_i^2 n_i = 1749000.$$

• Kỳ vọng mẫu của X là

$$\bar{X} = \frac{1}{n} \sum X_i n_i = 131(\text{cm}).$$

• Phương sai mẫu của X là:

$$\hat{S}^2 = \frac{1}{n} \sum X_i^2 n_i - \bar{X}^2 = (18,1384)^2 (\text{cm}^2).$$

• Phương sai mẫu hiệu chỉnh của X là:

$$S^2 = \frac{n}{n-1} \hat{S}^2 = (18,2297)^2 (\text{cm}^2).$$

a) Một tài liệu thống kê cũ cho rằng chiều cao trung bình của giống cây trồng trên là 127cm. Hãy cho kết luận về tài liệu đó với mức ý nghĩa 1%.

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 1\% = 0,01$:

$H_0: \mu = 127$ với giả thiết đối $H_1: \mu \neq 127$

Vì $n \geq 30$; σ^2 chưa biết, nên ta có qui tắc kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} = \frac{(131 - 127)\sqrt{100}}{18,2297} = 2,1942.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_{α} thoả

$$\varphi(z_{\alpha}) = (1 - \alpha)/2 = 0,99/2 = 0,495$$

ta được $z_{\alpha} = 2,58$.

Bước 3: Kiểm định. Vì $|z| = 2,1942 < 2,58 = z_{\alpha}$ nên ta chấp nhận $H_0: \mu = 127$.

Kết luận: Với mức ý nghĩa 1%, tài liệu cũ về chiều cao trung bình của giống cây trồng trên còn phù hợp với thực tế.

b) Những cây trồng có chiều cao từ 135cm trở lên được gọi là những cây “cao”. Trước đây, tỉ lệ những cây “cao” của loại cây trồng trên là 40%.

Các số liệu trên thu thập được sau khi đã áp dụng một kỹ thuật mới. Hãy cho kết luận về kỹ thuật mới với mức ý nghĩa 5%. Đây là bài toán kiểm định giả thiết về tỉ lệ p các cây cao với mức ý nghĩa $\alpha = 5\% = 0,05$:

$$H_0: p = 40\% = 0,4 \text{ với giả thiết đối } H_1: p \neq 0,4$$

Ta có qui tắc kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(F_n - p_0)\sqrt{n}}{\sqrt{p_0q_0}} = \frac{(0,35 - 0,4)\sqrt{100}}{\sqrt{0,4(1 - 0,4)}} = -1,0206.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thoả

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0,95/2 = 0,475$$

ta được $z_\alpha = 1,96$.

Bước 3: Kiểm định. Vì $|z| = 1,0206 < 1,96 = z_\alpha$ nên ta chấp nhận giả thiết $H_0: p = 0,4$.

Kết luận: Với mức ý nghĩa 5%, phương pháp mới không có tác dụng làm thay đổi tỉ lệ các cây cao.

c) Những cây trồng có chiều cao từ 105cm đến 125cm được gọi là những cây loại A. Bằng phương pháp mới, sau một thời gian người ta thấy chiều cao trung bình của những cây loại A là 119,5cm. Hãy cho kết luận về phương pháp mới với mức ý nghĩa 1% (GS X có phân phối chuẩn).

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu_A = M(X_A)$ của chiều cao $X = X_A$ của các cây loại A với mức ý nghĩa $\alpha = 1\% = 0,01$:

$$H_0: \mu_A = 119,5 \text{ với giả thiết đối } H_1: \mu_A \neq 119,5.$$

Ta lập bảng số liệu của X_A :

X_{Ai}	110	120
N_{Ai}	10	15

Từ bảng trên ta tính được:

$$n_A = 25; \sum X_{Ai}n_{Ai} = 2900; \sum X_{Ai}^2n_{Ai} = 337000.$$

- Kỳ vọng mẫu của X_A là

$$\bar{X}_A = \frac{1}{n} \sum X_{Ai}n_{Ai} = 116(\text{cm}).$$

- Phương sai mẫu của X_A là:

$$\hat{S}_A^2 = \frac{1}{n} \sum X_{Ai}^2n_{Ai} - \bar{X}_A^2 = (4,8990)^2(\text{cm}^2).$$

- Phương sai mẫu hiệu chỉnh của X_A là:

$$S_A^2 = \frac{n_A}{n_A - 1} \hat{S}_A^2 = 5^2(\text{cm}^2).$$

Vì $n_A = 25 < 30$, X_A có phân phối chuẩn, $\sigma_A^2 = D(X_A)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X}_A - \mu_0)\sqrt{n_A}}{S_A} = \frac{(116 - 119,5)\sqrt{25}}{5} = -3,5.$$

Bước 2: Đặt $k = n_A - 1 = 24$. Tra bảng phân phối Student ứng với $k = 24$ và $\alpha = 0,01$ ta được $t_\alpha = t_\alpha^k = 2,797$.

Bước 3: Kiểm định. Vì $|z| = 3,5 > 2,797 = t_\alpha$ nên ta bác bỏ giả thiết $H_0: \mu_A = 119,5$, nghĩa là chấp nhận $H_1: \mu_A \neq 119,5$. Cụ thể, ta nhận định $\mu_A < 119,5$ (vì $\bar{X}_A = 116 < 119,5$).

Kết luận: Với mức ý nghĩa 1%, phương pháp mới có tác dụng làm thay đổi chiều cao trung bình của các cây loại A, theo hướng làm tăng chiều cao trung bình của các cây loại này.

Bài 4.14. Cho các số liệu như Bài 4.13.

- Giả sử trung bình tiêu chuẩn của chiều cao X là 125cm. Có thể khẳng định rằng việc canh tác làm tăng chiều cao trung bình của giống cây trồng trên với mức ý nghĩa 1% hay không?
- Giả sử trung bình tiêu chuẩn của chiều cao X là 134cm. Có thể khẳng định rằng việc canh tác làm giảm chiều cao trung bình của giống cây trồng trên với mức ý nghĩa 2% hay không?
- Sau khi áp dụng phương pháp canh tác mới, người ta thấy chiều cao trung bình của các cây loại A là 114cm. Hãy kết luận xem phương pháp mới có làm giảm chiều cao trung bình của các cây loại A hay không với mức ý nghĩa 3% (Giả sử X có phân phối chuẩn).
- Trước đây, chiều cao trung bình của các cây loại A là 120cm. Các số liệu trên thu thập được sau khi đã áp dụng một kỹ thuật mới. Hãy kết luận xem kỹ thuật mới có làm giảm chiều cao trung bình của các cây loại A hay không với mức ý nghĩa 2% (Giả sử X có phân phối chuẩn).
- Sau khi áp dụng một phương pháp sản xuất, người ta thấy tỉ lệ cây loại A là 35%. Hãy kết luận xem phương pháp mới có làm tăng tỉ lệ cây loại A lên hay không với mức ý nghĩa 2%.
- Một tài liệu thống kê cũ cho rằng tỉ lệ cây loại A là 20%. Hãy xét xem việc canh tác có làm tăng tỉ lệ cây loại A hay không với mức ý nghĩa 5%?

Lời giải

Ta có:

- Cỡ mẫu là $n = 100$.
- Kỳ vọng mẫu của X là

$$\bar{X} = \frac{1}{n} \sum X_i n_i = 131(\text{cm}).$$

- Phương sai mẫu của X là

$$\hat{S}^2 = \frac{1}{n} \sum X_i^2 n_i - \bar{X}^2 = (18,1384)^2 (\text{cm}^2).$$

- Phương sai mẫu hiệu chỉnh của X là

$$S^2 = \frac{n}{n-1} \hat{S}^2 = (18,2297)^2 (\text{cm}^2).$$

- a) Giả sử trung bình tiêu chuẩn của chiều cao X là 125cm. Có thể khẳng định rằng việc canh tác làm tăng chiều cao trung bình của giống cây trồng trên với mức ý nghĩa 1% hay không?

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 1\% = 0,01$:

$$H_0: \mu = 125 \text{ với giả thiết đối } H_1: \mu > 125.$$

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} = \frac{(131 - 125)\sqrt{100}}{18,2297} = 3,2913.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả $\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,98/2 = 0,49$ ta được $z_{2\alpha} = 2,33$.

Bước 3: Kiểm định. Vì $z = 3,2913 > 2,33 = z_{2\alpha}$ nên ta bác bỏ giả thiết $H_0: \mu=125$, nghĩa là chấp nhận $H_1: \mu > 125$.

Kết luận: Với mức ý nghĩa 1%, có thể kết luận rằng việc canh tác làm tăng chiều cao trung bình của giống cây trồng trên.

- b) Giả sử trung bình tiêu chuẩn của chiều cao X là 134cm. Có thể khẳng định rằng việc canh tác làm giảm chiều cao trung bình của giống cây trồng trên với mức ý nghĩa 2% hay không?

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu = M(X)$ với mức ý nghĩa $\alpha = 2\% = 0,02$:

$$H_0: \mu = 134 \text{ với giả thiết đối } H_1: \mu < 134.$$

Vì $n \geq 30$; $\sigma^2 = D(X)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} = \frac{(131 - 134)\sqrt{100}}{18,2297} = -1,6457.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thoả $\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,96/2 = 0,48$ ta được $z_{2\alpha} = 2,06$.

Bước 3: Kiểm định. Vì $-z = 1,6457 < 2,06 = z_{2\alpha}$ nên ta chấp nhận giả thiết $H_0: \mu = 134$.

Kết luận: Với mức ý nghĩa 2%, không thể kết luận rằng việc canh tác làm giảm chiều cao trung bình của giống cây trồng trên.

- c) Sau khi áp dụng phương pháp canh tác mới, người ta thấy chiều cao trung bình của các cây loại A là 114cm. Hãy kết luận xem phương pháp mới có làm giảm chiều cao trung bình của các cây loại A hay không với mức ý nghĩa 3% (Giả sử X có phân phối chuẩn).

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu_A = M(X_A)$ của chỉ tiêu $X = X_A$ của các cây loại A với mức ý nghĩa $\alpha = 3\% = 0,03$:

$$H_0: \mu_A = 114 \text{ với giả thiết đối } H_1: \mu_A > 114.$$

Ta lập bảng số liệu của X_A :

X_{Ai}	110	120
N_{Ai}	10	15

Từ bảng trên ta tính được:

$$n_A = 25; \sum X_{Ai} n_{Ai} = 2900; \sum X_{Ai}^2 n_{Ai} = 337000.$$

- Kỳ vọng mẫu của X_A là

$$\bar{X}_A = \frac{1}{n} \sum X_{Ai} n_{Ai} = 116(\text{cm}).$$

- Phương sai mẫu của X_A là:

$$\hat{S}_A^2 = \frac{1}{n} \sum X_{Ai}^2 n_{Ai} - \bar{X}_A^2 = (4,8990)^2 (\text{cm}^2).$$

- Phương sai mẫu hiệu chỉnh của X_A là:

$$S_A^2 = \frac{n_A}{n_A - 1} \hat{S}_A^2 = 5^2 (\text{cm}^2).$$

Vì $n_A < 30$, X_A có phân phối chuẩn, $\sigma_A^2 = D(X_A)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(\bar{X}_A - \mu_0)\sqrt{n_A}}{S_A} = \frac{(116 - 114)\sqrt{25}}{5} = 2.$$

Bước 2: Đặt $k = n_A - 1 = 24$. Tra bảng phân phối Student ứng với $k = 24$ và $2\alpha = 0,06$ ta được $t_{2\alpha} = 1,974$.

Bước 3: Kiểm định. Vì $z = 2 > 1,974 = t_{2\alpha}$ nên ta bác bỏ giả thiết $H_0: \mu_A = 114$, nghĩa là chấp nhận $H_1: \mu_A > 114$.

Kết luận: Với mức ý nghĩa 3%, phương pháp mới làm giảm chiều cao trung bình của các cây loại A.

- d) Trước đây, chiều cao trung bình của các cây loại A là 120cm. Các số liệu trên thu thập được sau khi đã áp dụng một kỹ thuật mới. Hãy kết

luyện xem kỹ thuật mới có làm giảm chiều cao trung bình của các cây loại A hay không với mức ý nghĩa 2% (Giả sử X có phân phối chuẩn).

Đây là bài toán kiểm định giả thiết về kỳ vọng $\mu_A = M(X_A)$ của chỉ tiêu $X = X_A$ của các cây loại A với mức ý nghĩa $\alpha = 2\% = 0,02$:

$H_0: \mu_A = 120$ với giả thiết đối $H_1: \mu_A < 120$.

Vì $n_A < 30$, X_A có phân phối chuẩn, $\sigma_A^2 = D(X_A)$ chưa biết, nên ta kiểm định như sau:

Bước 1: Ta có
$$z = \frac{(\bar{X}_A - \mu_0)\sqrt{n_A}}{S_A} = \frac{(116 - 120)\sqrt{25}}{5} = -4.$$

Bước 2: Đặt $k = n_A - 1 = 24$. Tra bảng phân phối Student ứng với $k = 24$ và $2\alpha = 0,04$ ta được $t_{2\alpha} = 2,1715$.

Bước 3: Kiểm định. Vì $-z = 4 > 2,1715 = t_{2\alpha}$ nên ta bác bỏ giả thiết $H_0: \mu_A = 120$, nghĩa là chấp nhận $H_1: \mu_A < 120$.

Kết luận: Với mức ý nghĩa 2%, kỹ thuật mới làm giảm chiều cao trung bình của các cây loại A.

e) Sau khi áp dụng một phương pháp sản xuất, người ta thấy tỉ lệ cây loại A là 35%. Hãy kết luận xem phương pháp mới có làm tăng tỉ lệ cây loại A lên hay không với mức ý nghĩa 2% .

Đây là bài toán kiểm định giả thiết về tỉ lệ p các sản phẩm loại A với mức ý nghĩa $\alpha = 2\% = 0,02$:

$H_0: p = 35\% = 0,35$ với giả thiết đối $H_1: p < 0,35$.

Ta có tỉ lệ mẫu các cây loại A là $F_n = 25/100 = 0,25$. Ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(F_n - p_0)\sqrt{n}}{\sqrt{p_0q_0}} = \frac{(0,25 - 0,35)\sqrt{100}}{\sqrt{0,35(1 - 0,35)}} = -2,0966.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thỏa

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,96/2 = 0,48$$

ta được $z_{2\alpha} = 2,06$.

Bước 3: Kiểm định. Vì $-z = 2,0966 > 2,06 = z_{2\alpha}$ nên ta bác bỏ giả thiết $H_0: p = 0,35$, nghĩa là chấp nhận $H_1: p < 0,35$.

Kết luận: Với mức ý nghĩa 2%, phương pháp mới làm tăng tỉ lệ cây loại A.

f) Một tài liệu thống kê cũ cho rằng tỉ lệ cây loại A là 20%. Hãy xét xem việc canh tác có làm tăng tỉ lệ cây loại A hay không với mức ý nghĩa 5%?

Đây là bài toán kiểm định giả thiết về tỉ lệ p các sản phẩm loại A với mức ý nghĩa $\alpha = 5\% = 0,05$:

$H_0: p = 20\% = 0,20$ với giả thiết đối $H_1: p > 0,20$.

Ta kiểm định như sau:

Bước 1: Ta có

$$z = \frac{(F_n - p_0)\sqrt{n}}{\sqrt{p_0q_0}} = \frac{(0,25 - 0,20)\sqrt{100}}{\sqrt{0,20(1 - 0,20)}} = 1,25.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thỏa

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,90/2 = 0,45$$

ta được $z_{2\alpha} = 1,65$.

Bước 3: Kiểm định. Vì $z = 1,25 < 1,65 = z_{2\alpha}$ nên ta chấp nhận giả thiết $H_0: p = 0,20$.

Kết luận: Với mức ý nghĩa 5%, việc canh tác không làm tăng tỉ lệ các cây loại A.

Bài 4.15. Để khảo sát đường kính của một chi tiết máy người ta kiểm tra một số sản phẩm của hai nhà máy. Trong kết quả sau đây, X là đường kính của chi tiết máy do nhà máy 1 sản xuất còn Y là đường kính của chi tiết máy do nhà máy 2 sản xuất. Những sản phẩm có chi tiết máy nhỏ hơn 19cm được xếp vào loại C.

X(cm)	11-15	15-19	19-23	23-27	27-31	31-35	35-39
Số sản phẩm	9	19	20	26	16	13	18

Y(cm)	13-16	16-19	19-22	22-25	25-28	28-31	31-34
Số sản phẩm	7	9	25	26	18	15	11

- Có thể kết luận rằng đường kính trung bình của một chi tiết máy do hai nhà máy sản xuất bằng nhau hay không với mức ý nghĩa 1%?
- Có thể cho rằng đường kính trung bình của một chi tiết máy do nhà máy thứ 1 sản xuất lớn hơn đường kính trung bình của một chi tiết máy do nhà máy thứ 2 sản xuất hay không với mức ý nghĩa 5%?
- Xét xem đường kính trung bình của một chi tiết máy do nhà máy thứ 2 sản xuất có nhỏ hơn đường kính trung bình của một chi tiết máy do nhà máy thứ 1 sản xuất hay không với mức ý nghĩa 2%?
- Với mức ý nghĩa 4%, tỉ lệ sản phẩm loại C do hai nhà máy sản xuất có như nhau không?
- Với mức ý nghĩa 3%, có thể cho rằng tỉ lệ sản phẩm loại C do nhà máy thứ 1 sản xuất lớn hơn tỉ lệ sản phẩm loại C do nhà máy thứ 2 sản xuất hay không?

- f) Hãy nhận xét về ý kiến cho rằng tỉ lệ sản phẩm loại C do nhà máy thứ 2 sản xuất nhỏ hơn tỉ lệ sản phẩm loại C do nhà máy thứ 1 sản xuất với mức ý nghĩa 5%?

Lời giải

- 1) Đối với X ta có bảng số liệu:

X_i	13	17	21	25	29	33	37
n_i	9	19	20	26	16	13	18

Ta có:

$$n_X = 121; \sum X_i n_{X_i} = 3069; \sum X_i^2 n_{X_i} = 84337.$$

- Kỳ vọng mẫu của X là

$$\bar{X} = \frac{1}{n_X} \sum X_i n_{X_i} = 25,3636(\text{cm}).$$

- Phương sai mẫu của X là

$$\hat{S}_X^2 = \frac{1}{n_X} \sum X_i^2 n_{X_i} - \bar{X}^2 = (7,3271)^2 (\text{cm}^2).$$

- Phương sai mẫu hiệu chỉnh của X là

$$S_X^2 = \frac{n_X}{n_X - 1} \hat{S}_X^2 = (7,3575)^2 (\text{cm}^2).$$

- Tỉ lệ sản phẩm loại C là

$$F_{Xn} = \frac{m_X}{n_X} = \frac{9+19}{121} = 0,2314.$$

- 2) Đối với Y ta có bảng số liệu:

Y_i	14,5	17,5	20,5	23,5	26,5	29,5	32,5
n_i	7	9	25	26	18	15	11

Ta có:

$$n_Y = 111; \sum Y_i n_{Y_i} = 2659,5; \sum Y_i^2 n_{Y_i} = 66405,75.$$

- Kỳ vọng mẫu của Y là

$$\bar{Y} = \frac{1}{n_Y} \sum Y_i n_{Y_i} = 23,9595(\text{cm}).$$

- Phương sai mẫu của Y là

$$\hat{S}_Y^2 = \frac{1}{n_Y} \sum Y_i^2 n_{Y_i} - \bar{Y}^2 = (4,9188)^2 (\text{cm}^2).$$

- Phương sai mẫu hiệu chỉnh của Y là

$$S_Y^2 = \frac{n_Y}{n_Y - 1} \hat{S}_Y^2 = (4,9411)^2 (\text{cm}^2).$$

- Tỉ lệ sản phẩm loại C là

$$F_{Yn} = \frac{m_Y}{n_Y} = \frac{7+9}{111} = 0,1441.$$

- a) Có thể kết luận rằng đường kính trung bình của một chi tiết máy do hai nhà máy sản xuất bằng nhau hay không với mức ý nghĩa 1%?

Đây là bài toán kiểm định so sánh hai kỳ vọng với mức ý nghĩa $\alpha = 1\% = 0,01$:

$$H_0: \mu_X = \mu_Y \text{ với giả thiết đối } H_1: \mu_X \neq \mu_Y.$$

Vì $n_X > 30$; $n_Y > 30$ nên ta kiểm định như sau:

Bước 1: Ta có:

$$z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} = \frac{25,3636 - 23,9595}{\sqrt{\frac{(7,3575)^2}{121} + \frac{(4,9411)^2}{111}}} = 1,7188.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thỏa

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0,99/2 = 0,495$$

ta được $z_\alpha = 2,58$.

Bước 3: Kiểm định. Vì $|z| = 1,7188 < 2,58 = z_\alpha$ nên ta chấp nhận giả thiết $H_0: \mu_X = \mu_Y$.

Kết luận: Với mức ý nghĩa 1%, có thể xem đường kính trung bình của một chi tiết máy do hai nhà máy sản xuất là bằng nhau.

- b) Có thể cho rằng đường kính trung bình của một chi tiết máy do nhà máy thứ 1 sản xuất lớn hơn đường kính trung bình của một chi tiết máy do nhà máy thứ 2 sản xuất hay không với mức ý nghĩa 5%?

Đây là bài toán kiểm định so sánh hai kỳ vọng với mức ý nghĩa $\alpha = 5\% = 0,05$:

$$H_0: \mu_X = \mu_Y \text{ với giả thiết đối } H_1: \mu_X > \mu_Y.$$

Vì $n_X > 30$; $n_Y > 30$ nên ta kiểm định như sau:

Bước 1: Tương tự câu a), ta có:

$$z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} = 1,7188.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thỏa

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,90/2 = 0,45$$

ta được $z_{2\alpha} = 1,65$.

Bước 3: Kiểm định. Vì $z = 1,7188 > 1,65 = z_{2\alpha}$ nên ta bác bỏ giả thiết $H_0: \mu_X = \mu_Y$, nghĩa là chấp nhận $H_1: \mu_X > \mu_Y$.

Kết luận: Với mức ý nghĩa 5%, có thể xem đường kính trung bình của một chi tiết máy do nhà máy thứ 1 sản xuất lớn hơn đường kính trung bình của một chi tiết máy do nhà máy thứ 2 sản xuất.

c) Xét xem đường kính trung bình của một chi tiết máy do nhà máy thứ 2 sản xuất có nhỏ hơn đường kính trung bình của một chi tiết máy do nhà máy thứ 1 sản xuất hay không với mức ý nghĩa 2%?

Đây là bài toán kiểm định so sánh hai kỳ vọng với mức ý nghĩa $\alpha = 2\% = 0,02$:

$H_0: \mu_X = \mu_Y$ với giả thiết đối $H_1: \mu_X > \mu_Y$.

Vì $n_X > 30$; $n_Y > 30$ nên ta kiểm định như sau:

Bước 1: Tương tự câu a), ta có:

$$z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} = 1,7188.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thỏa

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,96/2 = 0,48$$

ta được $z_{2\alpha} = 2,06$.

Bước 3: Kiểm định. Vì $z = 1,7188 < 2,06 = z_{2\alpha}$ nên ta chấp nhận giả thiết $H_0: \mu_X = \mu_Y$.

Kết luận: Với mức ý nghĩa 2%, chưa thể xem đường kính trung bình của một chi tiết máy do nhà máy thứ 2 sản xuất nhỏ hơn đường kính trung bình của một chi tiết máy do nhà máy thứ 1 sản xuất.

d) Với mức ý nghĩa 4%, tỉ lệ sản phẩm loại C do hai nhà máy sản xuất có như nhau không?

Đây là bài toán kiểm định giả thiết so sánh hai tỉ lệ với mức ý nghĩa $\alpha = 4\% = 0,04$:

$H_0: p_1 = p_2$ với giả thiết đối $H_1: p_1 \neq p_2$

Ta kiểm định như sau:

Bước 1: Ta có:

$$p_0 = \frac{n_1 F_{n1} + n_2 F_{n2}}{n_1 + n_2} = \frac{28 + 16}{121 + 111} = 0,1897.$$

$$z = \frac{F_{n1} - F_{n2}}{\sqrt{p_0(1-p_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0,2314 - 0,1441}{\sqrt{0,1897(1-0,1897)\left(\frac{1}{121} + \frac{1}{111}\right)}} = 1,6942.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm z_α thỏa

$$\varphi(z_\alpha) = (1 - \alpha)/2 = 0,96/2 = 0,48$$

ta được $z_\alpha = 2,06$.

Bước 3: Kiểm định. Vì $|z| = 1,6942 < 2,06 = z_\alpha$ nên ta chấp nhận giả thiết $H_0: p_1 = p_2$.

Kết luận: Với mức ý nghĩa 4%, có thể xem tỉ lệ sản phẩm loại C do hai nhà máy sản xuất là như nhau.

e) Với mức ý nghĩa 3%, có thể cho rằng tỉ lệ sản phẩm loại C do nhà máy thứ 1 sản xuất lớn hơn tỉ lệ sản phẩm loại C do nhà máy thứ 2 sản xuất hay không?

Đây là bài toán kiểm định giả thiết so sánh hai tỉ lệ với mức ý nghĩa $\alpha = 3\% = 0,03$:

$H_0: p_1 = p_2$ với giả thiết đối $H_1: p_1 > p_2$

Ta kiểm định như sau:

Bước 1: Tương tự câu d), ta có:

$$z = \frac{F_{n1} - F_{n2}}{\sqrt{p_0(1-p_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 1,6942.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thỏa

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,94/2 = 0,47$$

ta được $z_{2\alpha} = 1,88$.

Bước 3: Kiểm định. Vì $z = 1,6942 < 1,88 = z_{2\alpha}$ nên ta chấp nhận giả thiết $H_0: p_1 = p_2$.

Kết luận: Với mức ý nghĩa 3%, chưa thể cho rằng tỉ lệ sản phẩm loại C do nhà máy thứ 1 sản xuất lớn hơn tỉ lệ sản phẩm loại C do nhà máy thứ 2 sản xuất.

f) Hãy nhận xét về ý kiến cho rằng tỉ lệ sản phẩm loại C do nhà máy thứ 2 sản xuất nhỏ hơn tỉ lệ sản phẩm loại C do nhà máy thứ 1 sản xuất với mức ý nghĩa 5%?

Đây là bài toán kiểm định giả thiết so sánh hai tỉ lệ với mức ý nghĩa $\alpha = 5\% = 0,05$:

$H_0: p_1 = p_2$ với giả thiết đối $H_1: p_1 > p_2$

Ta kiểm định như sau:

Bước 1: Tương tự câu d), ta có:

$$z = \frac{F_{n1} - F_{n2}}{\sqrt{p_0(1-p_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 1,6942.$$

Bước 2: Tra bảng giá trị hàm Laplace để tìm $z_{2\alpha}$ thỏa

$$\varphi(z_{2\alpha}) = (1 - 2\alpha)/2 = 0,90/2 = 0,45$$

ta được $z_{2\alpha} = 1,65$.

Bước 3: Kiểm định. Vì $z = 1,6942 > 1,65 = z_{2\alpha}$ nên ta bác bỏ giả thiết $H_0: p_1 = p_2$, nghĩa là chấp nhận $H_1: p_1 > p_2$.

Kết luận: Với mức ý nghĩa 5%, có thể chấp nhận ý kiến cho rằng tỉ lệ sản phẩm loại C do nhà máy thứ 2 sản xuất nhỏ hơn tỉ lệ sản phẩm loại C do nhà máy thứ 1 sản xuất.