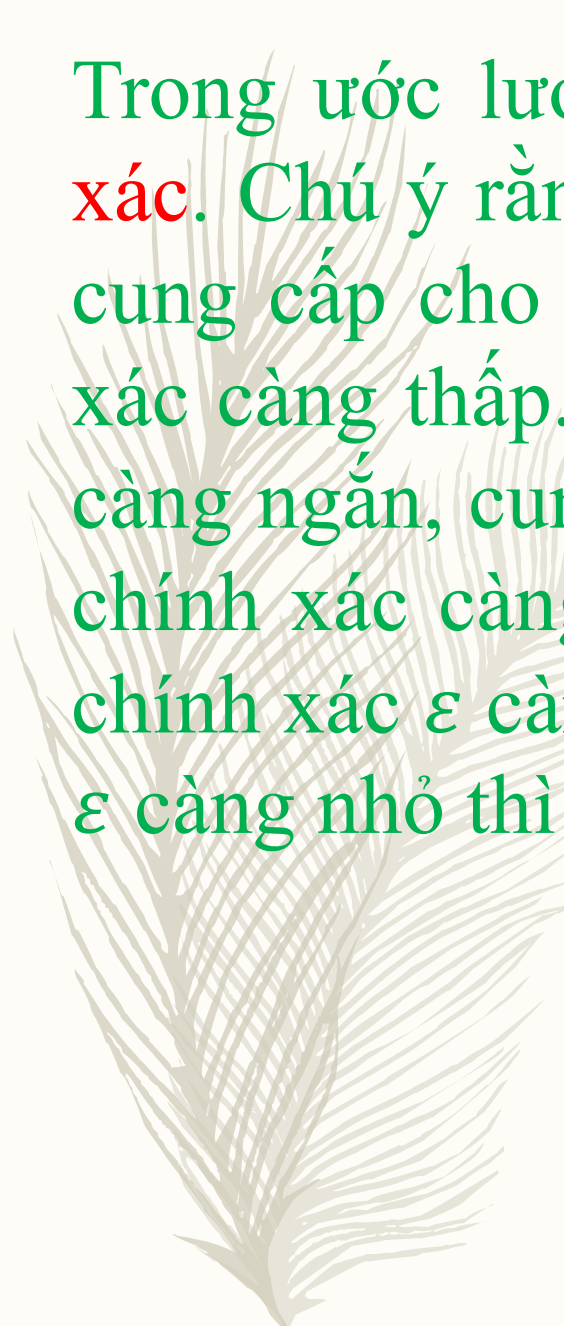


Bài 3. Ước lượng khoảng

3.1. Định nghĩa. Xét tổng thể X có đặc số θ chưa biết (θ có thể là kỳ vọng, phương sai, hay tỉ lệ) và mẫu X_1, X_2, \dots, X_n . Gọi $\hat{\theta}$ là ước lượng điểm của θ . Với mỗi $\gamma = 1 - \alpha$ với $0 < \alpha < 1$, α khá bé, ước lượng khoảng của θ với độ tin cậy γ là khoảng $(\theta - \varepsilon; \theta + \varepsilon)$ với $\varepsilon > 0$, sao cho

$$P(\hat{\theta} - \varepsilon \leq \theta \leq \hat{\theta} + \varepsilon) = \gamma.$$



Trong ước lượng khoảng $(\hat{\theta} - \varepsilon; \hat{\theta} + \varepsilon)$, ta gọi ε là **độ chính xác**. Chú ý rằng nếu ε càng lớn thì ước lượng khoảng càng dài, cung cấp cho chúng ta ít thông tin về đặc số θ nên tính chính xác càng thấp. Ngược lại nếu ε càng nhỏ thì ước lượng khoảng càng ngắn, cung cấp cho ta nhiều thông tin về đặc số θ nên tính chính xác càng cao. Như vậy, trong ước lượng khoảng, nếu độ chính xác ε càng lớn thì độ tin cậy γ càng cao. Nếu độ chính xác ε càng nhỏ thì độ tin cậy γ càng thấp.

2.2. Ước lượng khoảng cho kỳ vọng. Xét tổng thể X có kỳ vọng $\mu = E(X)$ chưa biết. Với mỗi $\gamma = 1 - \alpha$ ($0 < \alpha < 1$, α khá bé), hãy dựa vào mẫu (X_1, X_2, \dots, X_n) để đưa ra ước lượng khoảng cho kỳ vọng μ : $(\bar{X} - \varepsilon; \bar{X} + \varepsilon)$, $\varepsilon > 0$ với độ tin cậy γ , nghĩa là

$$P(\bar{X} - \varepsilon \leq \mu \leq \bar{X} + \varepsilon) = \gamma,$$

trong đó \bar{X} là kỳ vọng mẫu.

Trường hợp

Công thức

$$n \geq 30$$

$$\left(\underbrace{\bar{X} - z_{\alpha} \frac{S}{\sqrt{n}}}_{\varepsilon}; \underbrace{\bar{X} + z_{\alpha} \frac{S}{\sqrt{n}}}_{\varepsilon} \right)$$

$n < 30$ và X có phân
phối chuẩn

$$\left(\underbrace{\bar{X} - t_{\alpha} \frac{S}{\sqrt{n}}}_{\varepsilon}; \underbrace{\bar{X} + t_{\alpha} \frac{S}{\sqrt{n}}}_{\varepsilon} \right)$$

- z_{α} thỏa mãn $\varphi(z_{\alpha}) = \frac{(1-\alpha)}{2} = \frac{\gamma}{2}$ được tra từ bảng giá trị hàm Laplace.
- t_{α} được tra từ bảng giá trị hàm Student ứng với $k = n - 1$ và $\alpha = 1 - \gamma$.

Ví dụ. Để khảo sát chỉ tiêu của một loại sản phẩm X, người ta chọn ngẫu nhiên một mẫu và thu được 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 19 (cm) trở xuống được xếp loại B.

- Hãy ước lượng giá trị trung bình của của chỉ tiêu X với độ tin cậy 95%.
- Hãy ước lượng giá trị trung bình của chỉ tiêu X của những sản phẩm loại B với độ tin cậy 99% (giả sử X có phân phối chuẩn).

Giải. a) Đây là bài toán ước lượng khoảng cho kỳ vọng $\mu = E(X)$ với độ tin cậy $\gamma = 1 - \alpha = 95\% = 0,95$. Từ bảng số liệu ta có: cỡ mẫu $n = 100$; $\bar{X} = 26,36$; $S^2 = (7,4827)^2$. Vì $n \geq 30$ nên ta có công thức ước lượng khoảng cho kỳ vọng như sau

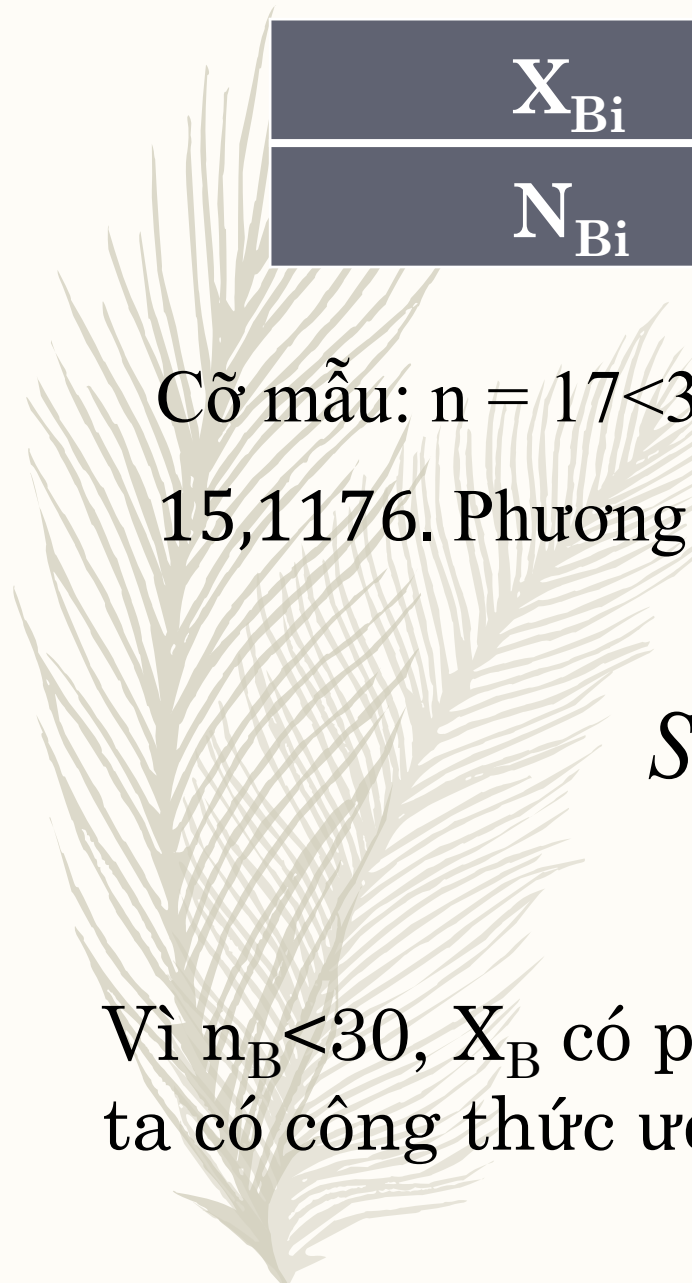
$$\left(\underbrace{\bar{X} - z_{\alpha} \frac{S}{\sqrt{n}}}_{\varepsilon}; \underbrace{\bar{X} + z_{\alpha} \frac{S}{\sqrt{n}}}_{\varepsilon} \right)$$

trong đó $\varphi(z_{\alpha}) = (1 - \alpha)/2 = 0,95/2 = 0,475$. Tra bảng giá trị hàm Laplace ta được $z_{\alpha} = 1,96$. Vậy ước lượng khoảng cho kỳ vọng là

$$\left(26,36 - 1,96 \frac{7,4827}{\sqrt{100}}; 26,36 + 1,96 \frac{7,4827}{\sqrt{100}} \right) = (24,89; 27,83).$$

Nói cách khác, với độ tin cậy 95% giá trị trung bình của chỉ tiêu X của tổng thể nằm trong khoảng từ 24,89 đến 27,83.

b) Đây là bài toán ước lượng khoảng cho kỳ vọng $\mu_B = E(X_B)$ với chỉ tiêu đang xét bây giờ là X_B , những sản phẩm loại B, với độ tin cậy $\gamma = 1 - \alpha = 99\% = 0,99$. Từ bảng số liệu đã cho ta có bảng số liệu sau đây

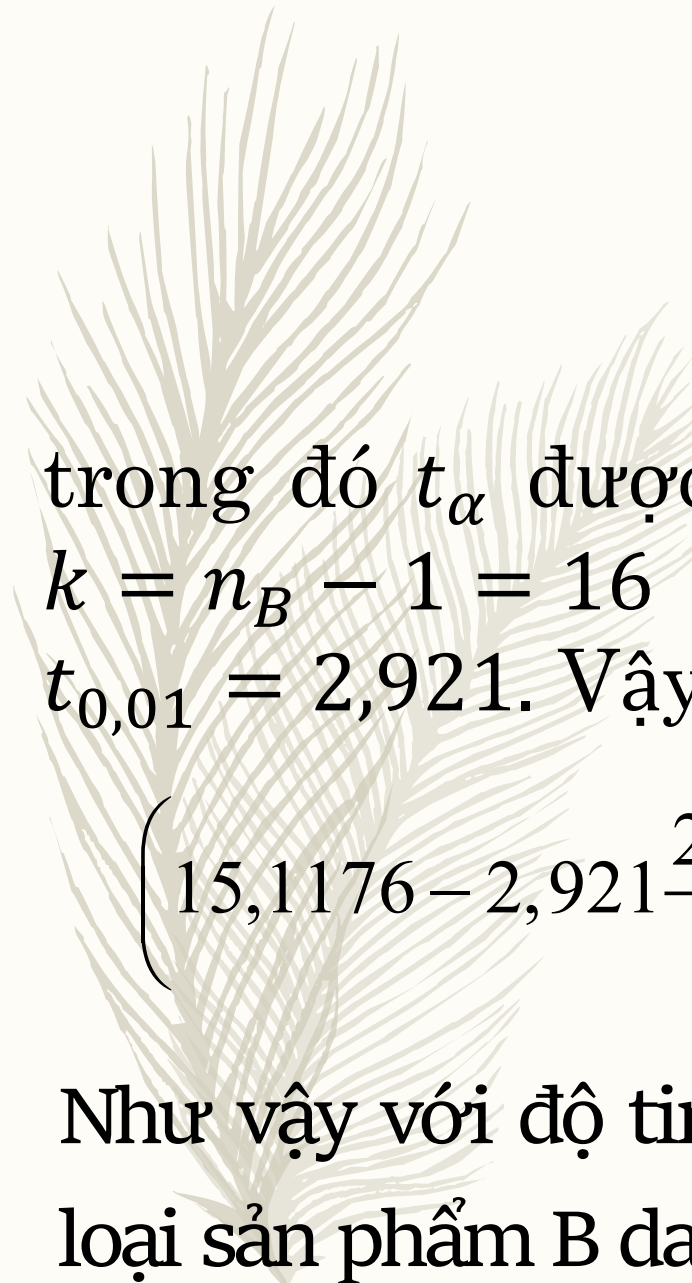


X_{Bi}	13	17
N_{Bi}	8	9

Cỡ mẫu: $n = 17 < 30$. Kỳ vọng của mẫu $\bar{X}_B = \left(\frac{1}{n_B}\right) \sum X_{Bi}n_{Bi} = 15,1176$. Phương sai mẫu của X_B là

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

Vì $n_B < 30$, X_B có phân phối chuẩn, $\sigma_B^2 = Var(X_B)$ chưa biết nên ta có công thức ước lượng khoảng cho kỳ vọng


$$\left(\underbrace{\bar{X} - t_{\alpha} \frac{S}{\sqrt{n}}}_{\varepsilon}; \underbrace{\bar{X} + t_{\alpha} \frac{S}{\sqrt{n}}}_{\varepsilon} \right)$$

trong đó t_{α} được tra từ bảng giá trị tới hạn Student với $k = n_B - 1 = 16$ và $\alpha = 1 - \gamma = 0,01$. Tra bảng ta được $t_{0,01} = 2,921$. Vậy ước lượng khoảng là

$$\left(15,1176 - 2,921 \frac{2,0580}{\sqrt{17}}; 15,1176 + 2,921 \frac{2,0580}{\sqrt{17}} \right) = (13,66; 16,58).$$

Như vậy với độ tin cậy 99%, giá trị trung bình của chỉ tiêu X của loại sản phẩm B dao động từ 13,66 đến 16,58.

3.3. Ước lượng khoảng cho tỉ lệ. Xét tổng thể X có tỉ lệ p chưa biết. Với mỗi $\gamma = 1 - \alpha$ ($0 < \alpha < 1, \alpha$ khá bé), hãy dựa vào mẫu (X_1, X_2, \dots, X_n) để đưa ra ước lượng khoảng cho tỉ lệ p : $(F_n - \varepsilon; F_n + \varepsilon)$, $\varepsilon > 0$ với độ tin cậy γ , nghĩa là

$$P(F_n - \varepsilon \leq p \leq F_n + \varepsilon) = \gamma,$$

Trong đó F_n là tỉ lệ mẫu.

Ước lượng khoảng cho tỉ lệ $p = P(A)$ (độ tin cậy $\gamma = 1 - \alpha$)

$$\left(F_n - \underbrace{z_\alpha \sqrt{\frac{F_n(1-F_n)}{n}}}_{\varepsilon}; F_n + \underbrace{z_\alpha \sqrt{\frac{F_n(1-F_n)}{n}}}_{\varepsilon} \right)$$

z_α thỏa $\varphi(z_\alpha) = \frac{1-\alpha}{2} = \frac{\gamma}{2}$ tra từ bảng giá trị hàm Laplace.

Ví dụ. Để khảo sát trọng lượng của một loại vật nuôi, 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 con có trọng lượng từ 145 kg trở lên được xếp loại A. Hãy ước lượng tỉ lệ con vật loại A với độ tin cậy 97%.

Giải. Đây là bài toán ước lượng tỉ lệ con loại A với độ tin cậy $\gamma = 1 - \alpha = 0,97$. Ta có công thức ước lượng khoảng

$$\left(F_n - z_\alpha \sqrt{\frac{F_n(1-F_n)}{n}}; F_n + z_\alpha \sqrt{\frac{F_n(1-F_n)}{n}} \right)$$

Trong đó $\varphi(z_\alpha) = \frac{\gamma}{2} = 0,485$. Tra bảng giá trị hàm Laplace ta được $z_\alpha = 2,17$. Cỡ mẫu $n = 189$, trong đó có 15 con có trọng lượng từ 145 kg trở lên. Do đó tỉ lệ mẫu các con loại A là

$$F_n = \frac{m}{n} = \frac{15}{189} = 0,0794.$$

Vậy ước lượng khoảng là

$$\left(0,0794 - 2,17\sqrt{\frac{0,0794(1-0,0794)}{189}}; 0,0794 + 2,17\sqrt{\frac{0,0794(1-0,0794)}{189}} \right) \\ = (0,0367; 0,1221) = (3,67\%; 12,21\%).$$

Nói cách khác, tỉ lệ con vật loại A ở độ chính xác 97% thì dao động từ 3,67% đến 12,21%.

Trong bài toán ước lượng khoảng đối xứng có 3 chỉ tiêu chính là

- Cỡ mẫu n
- Độ chính xác ε
- Độ tin cậy $\gamma = 1 - \alpha$.

3.2. Trường hợp ước lượng khoảng cho kỳ vọng. Ta xét trường hợp phổ biến nhất là $n \geq 30$; $\sigma^2 = \text{Var}(X)$ chưa biết. Khi đó, ta có công thức ước lượng khoảng cho kỳ vọng $\mu = E(X)$ với độ tin cậy $\gamma = 1 - \alpha$

$$\left(\bar{X} - z_{\alpha} \frac{S}{\sqrt{n}}; \bar{X} + z_{\alpha} \frac{S}{\sqrt{n}} \right); \varphi(z_{\alpha}) = \frac{1 - \alpha}{2} = \frac{\gamma}{2}.$$

Do đó ta có công thức ước lượng với độ chính xác là

$$\varepsilon = z_{\alpha} \frac{S}{\sqrt{n}}$$

- Nếu biết cỡ mẫu n và độ tin cậy γ ta có thể tính được độ chính xác ε .
- Nếu biết cỡ mẫu n và độ chính xác ε ta có thể suy ra độ tin cậy γ .

$$z_{\alpha} = \frac{\varepsilon \sqrt{n}}{S}$$

Tra từ bảng giá trị hàm Laplace ta tìm được $\varphi(z_{\alpha})$. Từ đó suy ra độ tin cậy $\gamma = 2\varphi(z_{\alpha})$.

Xác định các chỉ tiêu chính

Trường hợp ước lượng khoảng cho kỳ vọng $\mu = E(X)$

Chỉ tiêu đã biết	Chỉ tiêu cần tìm	Công thức
Cỡ mẫu n , độ tin cậy $\gamma = 1 - \alpha$	Độ chính xác ε	$\varepsilon = z_{\alpha} \frac{S}{\sqrt{n}}$
Cỡ mẫu n , độ chính xác ε	Độ tin cậy $\gamma = 1 - \alpha$	$\gamma = 2\varphi\left(\frac{\varepsilon\sqrt{n}}{S}\right)$
Độ tin cậy $\gamma = 1 - \alpha$, độ chính xác ε	Cỡ mẫu n	$n \geq \left\lceil \left(\frac{z_{\alpha}S}{\varepsilon}\right)^2 \right\rceil$

- z_{α} thỏa $\varphi(z_{\alpha}) = \frac{1-\alpha}{2} = \frac{\gamma}{2}$ được tra từ bảng giá trị hàm Laplace

- $\left\lceil \left(\frac{z_{\alpha}S}{\varepsilon}\right)^2 \right\rceil$ là số nguyên nhỏ nhất lớn hơn hoặc bằng $\left(\frac{z_{\alpha}S}{\varepsilon}\right)^2$

Ví dụ. Để khảo sát một loại chỉ tiêu X người ta quan sát một mẫu và thu được 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

- a) Nếu muốn ước lượng giá trị trung bình của chỉ tiêu X của loại sản phẩm trên với độ chính xác là 1,8 cm thì độ tin cậy sẽ là bao nhiêu?
- b) Nếu muốn ước lượng giá trị trung bình của chỉ tiêu X của loại sản phẩm trên với độ chính xác là 1,5 cm và độ tin cậy là 97% thì phải điều tra thêm ít nhất bao nhiêu sản phẩm nữa?

Giải. Ta tìm được cỡ mẫu $n = 100$; $\bar{X} = 26,36 \text{ cm}$; $S^2 = (7,4827)^2 \text{ (cm}^2\text{)}$.

Yêu cầu bài toán xác định độ tin cậy $\gamma = 1 - \alpha$. Từ giả thiết ta có trung bình mẫu và độ chính xác. Ta có

$$\gamma = 2\varphi\left(\frac{\varepsilon\sqrt{n}}{S}\right) = 2\varphi(2,41) = 2 \times 0,4920 = 98,40\%.$$

b) Yêu cầu bài toán xác định cỡ mẫu n . Biết độ chính xác $\varepsilon = 1,5 \text{ cm}$ độ tin cậy $\gamma = 1 - \alpha = 97\% = 0,97$. Vì $n \geq 30, \sigma^2 = \text{Var}(X)$ chưa biết nên ta có công thức tính độ chính xác của ước lượng

$$\varepsilon = z_{\alpha} \frac{S}{\sqrt{n}}$$

trong đó $\varphi(z_{\alpha}) = \frac{1-\alpha}{2} = \frac{0,97}{2} = 0,485$. Tra bảng giá trị hàm Laplace ta được $z_{\alpha} = 2,17$. Suy ra

$$n = \left(\frac{z_{\alpha} S}{\varepsilon} \right)^2 = \left(\frac{2,17 \times 7,4827}{1,5} \right)^2 \approx 117,18.$$

Thực tế yêu cầu

$$n \geq \lceil 117,18 \rceil = 118.$$

Vì $n_1 = 118 > 100$ (100 là cỡ mẫu đang có) nên ta cần điều tra thêm ít nhất $118 - 100 = 18$ sản phẩm nữa.

3.3. Trường hợp ước lượng khoảng cho tỉ lệ. Ta chỉ xét trường hợp n khá lớn $n \geq 30$. Khi đó ta có công thức ước lượng khoảng cho tỉ lệ p với độ tin cậy $\gamma = 1 - \alpha$

$$\left(F_n - z_\alpha \sqrt{\frac{F_n(1-F_n)}{n}}; F_n + z_\alpha \sqrt{\frac{F_n(1-F_n)}{n}} \right)$$

Với

$$\varphi(z_\alpha) = \frac{1-\alpha}{2} = \frac{\gamma}{2}.$$

$$\varepsilon = z_\alpha \sqrt{\frac{F_n(1-F_n)}{n}}$$

Ví dụ. Để khảo sát một loại chỉ tiêu X người ta quan sát một mẫu và thu được kết quả sau

$X(\text{cm})$	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 19 cm trở xuống được xếp loại B.

- Nếu muốn ước lượng tỉ lệ sản phẩm loại B với độ chính xác 8% thì phải đạt được độ tin cậy là bao nhiêu?
- Nếu muốn ước lượng tỉ lệ sản phẩm loại B với độ chính xác 9% và độ tin cậy là 96% thì phải điều tra thêm ít nhất bao nhiêu sản phẩm nữa?



Giải.

a) Vậy độ tin cậy đạt được là 98,68%.

b) Từ thực tế ta suy ra $n \geq 74$. Vì cỡ mẫu đang có là 100 sản phẩm nên ta không cần phải điều tra thêm.