Chương 4:

ƯỚC LƯỢNG THAM SỐ CỦA TỔNG THỂ

I. Bài toán ước lượng

Cho một tổng thể có 1 tham số θ chưa biết:

$$\theta : TB, TL, PS$$

Ta chọn một mẫu để ước lượng cho θ (*Theta*), khi đó ta có 02 phương pháp:

1. Ước lượng điểm

Ta cần chọn ra: $\theta := \theta_0$

Ví dụ: Chọn ra một mẫu, tính:

- $\overline{x}$: ước lượng Trung bình của Tổng thể.
- f : ước lượng Tỷ lệ của Tổng thể.
- s^2 : ước lượng Phương sai của Tổng thể

Nhận xét: Ước lượng điểm dù tốt nhất cũng chỉ cho ta biết một giá trị trong tập vô hạn các giá trị.

- Dễ thực hiện.
- Không xác định được xác suất ước lượng đúng là bao nhiêu?

$\Rightarrow$Do đó không đánh giá được mức độ sai lầm khi ta dùng θ_0 thay cho θ.

2. Ước lượng khoảng

Ta cần chỉ ra $(\theta_1 ; \theta_2)$ sao cho:

$$P(\theta_1 < \theta < \theta_2) = 1 - \alpha$$

$1-\alpha$: độ tin cậy của ước lượng $(90\% < 1-\alpha < 100\%)$

α : mức ý nghĩa, là khả năng có thể mắc phải sai lầm khi ước lượng.

$(\theta_1 ; \theta_2)$: khoảng ước lượng hay khoảng tin cậy của ước lượng.

$|\theta_2 - \theta_1|$: độ dài khoảng tin cậy của ước lượng.

Ví dụ: Ước lượng điểm môn xác suất thống kê

- Điểm: 4, 6, 7, ...
- Khoảng: (4 – 6), (4 – 8), (0 – 10), ...

II. Ước lượng tỷ lệ

Giả sử một tổng thể có hai loại phần tử bao gồm loại phần tử có tính chất *A* và loại phần tử không có tính chất *A*. Khi đó:

Gọi *p* là tỷ lệ phần tử có tính chất *A* của tổng thể đã cho.

<u>Giải</u>

Giả sử *p* chưa biết. Ta cần ước lượng *p* với độ tin cậy $1-\alpha$.

Bước 1: Chọn một mẫu có kích thước *n* tính *f*

Bước 2: Tính sai số ước lượng (*Độ chính xác*)

$$\varepsilon = Z_{1-\frac{\alpha}{2}}\sqrt{\frac{f(1-f)}{n}}$$

** Z_{α} gọi là phân vị chuẩn và các giá trị thường dùng được xác định trong **Phụ lục 2** (Trang 182).*

Bước 3: Khoảng ước lượng cần tìm

$$(f-\varepsilon; f+\varepsilon)$$

Ví dụ 1: Trong một đợt kiểm tra y tế ở một địa phượng người ta chọn ngẫu nhiên 300 em bé, thì thấy có 180 em bé bị suy dinh dưỡng.

a) Hãy ước lượng tỷ lệ suy dinh dưỡng ở địa phương đã cho với độ tin cậy 95%.

b) Nếu muốn độ chính xác của ước lượng trên là 0.04 thì độ tin cậy của ước lượng là bao nhiêu?

<u>Giải</u>

Trẻ em ở địa phương - Bị suy dinh dưỡng: 180

- Không bị suy dinh dưỡng: 120

- Gọi *p* là tỷ lệ trẻ suy dinh dưỡng ở địa phương đã cho.

a) Ta cần ước lượng *p* với độ tin cậy $1-\alpha = 95\% \Leftrightarrow \alpha = 0.05$

Bước 1: Phân tích mẫu có: *n = 300* và trẻ em bị suy dinh dưỡng ở địa phương có 180

$$\Rightarrow f = \frac{180}{300} = 0.6$$

Bước 2: Tính sai số ước lượng

$$\varepsilon = Z_{1-\frac{\alpha}{2}}\sqrt{\frac{f(1-f)}{n}} = Z_{0.975}\sqrt{\frac{f(1-f)}{n}} = 1.96\sqrt{\frac{0.6(1-0.6)}{300}} = 0.055$$

Bước 3: Khoảng ước lượng cần tìm:

$$(f-\varepsilon; f+\varepsilon) = (0.6-0.055; 0.6+0.055) = (0.545; 0.655) = (54.5\%; 65.5\%)$$

b) Tìm độ tin cậy của ước lượng

Từ công thức sai số ước lượng

$$\varepsilon = z_{1-\frac{\alpha}{2}}\sqrt{\frac{f(1-f)}{n}} \Rightarrow z_{1-\frac{\alpha}{2}} = \varepsilon\sqrt{\frac{n}{f(1-f)}}$$

Ta có: $\varepsilon = 0.04, n = 300, f = 0.6 \Rightarrow z_{1-\frac{\alpha}{2}} = 0.04\sqrt{\frac{300}{0.6(1-0.6)}} = 1.414 \approx z_{0.92}$

Khi đó,

$$1-\frac{\alpha}{2}=0.92 \Rightarrow 1-\alpha=0.84$$

Vậy độ tin cậy của ước lượng là 84%.

Ví dụ 2: Để kiểm tra chất lượng của một kho hàng. Người ta chọn ngẫu nhiên ra 1000 sản phẩm, thì thấy 70 phế phẩm. Với độ tin cậy 90%, hãy ước lượng tỷ lệ phế phẩm của kho hàng đã cho.

III. Ước lượng Trung bình

Giả sử một tổng thể có kích thước *N*. Gọi *m* là trung bình của một dấu hiệu nào đó của tổng thể.

Giả sử *m* chưa biết, ta cần ước lượng *m* với độ tin cậy $1-\alpha$.

Bước 1: Chọn một mẫu có kích thước *n*, tính $\overline{x}$ và s.

Bước 2: Tính sai số ước lượng (*Độ chính xác*)

Trường hợp 1: Khi phương sai đã biết ($V(X)=\sigma^2$) và kích thước mẫu $n \geq 30$ hoặc $n<30$

$$\varepsilon = Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

Trường hợp 2: Phương sai chưa biết và kích thước mẫu $n \geq 30$.

$$\varepsilon = Z_{1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

Trường hợp 3: Phương sai chưa biết và kích thước mẫu $n<30$.

$$\varepsilon = t_{\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}}$$

* $t_{\alpha}(n)$ *phân vị Student và các giá trị thường dùng được xác định trong bảng* ***Phụ lục 4*** *(Trang 184).*

Bước 3: Khoảng ước lượng cần tìm

$$\left(\overline{x}-\varepsilon;\overline{x}+\varepsilon\right)$$

Ví dụ 3: Tiến hành điều tra trên 148 hộ gia đình về mức tiêu thụ sản phẩm trung bình ở một địa phương. Hãy ước lượng khoảng tin cậy 99% về mức tiêu thụ sản phẩm trung bình của hộ gia đình.

Mức tiêu thụ Kg/hộ	0-1	1-2	2-3	3-4	4-5	>5
Số hộ	17	25	40	37	18	11

- Gọi *m* là mức tiêu thụ sản phẩm Trung bình của một hộ ở địa phương.

m: chưa biết

- Ta cần ước lượng *m* với độ tin cậy $1-\alpha=99\% \Leftrightarrow \alpha=0.01$

Bước 1: Phân tích mẫu có: $n = 148; \overline{x} = 2.817; s = 1.395$

Bước 2: Sai số ước lượng

Vì $n = 148 > 30$ nên $\varepsilon = Z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \Rightarrow \varepsilon = Z_{0.995} \frac{s}{\sqrt{n}} = 2.576 \frac{1.395}{\sqrt{148}} = 0.295$

Bước 3: Khoảng ước lượng cần tìm:

$$\left(\overline{x} - \varepsilon; \overline{x} + \varepsilon\right) = \left(2.817 - 0.295; 2.817 + 0.295\right) = \left(2.522; 3.112\right)$$

Ví dụ 4: Người ta muốn điều tra xem năng suất lúa tại một vùng sau khi sử dụng giống mới là bao nhiêu. Tiến hành kiểm tra trên diện tích 100ha trồng giống lúa mới tại vùng này và được bảng số liệu sau:

Năng suất (tấn/ha)	21	24	25	26	28	32	34
Diện tích (ha)	8	22	30	15	13	7	5

a) Hãy ước lượng năng suất lúa trung bình của giống lúa mới với độ tin cậy 99%.

b) Nếu muốn độ chính xác của ước lượng trên là 0.51 tấn/ha thì độ tin cậy của ước lượng là bao nhiêu?

c) Biết những thửa ruộng có năng suất từ 28 tấn/ha trở lên là những thửa ruộng có năng suất cao. Hãy ước lượng năng suất lúa trung bình của những thửa ruộng có năng suất cao với mức ý nghĩa 5%.

Giải

Gọi *m* là năng suất lúa trung bình của vùng

m: chưa biết

a) Ta cần ước lượng *m* với độ tin cậy $1 - \alpha = 99\% \Leftrightarrow \alpha = 0.01$

Bước 1: Phân tích mẫu có: $n = 100; \overline{x} = 25.940; s = 3.084$

Bước 2: Sai số ước lượng

Vì $n = 100 > 30$ nên $\varepsilon = Z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \Rightarrow \varepsilon = Z_{0.995} \frac{s}{\sqrt{n}} = 2.576 \frac{3.084}{\sqrt{100}} = 0.794$

Bước 3: Khoảng ước lượng cần tìm:

$$\left(\overline{x} - \varepsilon; \overline{x} + \varepsilon\right) = \left(25.94 - 0.794; 25.94 + 0.794\right) = \left(25.146; 26.734\right)$$

b) Tìm độ tin cậy của ước lượng

Từ công thức sai số ước lượng

$$\varepsilon = Z_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \Rightarrow z_{1-\frac{\alpha}{2}} = \frac{\varepsilon \times \sqrt{n}}{s}$$

Ta có: $\varepsilon = 0.51, n = 100, s = 3.084 \Rightarrow z_{1-\frac{\alpha}{2}} = \frac{\varepsilon \times \sqrt{n}}{s} = \frac{0.51 \times \sqrt{100}}{3.084} = 1.653 \approx z_{0.95}$

Khi đó,

$$1-\frac{\alpha}{2}=0.95 \Rightarrow 1-\alpha=0.90$$

Vậy độ tin cậy của ước lượng là 90%.

c) Ta cần ước lượng m với mức ý nghĩa $\alpha=0.05$

Bước 1: Phân tích mẫu có: $n=25; \overline{x}=30.32; s=2.561$

Bước 2: Sai số ước lượng

Vì $n=25<30$ nên $\varepsilon=t_{\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}}=t_{0.025}(24)\frac{s}{\sqrt{n}}=2.064\frac{2.561}{\sqrt{25}}=1.057$

Bước 3: Khoảng ước lượng cần tìm:

$$\left(\overline{x}-\varepsilon;\overline{x}+\varepsilon\right)=\left(30.32-1.057;30.32+1.057\right)=\left(29.263;31.377\right)$$

Ví dụ 5: Điều tra năng suất lúa trên diện tích 100 ha ở một vùng được bảng số liệu sau:

Năng suất (tấn/ha)	21	24	25	26	28	32	34
Diện tích (ha)	10	20	30	15	10	10	5

a) Ước lượng năng suất lúa trung bình của vùng đó với độ tin cậy 95%.

b) Nếu muốn độ chính xác khi ước lượng trung bình là 0.5 tấn thì độ tin cậy của ước lượng là bao nhiêu?

c) Với độ tin cậy 99%, nếu muốn độ chính xác khi ước lượng trung bình là 0.7 thì cần điều tra bao nhiêu diện tích nữa?

III. Ước lượng phương sai (Tự xem sách)