CHƯƠNG 6: ƯỚC LƯỢNG THAM SỐ CỦA TỔNG THỂ

Free download
2015

1

- Tổng thể được đặc trưng bởi *dấu hiệu nghiên cứu X*, là đại lượng ngẫu nhiên. Tổng thể có ba đặc trưng số quan trọng là:
  - $E(X)=\mu$: trung bình tổng thể (*định lượng*)
  - $var(X)=\sigma^2$: phương sai tổng thể (*định lượng*)
  - p: tỷ lệ tổng thể (*định tính*)
- Ta gọi chung các đặc trưng số của tổng thể là θ.
- Đặc trưng số của *tổng thể* là một giá trị số ***cố định*** nhưng ***chưa biết***, còn đặc trưng số của *mẫu* là giá trị số ***biết*** nhưng ***không cố định***. Ta phải dự đoán (ước lượng) θ.
- Có hai dạng ước lượng cơ bản là *ước lượng điểm* và *ước lượng khoảng*.

2

1) Ước lượng điểm

Từ kết quả khảo sát của mẫu, ta có thể đưa ra *một con số* $\hat{\theta}$ để ước lượng (dự đoán) cho θ.

Khi đó $\hat{\theta}$ được gọi là ước lượng điểm của θ.

Thí dụ: người ta hay dùng:

- trung bình mẫu $\bar{x}$ để ước lượng trung bình tổng thể μ
- phương sai mẫu (đã hiệu chỉnh) s^2 để ước lượng phương sai đám đông σ^2
- tỷ lệ mẫu f để ước lượng tỷ lệ đám đông p

3

2) Ước lượng khoảng

Từ kết quả khảo sát mẫu, ta đưa ra khoảng $(\hat{\theta}_1, \hat{\theta}_2)$, với *mong muốn* là tham số tổng thể θ sẽ thuộc vào khoảng này với một xác suất nhất định $\gamma = 1-\alpha$, nghĩa là:

$$P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = P[\theta \in (\hat{\theta}_1, \hat{\theta}_2)] = 1-\alpha$$

thì $(\hat{\theta}_1, \hat{\theta}_2)$ gọi là khoảng tin cậy, khoảng ước lượng hay ước lượng khoảng của θ.

$(1-\alpha)$ được gọi là độ tin cậy của khoảng ước lượng.

4

Khi đưa ra ước lượng khoảng $(\hat{\theta}_1, \hat{\theta}_2)$ từ mẫu thì có hai trường hợp xảy ra:

- Khoảng ước lượng này thực sự chứa θ, tức là ta ước lượng đúng.
- Khoảng ước lượng này không chứa θ, tức là ta ước lượng sai.

Xác suất ước lượng sai là $\alpha = P[\theta \notin (\hat{\theta}_1, \hat{\theta}_2)]$, gọi là xác suất mắc sai lầm khi ước lượng.

5

Bình loạn:

- Ai lấy vợ cũng đều mong ước / ao ước / kỳ vọng vợ mình đẹp, hiền, nết na, thùy mị, đoan trang, giỏi giang, cẩn thận,... nói chung là hết ý!!!
- Ta "ước lượng" người "ấy" đạt những điều ao ước trên thì ta mới rước nàng về "dinh".
- Sau khi cưới xong, có 2 trường hợp xảy ra:
  - Thực tế người "ấy" có những đức tính trên: Ta ước lượng đúng. *Hoan hô, cuộc đời vẫn đẹp sao…* !!!
  - Thực tế người "ấy" không có các đức tính trên, nhưng *giả bộ có*, làm ta mất phương hướng: Ta ước lượng sai. *Thành thật chia bùn* !!!

6

❖Ta có các dạng ước lượng cơ bản sau:

- Ước lượng giá trị trung bình
- Ước lượng tỷ lệ

- Trong thực hành, để ước lượng *giá trị trung bình* người ta căn cứ vào cỡ mẫu n *(lớn hoặc nhỏ)* và phương sai var(X)= σ^2 *(biết hoặc không)* để đưa ra phương pháp ước lượng tương ứng.
- Còn *ước lượng tỷ lệ* đòi hỏi mẫu lớn (n>=30).

7

A. ƯỚC LƯỢNG GIÁ TRỊ TRUNG BÌNH μ

1a) <u>n ≥ 30 , biết σ^2</u>

$\mu = \bar{x} \pm t_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ hay $\bar{x} - t_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

tra bảng F, với $\gamma = 1-\alpha = 2.\phi(t_{\alpha/2})$

1b) <u>n ≥ 30 , *nếu không biết σ :*</u> thay σ bằng s

$\mu = \bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$

2a) <u>n < 30, biết σ^2 *(X có phân phối chuẩn)*</u>

$\mu = \bar{x} \pm t_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

2b) **<u>n < 30, không biết σ^2</u>** <u>*(X có phân phối chuẩn)*</u>

$\mu = \bar{x} \pm t_{\alpha/2}^{(n-1)}\frac{s}{\sqrt{n}}$

tra bảng H, bậc tự do n–1

8

Độ chính xác (sai số) của ước lượng:

$\varepsilon = t_{\alpha/2}\dfrac{\sigma(s)}{\sqrt{n}}$ hoặc $\varepsilon = t_{\alpha/2}^{(n-1)}\dfrac{s}{\sqrt{n}}$

Ta có:

$P(\bar{X} - \varepsilon < \mu < \bar{X} + \varepsilon) = 1 - \alpha = \gamma$

$\Leftrightarrow P(|\ \bar{X}\ -\ \mu\ | <\ \varepsilon) = 1 - \alpha$

tính toán thực tế sai số

Ưu điểm của UL Khoảng so với UL Điểm?

Bảng F ↔ Phụ lục 2 sách ôn Cao học

Biết độ tin cậy $\gamma = 1-\alpha$, tìm $t_{\alpha/2}$ =?

- ***Với độ tin cậy $\gamma = 0,95$*** → $\gamma/2 = 0,475$

Số 0,475 ở dòng **1.9** và cột **6**. Vậy $t_{\alpha/2} = 1,96$

- ***Với độ tin cậy $\gamma = 0,94$*** → $\gamma/2 = 0,47$

Không thấy số 0,47 trong bảng F.

Số 0,4699 sai lệch so với 0,47 là ***nhỏ nhất***.

Vậy $t_{\alpha/2} = 1,88$

- ***Với độ tin cậy $\gamma = 0,90$*** → $\gamma/2 = 0,45$

Ta thấy có số 0,4495 → $t_{\alpha/2} = 1,64$

Ta thấy có số 0,4505 → $t_{\alpha/2} = 1,65$

10 Vậy $t_{\alpha/2} = 1,65$ hoặc $t_{\alpha/2} = 1,64$

Bảng H

1) ***Biết độ tin cậy $\gamma = 1-\alpha$, tìm $t_{\alpha/2}(n-1)$=?***

- **$\gamma = 0,95$, n= 20** $\Rightarrow$ $t_{\alpha/2}(n-1) = t_{0,025}(19) = 2,0930$

Dòng n-1= 19 và cột γ= 0.95 ta có giá trị 2.0930

- **$\gamma = 0,99$, n = 5** $\Rightarrow$ $t_{\alpha/2}(n-1) = t_{0,005}(4) = 4,6041$

2) ***Biết $t_{\alpha/2}(n-1)$, tìm độ tin cậy $\gamma = 1-\alpha$ =?***

- *Với n= 20 và $t_{\alpha/2}(n-1) = 2,3457$*

Dòng n-1 = 19, số 2.3457 ở cột γ =0.97 nên γ= 0,97

- *Với n= 19 và $t_{\alpha/2}(n-1) = 2,0$*

Dòng n-1 = 18, số 2.0 ≈ 2.0071 nên $\gamma \approx 0,94$

11

Bảng phụ lục 4 (Sách ôn Cao học) ↔ Bảng H

1) ***Biết độ tin cậy $\gamma = 1-\alpha$, tìm $t_{\alpha/2}(n-1)$=?***

- **$\gamma = 0,95$** $\Rightarrow$ $\alpha = 0,05 \Rightarrow \alpha/2 = 0,025$, **n= 20**

  $\Rightarrow$ $t_{\alpha/2}(n-1) = t_{0,025}(19) = 2,093$

Dòng k= 19 và ***cột α= 0,025*** ta có giá trị 2.0930

2) ***Biết $t_{\alpha/2}(n-1)$, tìm độ tin cậy $\gamma = 1-\alpha$ =?***

- *Với n= 20 và $t_{\alpha/2}(n-1) = 2,3457 \approx 2,346$*

Dòng k= 19, số 2.346 ở ***cột α= 0.015***

nên ***$\alpha/2 = 0,015$*** $\Rightarrow \alpha = 0,03 \Rightarrow \gamma \approx 0,97$

- *Với n= 19 và $t_{\alpha/2}(n-1) = 2,0$*

Dòng k= 18, số 2.0 ≈ 2.007 ở ***cột α= 0.03***

12 nên ***$\alpha/2 = 0,03$*** $\Rightarrow \alpha = 0,06 \Rightarrow \gamma \approx 0,94$

VD: Giả sử ta có n= 64, $\bar{x}$= 28, s= 6

Áp dụng công thức $\varepsilon = t_{\alpha/2} s/\sqrt{n}$

và khoảng tin cậy $(\bar{x}-\varepsilon, \bar{x}+\varepsilon)$ ta có bảng sau:

Độ tin cậy γ	$t_{\alpha/2}$	Độ chính xác (sai số) ε	Khoảng tin cậy	
99%	2,58	1.9350	26.0650	29.9350
95%	1,96	1.4700	26.5300	29.4700
90%	1,65	1.2375	26.7625	29.2375

26,065 26,53 26,7625 **28** 29,2375 29,47 29,935

Vậy : ĐTC cao $\Leftrightarrow$ giá trị ε lớn $\Leftrightarrow$ KTC rộng $\Leftrightarrow$ ĐCX kém

Nếu dự báo thời tiết (nhiệt độ) thì ta thích KTC rộng hay hẹp ?!

13

Phân biệt ước lượng điểm và UL khoảng?

VD1:

Điểm trung bình môn toán của 100 thí sinh dự thi vào ĐHKT là 5, với độ lệch chuẩn mẫu (đã hiệu chỉnh) s = 2,5.

1) Ước lượng điểm trung bình môn toán của toàn thể thí sinh của trường

2) Ước lượng điểm trung bình môn toán của toàn thể thí sinh với độ tin cậy là 95%.

3) Với sai số 0,25 điểm. Hãy xác định độ tin cậy.

14

Giải

1) Do $\bar{x}$= 5 nên đtb môn toán của toàn thể thí sinh là 5

2) Áp dụng trường hợp n ≥ 30 , σ chưa biết :

$\gamma = 95\% \Rightarrow t_{\alpha/2} = 1{,}96$

$$\mu = \bar{x} \pm \frac{t_{\alpha/2} s}{\sqrt{n}} = 5 \pm \frac{1{,}96*2{,}5}{\sqrt{100}} = 5 \pm 0{,}49$$

Vậy với độ tin cậy 95% KUL điểm trung bình môn toán của toàn thể thí sinh dự thi là (4,51 ; 5,51) điểm.

3) $\varepsilon = 0{,}25 \Rightarrow t_{\alpha/2} = \frac{\varepsilon\sqrt{n}}{s} = 0{,}25*10/2{,}5 = 1$

$\Rightarrow \phi(t_{\alpha/2}) = \phi(1{,}00) = 0{,}3413$ *(tra bảng F)*

$\gamma = 2\phi(t_{\alpha/2}) \Rightarrow \gamma = 0{,}6826 = 68{,}26\%$

15

VD2:

Tuổi thọ của một loại bóng đèn được biết theo quy luật chuẩn, *với độ lệch chuẩn 100 giờ*.

1) Chọn ngẫu nhiên 100 bóng để thử nghiệm, thấy mỗi bóng tuổi thọ trung bình là 1000 giờ. Hãy ước lượng tuổi thọ trung bình của bóng đèn xí nghiệp A sản xuất với độ tin cậy 95%

2) Với độ chính xác là 15 giờ. Hãy xác định độ tin cậy.

3) Với độ chính xác là 25 giờ và độ tin cậy là 95% thì cần thử nghiệm bao nhiêu bóng.

16

Giải: Áp dụng trường hợp n ≥ 30, σ đã biết

1) n = 100 ; $\bar{x}$ = 1000 ; γ = 95% ; σ = 100

$\gamma = 95\% \Rightarrow t_{\alpha/2} = 1{,}96$

$$\mu = \bar{x} \pm t_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 1000 \pm \frac{1{,}96*100}{\sqrt{100}} = 1000 \pm 19{,}6$$

Vậy với độ tin cậy 95% tuổi thọ trung bình của bóng đèn thuộc xí nghiệp A vào khoảng (980,4 ; 1019,6) giờ

2) $t_{\alpha/2} = \frac{\varepsilon\sqrt{n}}{\sigma} = \frac{15*\sqrt{100}}{100} = 1{,}5 \Rightarrow \phi(1{,}50) = 0{,}4332$ *(bảng F)*

$\gamma = 2\phi(t_{\alpha/2}) \Rightarrow \gamma = 0{,}8662 = 86{,}62\%$

3) $\gamma = 95\% \Rightarrow t_{\alpha/2} = 1{,}96$

$$n = \frac{t^2_{\alpha/2}\sigma^2}{\varepsilon^2} = \frac{(1{,}96)^2 \cdot 100^2}{25^2} = 61{,}466 \approx 62 \text{ (làm tròn lên)}$$

17

Làm tròn lên của 1 số thập phân là lấy phần nguyên của số đó cộng thêm 1.

Nhận xét: Các dạng toán UL cơ bản

Dạng toán:

Có 3 tham số : n, ε , γ =1−α (biết γ ↔ biết $t_{\alpha/2}$)

Các tham số mẫu: $\bar{x}$, s

1) Biết n, γ → ε = ?

2) Biết n, ε → γ = ?

3) Biết ε , γ → n = ?

Dùng công thức $\varepsilon = t_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ hay $\varepsilon = t_{\alpha/2}\frac{s}{\sqrt{n}}$

18

- **VD2bis:**
- Để khảo sát hàm lượng chất đạm X (%) trong một loại sữa hộp, người ta kiểm tra 100 hộp và thấy hàm lượng đạm trung bình là 18 (%) và độ lệch chuẩn mẫu (có hiệu chỉnh) là 4(%).
- Nếu muốn ước lượng trung bình của hàm lượng đạm trong 1 hộp sữa đạt độ tin cậy 95% và độ chính xác 0,5 (%) thì cần phải khảo sát *thêm* bao nhiêu hộp sữa *nữa*?
- **Giải:**
- $n = (t_{\alpha/2}.s/\varepsilon)^2 = (1{,}96*4 / \mathbf{0{,}5})^2 = 245{,}86 \approx 246$

Vậy cần phải khảo sát thêm 246-100 = 146 hộp

19

Tra bảng H, tại sao?

VD3 :

Trọng lượng các bao bột mì tại một cửa hàng lương thực theo quy luật chuẩn. Kiểm tra 20 bao, thấy trọng lượng trung bình của mỗi bao bột mì là 48kg, và phương sai mẫu hiệu chỉnh là $s^2 = (0{,}5\text{kg})^2$.

1) Với độ tin cậy 95% hãy ước lượng trọng lượng trung bình của một bao bột mì thuộc cửa hàng.

2) Với độ chính xác 260 g, xác định độ tin cậy.

20

Giải:

1) Áp dụng trường hợp n < 30 , chưa biết σ

$\gamma = 95\% \Rightarrow t_{\alpha/2}(n-1) = 2{,}0930$ *(tra bảng H)*

$$\mu = \bar{x} \pm \frac{t_{0,025}(19).s}{\sqrt{n}} = 48 \pm \frac{2{,}093*0{,}5}{\sqrt{20}} = 48 \pm 0{,}234$$

Vậy với độ tin cậy 95%, trọng lượng trung bình của một bao bột mì thuộc cửa hàng thuộc vào khoảng (47,766 ; 48,234) kg

2) ε = 260 g = 0,26 kg

$t_{\alpha/2}(n-1) = t_{\alpha/2}(19) = \frac{(0{,}26)\sqrt{20}}{0{,}5} = 2{,}325 \approx 2{,}3457$

(2,3457 là giá trị gần 2,325 nhất trong bảng tra, **cùng dòng 19**).

$\Rightarrow \gamma = 0{,}97 = 97\%$ *(tra bảng H)*

21

B. ƯỚC LƯỢNG TỶ LỆ p : *với n ≥ 30*

$$p = f \pm t_{\alpha/2}\sqrt{\frac{f(1-f)}{n}}$$

$\varepsilon = t_{\alpha/2}\sqrt{\frac{f(1-f)}{n}}$ độ chính xác (sai số) của UL

Điều kiện áp dụng : $\begin{cases} n.f > 10 \\ n.(1-f) > 10 \end{cases}$

Dạng toán:

Có 3 dạng toán giống ước lượng trung bình

Tham số mẫu: f

Dùng công thức $\varepsilon = t_{\alpha/2}\sqrt{\frac{f(1-f)}{n}}$

22

Phân biệt ước lượng điểm và UL khoảng?

VD4:

Để ước lượng tỷ lệ sản phẩm xấu của một kho đồ hộp, người ta kiểm tra ngẫu nhiên 100 hộp thấy có 11 hộp xấu.

1) Ước lượng tỷ lệ sản phẩm xấu của kho đồ hộp.

2) Ước lượng tỷ lệ sản phẩm xấu của kho đồ hộp với độ tin cậy 94%.

3) Với sai số cho phép ε = 3%, xác định độ tin cậy.

23

Giải

1) n = 100 , tỷ lệ mẫu $f = \frac{11}{100} = 0{,}11$

Vậy tỷ lệ hộp xấu của kho là 11%

2) $\gamma = 94\% = 0{,}94 \Rightarrow t_{\alpha/2} = 1{,}88$ *(tra bảng F)*

$$p = f \pm \frac{t_{\alpha/2}\sqrt{f(1-f)}}{\sqrt{n}} = 0{,}11 \pm \frac{1{,}88\sqrt{0{,}11(1-0{,}11)}}{\sqrt{100}} = 0{,}11 \pm 0{,}059$$

Vậy với độ tin cậy 94%, tỷ lệ sản phẩm xấu của kho đồ hộp vào khoảng (0,051 ; 0,169) ⇒ 5,1% < p < 16,9%

3) ε = 3% = **0,03**

$$t_{\alpha/2} = \frac{\varepsilon\sqrt{n}}{\sqrt{f(1-f)}} = \frac{0{,}03\sqrt{100}}{\sqrt{0{,}11(1-0{,}11)}} = 0{,}96$$

$\phi(0{,}96) = 0{,}3315 \Rightarrow \gamma = 2\times\phi(0{,}96) = 0{,}663 = 66{,}3\%$

24

VD5: Lô trái cây của một chủ hàng được đóng thành sọt, mỗi sọt 100 trái. Kiểm tra 50 sọt thấy có 450 trái không đạt tiêu chuẩn.

1) Ước lượng tỷ lệ trái cây không đạt tiêu chuẩn của lô hàng với độ tin cậy 95%.
2) Muốn ước lượng tỷ lệ trái cây không đạt tiêu chuẩn với độ chính xác 0,5% thì độ tin cậy đạt được là bao nhiêu?
3) Muốn ước lượng tỷ lệ trái cây không đạt tiêu chuẩn với độ tin cậy 99% và độ chính xác 1% thì cần kiểm tra bao nhiêu sọt?
4) Muốn ước lượng tỷ lệ trái cây không đạt TC với độ tin cậy 99,70% thì độ chính xác đạt được là bao nhiêu?

25

Giải:

1) Gọi p là tỷ lệ trái cây không đạt tiêu chuẩn.

$\gamma = 95\% \Rightarrow t_{\alpha/2} = 1{,}96$

Tỷ lệ mẫu $f = \dfrac{450}{5000} = 0{,}09$

$$\varepsilon = 1{,}96\sqrt{\frac{0{,}09\times(1-0{,}09)}{5000}} = 0{,}008$$

Khoảng ước lượng của p là: $0{,}082 < p < 0{,}098$

2) $$t_{\alpha/2} = \frac{\varepsilon\sqrt{n}}{\sqrt{f(1-f)}} = 0{,}005\sqrt{\frac{5000}{0{,}09(1-0{,}09)}} = 1{,}24$$

$\gamma = 2\phi(t_{\alpha/2}) = 2 \times 0{,}3925 = 0{,}785$. *(tra bảng F)*

Vậy độ tin cậy đạt được 78,5%.

26

3) Ta cần xác định kích thước mẫu n.

$\gamma = 99\% \Rightarrow t_{\alpha/2} = 2{,}58$ *(tra bảng F)*

$$n = t_{\alpha/2}^2 \frac{f(1-f)}{\varepsilon^2} = 2{,}58^2 \frac{0{,}09\times(1-0{,}09)}{(0{,}01)^2} = 5451{,}59 \text{ (trái)}$$

Vì mỗi sọt có 100 trái nên ta cần kiểm tra 5451,59/100 = 54,5159 ≈ 55 sọt.

4) Ta cần xác định độ chính xác ε với độ tin cậy 99,70% (ứng $t_{\alpha/2}$= 2,96) với kích thước mẫu n = 5000.

$$\varepsilon = t_{\alpha/2}\frac{\sqrt{f(1-f)}}{\sqrt{n}} = 2{,}96\times\frac{\sqrt{0{,}09\times(1-0{,}09)}}{\sqrt{5000}} = 0{,}012$$

Vậy độ chính xác đạt được 1,2%.

27

Câu hỏi:

- Qua 2 thí dụ trên bạn rút ra được các điều cần lưu ý chưa?
- "Chuyện nhỏ nhưng nếu không biết lại là chuyện lớn" (nhạc Rap VN)!

28

<u>VD6</u>: Một lô hàng có 5000 sản phẩm. Chọn ngẫu nhiên 400 sản phẩm từ lô hàng để kiểm tra thì thấy có 360 sản phẩm loại A.

1) Hãy ước lượng <u>tỷ lệ</u> sản phẩm loại A có trong lô hàng, với độ tin cậy 96%?
2) Hãy ước lượng <u>số</u> sản phẩm loại A có trong lô hàng, với độ tin cậy 96%?
3) Nếu muốn ước lượng <u>số</u> sản phẩm loại A của lô hàng đạt được độ chính xác 150 sản phẩm và độ tin cậy 99% thì phải kiểm tra bao nhiêu sản phẩm?

29

<u>Giải</u>:

1) Tỷ lệ mẫu f= 360 / 400 = 0,9

$$p = f \pm t_{\alpha/2}\sqrt{\frac{f(1-f)}{n}} = 0,9 \pm 2,05\frac{\sqrt{0,9.0,1}}{\sqrt{400}} = 0,9 \pm 0,0308$$

$\to 0,8692 < p < 0,9308$

2) Gọi M là số sản phẩm loại A có trong lô hàng:

$0,8692 * 5000 < M < 0,9308 * 5000$

3) Với ε = 150 / 5000 = 0,03

$$\varepsilon = t_{\alpha/2}\sqrt{\frac{f(1-f)}{n}} \to n = \left(\frac{t_{\alpha/2}}{\varepsilon}\right)^2 f(1-f)$$

$$n = 2,58^2\frac{0,9.0,1}{0,03^2} = 665,640 \cong 666 \text{ sản phẩm}$$

30

<u>Chứng minh</u>: gọi ε là độ chính xác của ước lượng khoảng ứng với 400 sản phẩm, và ε' là độ chính xác của ước lượng khoảng ứng với 5000 sản phẩm.

Ta có $p = f \pm \varepsilon$ ứng với ước lượng tỷ lệ của 400 sản phẩm. $Np = Nf \pm N\varepsilon$ là ước lượng ứng với N= 5000 sản phẩm, và độ chính xác là ε'= Nε= 150.

Vậy ε = ε'/N= 150/5000 = 0,03

31

- <u>VD6bis</u>:
- Một lô hàng có rất nhiều sản phẩm, trong đó có 4500 sản phẩm loại A. Lấy ngẫu nhiên 400 sản phẩm từ lô hàng thì thấy có 360 sản phẩm loại A.
- 1) Ước lượng số sản phẩm có trong lô hàng?
- 2) Ước lượng số sản phẩm có trong lô hàng, với độ tin cậy 96%?

32

- Giải:
- 1) Tỷ lệ mẫu: f= 360/400 = 0,9
- Gọi p= M/N là tỷ lệ sản phẩm loại A của lô hàng
- ước lượng điểm: p=f → 4500/N = 0,9 → N= 5000
- 2) Theo kết quả bài 6, ta có ước lượng khoảng:
- 0,8692 < p= 4500/N < 0,9308
- → 4835 ≈ 4834,55 < N < 5177,17 ≈ 5178
- Lưu ý:
  - p= M/N , p luôn luôn ước lượng được
  - Biết N tìm M: VD6
  - Biết M tìm N: VD6bis

33

Câu hỏi:

- Bạn đã rút ra được điều cần lưu ý từ 2 thí dụ này chưa?
- Các dạng toán tương tự làm giống như 2 thí dụ này.
- Hãy để chuyện nhỏ mãi mãi là chuyện nhỏ!

34

VD7: Điều tra năng suất lúa trên diện tích 100 hecta trồng lúa của một vùng, được bảng số liệu sau:

Năng suất (tạ/ha)	41	44	45	46	48	52	54
Số ha có năng suất tương ứng	10	20	30	15	10	10	5

1) Hãy ước lượng năng suất lúa trung bình của vùng đó, với độ tin cậy 95%?

2) Những thửa ruộng có năng suất từ 48tạ/ha trở lên là những thửa có năng suất cao.

Hãy ước lượng tỷ lệ diện tích có năng suất cao trong vùng, với độ tin cậy 97%.

35

Giải:

1) Ta lập bảng như sau

x_i	n_i	$n_i x_i$	$n_i x_i^2$
41	10	410	16.810
44	20	880	38.720
45	30	1350	60.750
46	15	690	31.740
48	10	480	23.040
52	10	520	27.040
54	5	270	14.580
Tổng	**n = 100**	**4600**	**212680**

36

1) Từ kết quả tính ở bảng trên ta có

Năng suất trung bình $\overline{x} = \frac{4600}{100} = 46$ tạ/ha

Phương sai của năng suất

$$s^2 = \frac{1}{100-1}\left(212680 - 100 * 46^2\right) = 10,910$$

$\Rightarrow$ s= 3,303

γ = 95% $\Rightarrow$ $t_{\alpha/2}$ = 1,96

$$\mu = \overline{x} \pm \frac{t_{\alpha/2}\, s}{\sqrt{n}} = 46 \pm \frac{1,96*3,303}{\sqrt{100}} = 46 \pm 0,647$$

Vậy năng suất lúa trung bình của vùng đó vào khoảng (45,353 ; 46,647) đơn vị tính tạ.

37

2) Tỷ lệ mẫu $f = \frac{25}{100} = 0,25$

γ = 0,97 $\Rightarrow$ $t_{\alpha/2}$ = 2,17 *(tra bảng F)*

$$p = 0,25 \pm \frac{2,17\sqrt{0,25*0,75}}{\sqrt{100}}$$

$$= 0,25 \pm 0,094$$

Vậy với độ tin cậy 97%, tỷ lệ diện tích lúa có năng suất cao trong vùng vào khoảng (0,156 ; 0, 344).

38

- **<u>VD7bis</u>:**
- Với giả thiết của VD 7, câu 2.
- Hãy ước lượng *diện tích lúa* có năng suất cao của vùng này, biết rằng vùng này có diện tích 10.000 ha? Với độ tin cậy 97%.
- **<u>Giải</u>:**
- Gọi M là diện tích lúa có năng suất cao của vùng này.
- Ta có 0,156 < p < 0,344
- $\rightarrow$ 0,156 * 10.000 < M < 0,344 * 10.000

39

<u>VD8</u>

Một công ty tiến hành khảo sát nhu cầu tiêu dùng về 1 loại sản phẩm do công ty sản xuất. Khảo sát trên 500 hộ gia đình ở 1 thành phố ta được bảng số liệu:

Số lượng (kg/tháng)	0	2–3	3–4	4–5	5–6	6–7	7–8
Số hộ	150	33	52	127	73	35	30

1) Hãy ước lượng số lượng sản phẩm của công ty được tiêu thụ tại thành phố trung bình trong 1 tháng, với độ tin cậy 94%. Cho biết tổng số hộ gia đình trong toàn thành phố là 500000 hộ.
2) Hãy ước lượng mức tiêu thụ trung bình trên mỗi hộ ở <u>các hộ có nhu cầu sử dụng</u>, với độ tin cậy 95%.
3) Ước lượng số lượng sản phẩm công ty tiêu thụ được ở thành phố trung bình trong 1 tháng? Biết tổng số hộ <u>có tiêu dùng sản phẩm</u> là 400000 hộ?

40

Hướng dẫn

1) n= 500 , $\sum nx = 1690$, $\sum nx^2 = 8789,5$, $\overline{x} = 3,38$, s = 2,483

Gọi a là nhu cầu trung bình của 1 hộ về loại sản phẩm này

Gọi M là nhu cầu tb của toàn thành phố về loại sp này

$\varepsilon = t_{\alpha/2}\dfrac{s}{\sqrt{n}} = 1,88\dfrac{2,483}{\sqrt{500}} = 0,209 \Rightarrow 3,171 < a < 3,589$

$\Rightarrow$ 500.000 × 3,171 < M < 3,589 × 500.000 (kg/tháng)

2) n= 350, $\sum nx = 1690$,$\sum nx^2 = 8789,5$, $\overline{x} = 4,829$, s= 1,341

$\varepsilon = t_{\alpha/2}\dfrac{s}{\sqrt{n}} = 1,96\dfrac{1,341}{\sqrt{350}} = 0,14 \Rightarrow 4,689 < a < 4,969$

3) Số lượng sản phẩm công ty tiêu thụ được trung bình ở thành phố là 400.000 * *4,829* = 1931600 kg/tháng

41

VD9: Đo đường kính của 100 chi tiết do một máy sản xuất kết quả cho ở bảng sau:

Đường kính (mm)	Số chi tiết
19,5 – 20,0	3
20,0 – 20,5	5
20,5 – 21,0	*16*
21,0 – 21,5	*28*
21,5 – 22,0	*23*
22,0 – 22,5	*14*
22,5 – 23,0	7
23,0 – 23,5	4

Quy định những chi tiết có đường kính từ 20,5 mm đến 22,5 mm là những chi tiết đạt tiêu chuẩn.

Khi *ước lượng trung bình* đường kính của chi tiết đạt tiêu chuẩn với độ chính xác đạt 0,08 mm và khi *ước lượng tỷ lệ* chi tiết đạt tiêu chuẩn với độ chính xác là 5%, với cùng độ tin cậy 99% thì cần *đo thêm bao nhiêu* chi tiết *nữa*.

42

Giải

x_i	n_i	n_ix_i	$n_ix_i^2$
20,75	16	332,00	6889,0000
21,25	28	595,00	12643,7500
21,75	23	500,25	10880,4375
22,25	14	311,50	6930,8750
Tổng	**n= 81**	**1738,750**	**37344,0625**

1) $\overline{x} = \dfrac{1738,75}{81} = 21,466$

$s^2 = \dfrac{1}{81-1}\left(37344,0625 - 81\times(21,466)^2\right) = 0,252 \rightarrow s = 0,502$

2) Tỷ lệ chi tiết đạt tiêu chuẩn f = 81/100 = 0,81

43

1) Để ước lượng đường kính trung bình của chi tiết đạt tiêu chuẩn với độ chính xác 0,08 mm và độ tin cậy 99% thì cần mẫu có kích thước n_1

$$n_1 = \left(\frac{t_{\alpha/2}s}{\varepsilon}\right)^2 = \left(2,58\times\frac{0,502}{0,08}\right)^2 \approx 263$$

2) Để ước lượng tỷ lệ chi tiết đạt tiêu chuẩn với độ chính xác là 5% và độ tin cậy 99% thì cần mẫu có kích thước n_2

$$n_2 = t_{\alpha/2}^2\frac{f(1-f)}{\varepsilon^2} = 2,58^2\frac{0,81\times 0,19}{(0,05)^2} \approx 410$$

3) Để thỏa mãn đồng thời các điều kiện của bài toán thì cần mẫu có kích thước: n = max{n_1, n_2} = 410

Vậy ta cần đo thêm 410 – 100 = 310 chi tiết nữa

44

VD10

Một khách sạn lớn muốn ước lượng tỷ lệ khách có nhu cầu nghỉ trọ nhiều hơn 1 ngày. Họ muốn có độ tin cậy 96% và sai số không quá 5%. Hỏi cần lấy mẫu với kích thước thích hợp là bao nhiêu:

1) Nếu chưa có bất kỳ thông tin nào cho phép ước lượng này.

2) Nếu dựa vào một tài liệu khảo sát trước đây, thông tin cho biết tỷ lệ này là 25%.

45

Giải:

1) Ta có $f.(1-f) \le \left(\frac{f+(1-f)}{2}\right)^2 = \frac{1}{4}$ (bđt Côsi)

Do đó $\varepsilon = t_{\alpha/2}\sqrt{\frac{f(1-f)}{n}} \le t_{\alpha/2}\sqrt{\frac{1}{n}}.\sqrt{\frac{1}{4}} \le 0,05$

$\rightarrow n \ge \left(\frac{t_{\alpha/2}}{0,05}\right)^2.\frac{1}{4} = \left(\frac{2,05}{0,05}\right)^2.\frac{1}{4} = 420,25 \approx 421$

2) $n = \left(\frac{t_{\alpha/2}}{\varepsilon}\right)^2 f(1-f) = \left(\frac{2,050}{0,05}\right)^2 \times 0,25 \times 0,75 \approx 316$

Nhận xét: Khi chưa có thông tin gì hết thì ta phải điều tra với cỡ mẫu nhiều hơn khi có *thông tin f*.

46

Mời ghé thăm trang web:

❖https://sites.google.com/a/ueh.edu.vn/phamtricao/

❖https://sites.google.com/site/phamtricao/

47