BUỔI 7: LÝ THUYẾT ƯỚC LƯỢNG

A. TÓM TẮT LÍ THUYẾT

1. Ước lượng điểm

Cho $X = (x_1, x_2,..., x_n)$ là một mẫu kích thước n từ tổng thể có kì vọng μ và phương sai σ^2.

Số $\hat{\theta} = \hat{\theta}(x_1, x_2,..., x_n)$ được gọi là một ước lượng điểm của đặc trưng θ của tổng thể nếu ta coi nó là một giá trị gần đúng của θ.

a) Ước lượng không chệch

Ước lượng $\hat{\theta}$ của θ được gọi là không chệch nếu $E(\hat{\theta}) = \theta$.

Với một mẫu bất kì, ta có

(1) f là ước lượng không chệch của p ;

(2) $\overline{X}$ là ước lượng không chệch của μ ;

(3) S^2 là ước lượng không chệch của σ^2.

b) Ước lượng hợp lí cực đại

Giả sử tổng thể đã biết phân phối nhưng chưa biết các tham số θ của nó. Khi đó, hàm mật độ xác suất của nó có dạng $f(x, \theta)$. Từ mẫu ta có hàm

$$L(\theta) = L(x_1, x_2,..., x_n, \theta) = f(x_1, \theta).f(x_2, \theta) ... f(x_n, \theta)$$

được gọi là hàm hợp lí của θ.

Số $\hat{\theta} = \hat{\theta}(x_1, x_2,..., x_n)$ được gọi là ước lượng hợp lí cực đại của θ, nếu ứng với giá trị này của θ hàm hợp lí đạt cực đại.

. Nếu tổng thể có phân phối Poisson và mẫu $X = (x_1, x_2,..., x_n)$ thì ước lượng hợp lí cực đại của tham số λ của tổng thể là $\overline{X}$.

. Nếu tổng thể có phân phối chuẩn và mẫu $X = (x_1, x_2,..., x_n)$ thì ước lượng hợp lí cực đại của kì vọng của tổng thể là $\overline{X}$, của phương sai của tổng thể là $\hat{S}^2$.

2. Ước lượng khoảng

a) Khái niệm

Giả sử $\hat{\theta}$ là một ước lượng điểm không chệch của θ.

Khoảng $(\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon)$ được gọi là khoảng ước lượng của θ với độ tin cậy γ, nếu

$P(\hat{\theta} - \varepsilon < \theta < \hat{\theta} + \varepsilon) = P(|\hat{\theta} - \theta| < \varepsilon) \geq \gamma$.

Số $\varepsilon > 0$ được gọi là độ chính xác của ước lượng.

Thông thường ta cần tìm khoảng ước lượng với $\gamma \geq 95\%$ cho trước.

b) Ước lượng khoảng của tỉ lệ

. Bài toán. Giả sử tổng thể gồm các phần tử có tính chất τ và không có tính chất τ. Cần ước lượng tỉ lệ p các phần tử có tính chất τ của tổng thể với độ tin cậy γ cho trước.

. Phương pháp

- Với mẫu định tính kích thước $n \geq 30$, ta tìm được tỉ lệ f các phần tử có tính chất τ của mẫu.

- Tra bảng hàm số Laplace tìm số z_γ sao cho $\Phi(Z_\gamma) = \frac{\gamma}{2}$.

- Tìm độ chính xác của ước lượng

$$\varepsilon = Z_\gamma \sqrt{\frac{f(1-f)}{n}}.$$

- Khi đó $p = f \pm \varepsilon$.

c) Ước lượng khoảng của kì vọng

. **Bài toán** : Tìm khoảng ước lượng của kì vọng μ của tổng thể với độ tin cậy γ cho trước.

.Phương pháp

** Trường hợp $n \geq 30$.*

- Với mẫu định lượng kích thước n ta tìm được $\overline{X}$, S.
- Tra bảng hàm số Laplace tìm số Z_γ sao cho $\Phi(Z_\gamma) = \frac{\gamma}{2}$.
- Tìm độ chính xác

$$\varepsilon = Z_\gamma \frac{s}{\sqrt{n}}.$$

- Khi đó $\mu = \overline{X} \pm \varepsilon$

** Trường hợp n tùy ý, tổng thể có phân phối chuẩn, đã biết phương sai σ^2.*

Khi đó $\varepsilon = Z_\gamma \frac{\sigma}{\sqrt{n}}$; $\mu = \overline{X} \pm \varepsilon$.

** Trường hợp $n < 30$, tổng thể có phân phối chuẩn chưa biết phương sai.*

- Tra bảng phân phối Student dòng n – 1, cột 1-γ, ta tìm được số T_γ.

Khi đó $\varepsilon = T_\gamma \frac{s}{\sqrt{n}}$ và $\mu = \overline{X} \pm \varepsilon$.

c) Ước lượng khoảng của phương sai

.Bài toán. Tìm khoảng ước lượng của phương sai σ^2 của tổng thể với độ tin cậy γ cho trước, biết tổng thể có phân phối chuẩn.

.Phương pháp

- Với mẫu định lượng kích thước n ta tìm được S^2.
- Tra bảng phân phối "khi bình phương" dòng n-1, các cột 1-γ, γ ta được hai số tương ứng là χ_1, χ_2.

- Khi đó, khoảng ước lượng của σ^2 với độ tin cậy γ là

$$\left(\frac{(n-1)S^2}{\chi_1}, \frac{(n-1)S^2}{\chi_2} \right).$$

B. CÁC BÀI GIẢI MẪU

1. Ước lượng tỉ lệ

Bài 1. *Trước ngày bầu cử tổng thống, người ta phỏng vấn ngẫu nhiên 1800 cử tri thì thấy có 1180 người ủng hộ ứng cử viên A. Với độ tin cậy 95%, hỏi ứng cử viên đó thu được tối thiểu bao nhiêu phần trăm số phiếu bầu ?*

Giải

Tổng thể là toàn bộ số cử tri cả nước, các phần tử của tổng thể gồm hai loại : ủng hộ và không ủng hộ ứng cử viên A. Ta ước lượng tỉ lệ p các phần tử ủng hộ ứng cử viên A của tổng thể với độ tin cậy 95%.

Ta có mẫu gồm 1800 phần tử, trong đó có 1180 phần tử ủng hộ ứng cử viên A nên tỉ lệ mẫu là

$$f = \frac{1180}{1800} = 0{,}6556.$$

Tra bảng hàm số Laplace ta thấy

$$\Phi(1{,}96) = \frac{0{,}95}{2} = 0{,}475 \quad \text{nên} \quad Z_\gamma = 1{,}96.$$

Độ chính xác của ước lượng là

$$\varepsilon = 1{,}96\sqrt{\frac{0{,}6556(1-0{,}6556)}{1800}} = 0{,}0220.$$

Do đó tỉ lệ tổng thể ủng hộ ứng cử viên A là

$$P = 0{,}6556 \pm 0{,}0220,$$

hay $0{,}6336 \le p \le 0{,}6776$.

Vậy, tối thiểu ứng cử viên A sẽ thu được 63,36% số phiếu bầu.

Bài 2. *Người ta bắt được 1500 con thú, đánh dấu rồi thả lại vào rừng. Sau một thời gian bắt lại 360 con thì thấy có 27 con bị đánh dấu. Hãy ước lượng số thú trong rừng với độ tin cậy 99%.*

Giải

Tổng thể là toàn bộ thú trong rừng. Số phần tử của tổng thể là N chưa biết nhưng được chia thành hai loại : bị đánh dấu và không bị đánh dấu. Số phần tử bị đánh dấu của tổng thể là M = 1500 đã biết. Do đó để tìm N, ta ước lượng tỉ lệ p các phần tử bị đánh dấu của tổng thể với độ tin cậy 99%.

Ta có số phần tử của mẫu là n = 360, trong đó số phần tử bị đánh dấu là k = 27.

Suy ra tỉ lệ mẫu

$$f = \frac{27}{360} = 0{,}075.$$

Tra bảng hàm số Laplace ta thấy

$$\Phi(2{,}58) = \frac{0{,}99}{2} = 0{,}495 \quad \text{nên} \quad Z_\gamma = 2{,}58.$$

Độ chính xác của ước lượng

$$\varepsilon = 2{,}58\sqrt{\frac{0{,}075(1-0{,}075)}{360}} = 0{,}0358.$$

Do đó, tỉ lệ các phần tử bị đánh dấu của tổng thể là

$$p = 0{,}075 \pm 0{,}0358\ ,$$

$$\text{hay}\quad 0{,}0392 \le p \le 0{,}1108.$$

Ta lại có

$$p = \frac{M}{N}\quad \text{nên}\quad N = \frac{M}{p}.$$

Vậy $\dfrac{1500}{0{,}1108} \le N \le \dfrac{1500}{0{,}0392}$,

hay $13538 \le N \le 38265$.

Như vậy, số thú hiện có trong rừng là từ 13538 đến 38265 con, với độ tin cậy 99%.

2. Xác định kích thước mẫu định tính

Từ công thức tính độ chính xác

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

suy ra số phần tử của mẫu mới

$$n = \left[\frac{Z_\gamma^2 f(1-f)}{\varepsilon^2}\right] + 1,$$

trong đó f là tỉ lệ mẫu ban đầu

Bài 3. *Ở một vùng, khi khám bệnh cho bệnh nhân, người ta thấy tỉ lệ mắc bệnh tai mũi họng là 15%. Để ước lượng xác suất mắc bệnh tai mũi họng của vùng đó với độ tin cậy 95% và sai số không vượt quá 2% thì cần khám tối thiểu bao nhiêu người ?*

Giải

Theo đề bài ta có

- tỉ lệ mắc bệnh của mẫu ban đầu là $f = 0{,}15$;
- độ tin cậy $\gamma = 95\%$ nên $Z_\gamma = 1{,}96$;
- sai số của ước lượng, hay độ chính xác là $\varepsilon \le 0{,}02$.

Vậy mẫu cần tìm phải có số phần tử

$$n \ge \left[\frac{1{,}96^2.\ \ 0{,}15.\ \ (1-0{,}15)}{0{,}02^2}\right] + 1 = 1225.$$

Do đó, tối thiểu cần khám cho 1225 người.

Bài 4. *Lô trái cây của một chủ hàng được đóng thành từng sọt, mỗi sọt 100 trái. Kiểm tra 50 sọt người ta thấy có 450 trái không đạt tiêu chuẩn.*

a) Hãy ước lượng tỉ lệ trái cây không đạt tiêu chuẩn của lô hàng đó với độ tin cậy 95%.

b) Muốn ước lượng tỉ lệ trái cây không đạt tiêu chuẩn với độ chính xác 0,5% thì độ tin cậy đạt được là bao nhiêu ?

c) Muốn ước lượng tỉ lệ trái cây không đạt tiêu chuẩn với độ tin cậy 99% và độ chính xác 1% thì cần kiểm tra bao nhiêu sọt ?

Giải

a) Tổng thể là toàn bộ trái cây của lô hàng. Gọi p là tỉ lệ trái cây không đạt tiêu chuẩn của tổng thể. Ta cần ước lượng p với độ tin cậy 95%.

Ta có kích thước mẫu

$$n = 50 \,.\, 100 = 5000,$$

tỉ lệ mẫu không đạt tiêu chuẩn

$$f = \frac{450}{5000} = 0{,}09.$$

Với độ tin cậy $\gamma = 0{,}95$, tra bảng hàm số Laplace, ta tìm được $Z_\gamma = 1{,}96$.

Từ đó, sai số của ước lượng

$$\varepsilon = 1{,}96\sqrt{\frac{0{,}09.(1-0{,}09)}{5000}} = 0{,}0079.$$

Vậy, tỉ lệ trái không đạt tiêu chuẩn của lô hàng là

$$P = 0{,}09 \pm 0{,}0079,$$

hay

$$8{,}21\% \leq p \leq 9{,}97\%.$$

b) Theo đề bài, ta cần tìm độ tin cậy γ khi biết độ chính xác $\varepsilon = 0{,}005$ và $n = 5000$,

$$f = 0{,}09.$$

Từ công thức

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

suy ra

$$Z_\gamma = \varepsilon\sqrt{\frac{n}{f(1-f)}} = 0{,}005\sqrt{\frac{5000}{0{,}09.0{,}91}} = 1{,}24.$$

Mà $\Phi\,(Z_\gamma) = \dfrac{\gamma}{2}$, nên

$$\gamma = 2\,\Phi\,(Z_\gamma) = 2\,\Phi\,(1{,}24) = 2\,.\,0{,}3925 = 0{,}785.$$

Vậy, độ tin cậy đạt được là 78,5%.

c) Ta cần tìm kích thước mẫu mới khi biết độ tin cậy $\gamma = 99\%$, độ chính xác $\varepsilon = 1\%$ và tỉ lệ mẫu ban đầu f = 0,09. Khi đó $Z_\gamma = 2{,}58$. Do đó

$$n = \left[\frac{2{,}58^2.0{,}09.0{,}91}{0{,}01^2}\right] + 1 = 5452.$$

Số trái cây cần kiểm tra là 5452, và vì mỗi sọt có 100 trái nên số sọt cần kiểm tra là

$$\left[\frac{5452}{100}\right] + 1 = 55.$$

3. Ước lượng giá trị trung bình

Bài 5. *Ở một cửa hàng chế biến thủy sản, theo dõi lượng nước mắm bán ra trong một số ngày, người ta ghi được bảng số liệu sau.*

Số lượng bán ra (lít)	Số ngày
20 – 30	3
30 – 40	8
40 – 50	30
50 – 60	45
60 – 70	20
70 – 80	25
80 – 90	17
90 – 100	9
> 100	4

Hãy ước lượng số lít nước mắm bán ra trung bình mỗi ngày với độ tin cậy 99% trong hai trường hợp

a) biết phương sai $\sigma^2 = 132{,}25$;

b) chưa biết phương sai.

Giải

Ta lập bảng tính các đặc trưng của mẫu định lượng.

x_i	n_i	$x_i n_i$	$x_i^2 n_i$

25	3	75	1875
35	8	280	9800
45	30	1350	60750
55	45	2475	136125
65	20	1300	84500
75	25	1875	140625
85	17	1445	122825
95	9	855	81225
105	4	420	44100
Σ	161	10075	681825

$\overline{X} = \frac{10075}{161} = 62,5776$;

$\overline{X^2} = \frac{681825}{161} = 4234,9379$;

$\hat{S}^2 = 318,9819$;

$S^2 = 320,9755$;

$S = 17,9158$.

Với $\gamma = 99\%$ ta có $Z_\gamma = 2,58$.

a) Theo đề bài, phương sai của tổng thể đã biết,

$$\sigma^2 = 132,65 \text{ nên } \sigma = 11,5174.$$

Do đó

$$\varepsilon = Z_\gamma \frac{\sigma}{\sqrt{n}} = 2,58 \, . \, \frac{11,5174}{\sqrt{161}} = 2,3419.$$

Vậy, số lít nước mắm bán ra trung bình mỗi ngày ở cửa hàng đó là

$$\mu = \overline{X} \pm \varepsilon = 62,5776 \pm 2,3419.$$

hay

$$60,2357 \le \mu \le 64,9195 \text{ (lít)}.$$

b) Phương sai của tổng thể chưa biết, $n \ge 30$ nên độ chính xác là

$$\varepsilon = Z_\gamma \frac{S}{\sqrt{n}} = 2,58 \, . \, \frac{17,9158}{\sqrt{161}} = 3,6429.$$

Vậy $\mu = 62,5776 \pm 3,6429$ (lít).

Bài 6. *Đo đường kính của 20 trục máy do một máy tiện tự động sản xuất ra, ta được kết quả sau (tính bằng mm)*

250 ; 249 ; 251 ; 253 ; 248 ; 250 ; 250 ; 252 ; 257 ; 245 ;
248 ; 247 ; 249 ; 249 ; 250 ; 280 ; 250 ; 247 ; 253 ; 256.

Giả sử đường kính của các trục máy là đại lượng ngẫu nhiên có phân phối chuẩn. Hãy ước lượng đường kính trung bình của các trục máy do máy tiện sản xuất với độ tin cậy 95%.

Giải

Tổng thể là toàn bộ trục máy do máy tiện sản xuất. Ta cần ước lượng đường kính trung bình μ của tổng thể với độ tin cậy 95% và kích thước mẫu n < 30, chưa biết phương sai.

Từ mẫu định lượng đã cho, ta tính được

$$\overline{X} = 251{,}7\ ; \quad S = 7{,}2670.$$

Tra bảng phân phối Student dòng 19, cột 0,05, ta được $T_\gamma = 2{,}093$.

Suy ra

$$\varepsilon = 2{,}093 \ . \ \frac{7{,}267}{\sqrt{20}} = 3{,}4010.$$

Vậy $\mu = 251{,}7 \pm 3{,}401$ (mm).

4. Xác định kích thước mẫu định lượng

Từ công thức tính độ chính xác suy ra công thức tìm số phần tử của mẫu mới với S được tính từ mẫu ban đầu.

Chẳng hạn, $n = \left[\dfrac{Z_\gamma^2 S^2}{\varepsilon^2}\right] + 1.$

Bài 7. *Sai số đo của một loại dụng cụ đo có phân phối chuẩn với độ lệch bằng 20. Cần phải tiến hành bao nhiêu phép đo độc lập để sai số phạm phải không vượt quá 5 với độ tin cậy 90%?*

Giải

Đây là bài toán tìm kích thước mẫu định lượng biết độ tin cậy $\gamma = 0{,}9$, độ chính xác $\varepsilon = 5$, độ lệch chuẩn của tổng thể $\sigma = 20$ đã biết.

Theo công thức tính độ chính xác, ta có

$$\varepsilon = Z_\gamma \frac{\sigma}{\sqrt{n}}.$$

Suy ra

$$n = \left[\left(\frac{Z_\gamma \sigma}{\varepsilon}\right)^2\right] + 1.$$

Tra bảng hàm số Laplace, ta thấy $\Phi\ (1{,}65) = \dfrac{0{,}9}{2} = 0{,}45$. Do đó $Z_\gamma = 1{,}65$.

Vậy,

$$n = \left[\left(\frac{1{,}65.20}{5}\right)^2\right] + 1 = 44.$$

Cần tiến hành 44 phép đo độc lập.

5. Ước lượng phương sai

Bài 8. *Theo dõi số hàng bán được trong một ngày ở một cửa hàng, ta được kết quả ghi ở bảng sau.*

Số hàng bán được (kg)	Số ngày
1900 – 1950	2
1950 – 2000	10
2000 – 2050	8
2050 - 2100	5

Hãy ước lượng phương sai của lượng hàng bán được mỗi ngày với độ tin cậy 95%

Giải

Từ mẫu định lượng đã cho, ta tính được n = 25, $S^2 = 2058,3333$.

Tra bảng phân phối "khi bình phương" dòng 24, cột 0,05 và cột 0,95, ta được

$$\chi_1 = 36,415 \text{ , } \chi_2 = 13,848.$$

Suy ra

$$\frac{(n-1)S^2}{\chi_1} = 1356,5838 \text{ ; } \quad \frac{(n-1)S^2}{\chi_2} = 3567,3021.$$

Vậy, khoảng ước lượng của phương sai của tổng thể là

$$1356,5838 < \sigma^2 < 3567,3021.$$

C. BÀI TẬP

1. Kiểm tra ngẫu nhiên 500 sản phẩm của một nhà máy thì thấy có 360 sản phẩm loại một. Hãy ước lượng tỉ lệ sản phẩm loại một tối thiểu của cả nhà máy với độ tin cậy 95%.

2. Lấy ngẫu nhiên 400 hộp từ một kho đồ hộp để kiểm tra thì thấy có 20 hộp bị hỏng. Từ kết quả kiểm tra đó, hãy ước lượng tỉ lệ phế phẩm của kho hàng với độ tin cậy 95,45%.

3. Tại một khu rừng nguyên sinh, người ta đeo vòng vào chân của 1200 con chim. Sau một thời gian bắt lại 250 con thì thấy 40 con có đeo vòng. Hãy ước lượng số chim trong khu rừng đó với độ tin cậy 99%.

4. Muốn biết số cá có trong một hồ lớn, người ta bắt lên 2000 con, đánh dấu xong lại thả chúng xuống hồ . Sau đó người ta bắt lên 400 con thì thấy có 55 con bị đánh dấu. Với độ tin cậy 0,95 hãy ước lượng số cá trong hồ. Cho biết mỗi con cá có khối lượng trung bình 800 gam và mỗi kilôgam cá bán được 22000đ. Tính doanh thu tối thiểu khi bán hết số cá trong hồ.

5. Để xác định định mức thời gian gia công một chi tiết máy, người ta tiến hành thử nghiệm gia công 25 chi tiết. Kết quả trên tập mẫu thu được như sau.

Thời gian trung bình $\overline{X}$ = 20 phút, độ lệch mẫu hiệu chỉnh S = 2,02 phút.

Với độ tin cậy 90%, hãy xác định thời gian gia công trung bình tối đa đối với loại chi tiết đó, giả sử thời gian gia công tuân theo quy luật phân phối chuẩn.

6. Cân thử khối lượng của một số gia cầm ở một trại chăn nuôi, ta được kết quả sau (tính bằng kilôgam).

3,25 ; 2,5 ; 4 ; 3,75 ; 3,8 ; 3,9 ; 4,02 ;
3,8 ; 3,2 ; 3,82 ; 3,4 ; 3,6 ; 3,75 ; 4 ; 3,5

Giả sử khối lượng gia cầm tuân theo quy luật phân phối chuẩn với phương sai 0,01. Hãy ước lượng khối lượng trung bình của một con gia cầm với độ tin cậy 99%.

7. Nghiên cứu điểm trung bình môn Xác suất thống kê và ứng của 50 sinh viên ta có kết quả : $\overline{X} = 6,1$; $S = 1,0$. Tìm khoảng ước lượng cho điểm trung bình với độ tin cậy 95%. Nếu khoảng ước lượng có độ dài bằng 2 thì độ tin cậy đạt được là bao nhiêu ?

8. Đường kính trục của một loại sản phẩm là đại lượng ngẫu nhiên có phân phối chuẩn. Kiểm tra 6 trục chọn ngẫu nhiên ta có kết quả (tính bằng centimet) 7,1 ; 6,6 ; 9,7 ; 10,6 ; 7,5 ; 9,1.

a) Tìm ước lượng không chệch cho phương sai đường kính trục.

b) Tìm khoảng ước lượng cho phương sai đó với độ tin cậy 0,95.

9. Đo áp lực X (tính bằng kg/cm^2) của một số thùng chứa, ta được bảng kết quả sau.

Áp lực	200	210	220	230	240	250
Số thùng	10	26	56	64	30	14

Biết rằng áp lực là một đại lượng có phân phối chuẩn.

a) Với $\gamma = 0,99$, hãy tìm khoảng ước lượng của áp lực trung bình.

b) Tìm khoảng ước lượng của phương sai của áp lực với độ tin cậy 0,95.

10. Cân thử 100 trái cây của một nông trường, ta có kết quả sau đây.

Khối lượng (g)	Số trái
35 – 55	3
55 – 75	10
75 – 95	25
95 – 115	35
115 – 135	20
135 – 155	6
155 - 175	1

a) Tìm ước lượng không chệch cho khối lượng trung bình của một trái cây trong nông trường.

b) Tìm ước lượng không chệch cho phương sai của khối lượng trái cây trong nông trường.

c) Xem các trái có khối lượng không quá 95gam là trái cây loại hai.

Tìm ước lượng không chệch cho tỉ lệ trái cây loại hai trong nông trường.