Lý thuyết ước lượng:

Khái niệm:

+ thống kê mô tả: phương pháp tổ chức và tóm tắc dữ liệu dưới dạng đồ thị, biểu đồ, ..

+ thống kê suy diễn: rút ra kết luận từ dữ liệu

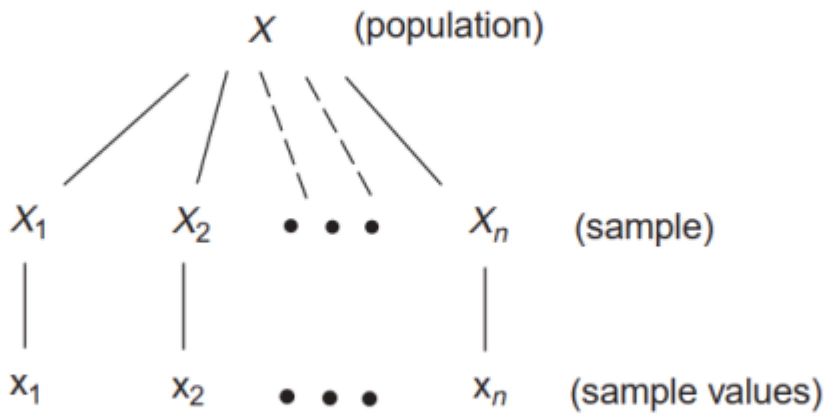
Bài toán: khảo sát chiều cao tất cả thanh niên Việt Nam

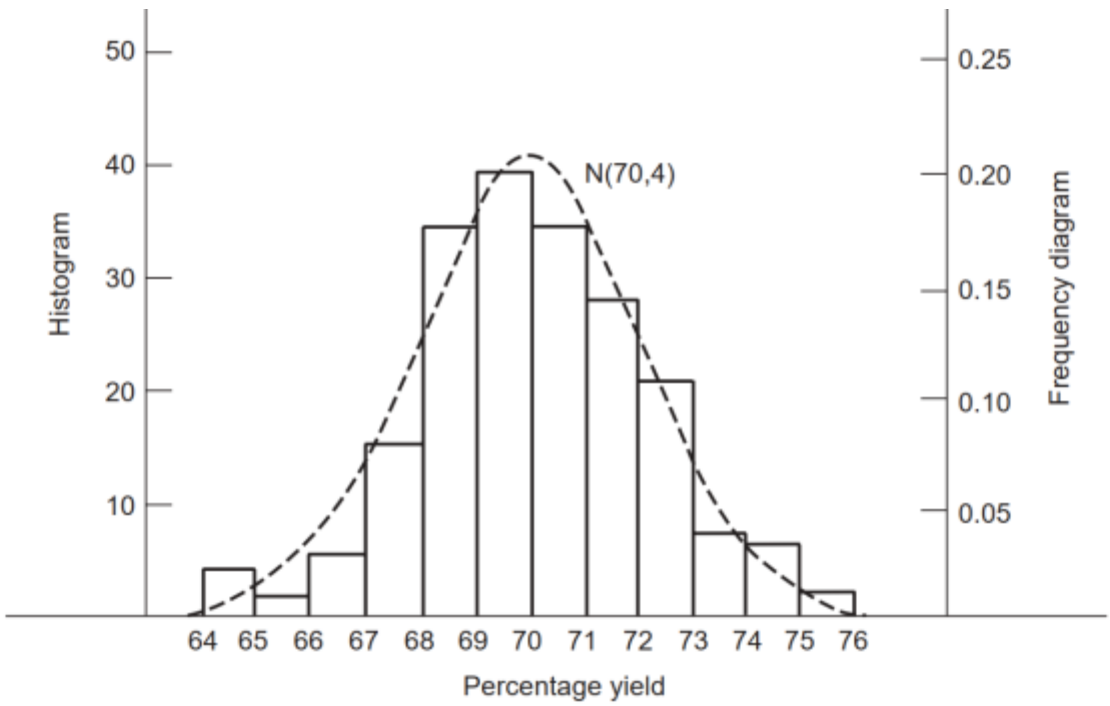
Chọn ra ngẫu nhiên n thanh niên đo chiều cao → $\textit{ước lượng chiều tb của ng VN}$

+ X: chiều cao tổng của người Việt Nam. → $\textit{tổng thể}\ (\textit{population})$

+ $X_1, X_2, \cdots, X_n$: $\textit{bnn có cùng phân phối } X \ \rightarrow \textit{mẫu}$

+ $x_1, x_2, \cdots, x_n \rightarrow \textit{giá trị mẫu}$





$$X \sim N(\mu, \sigma^2)$$

Tham số của mẫu:

Trung bình mẫu:

$$\overline{X} = \frac{X_1 + \cdots + X_n}{n}$$

Phương sai mẫu

$$S^2 = \frac{1}{n-1}\sum_{1}^{n}\left(X_i - \overline{X}\right)^2$$

Ví dụ: So sánh giá cà phê tại 4 cửa hàng tạp hóa được chọn ngẫu nhiên ở Thủ Đức cho thấy các mức tăng so với tháng trước là 12, 15, 17 và 20 nghìn đồng cho một túi 1 kg. Tìm trung bình, phương sai của mẫu này.

$$\overline{x} = \frac{12 + 15 + 17 + 20}{4} = 16,$$

$$s^2 = \frac{(12-16)^2 + (15-16)^2 + (17-16)^2 + (20-16)^2}{3}$$

Ước lượng:

+ đánh giá: tiêu chuẩn

$$f(x, \mu, \sigma, ..)$$

Giả sử tham số tổng thể cần ước lượng θ

$$\Theta = h(X_1, X_2, \ldots., X_n)$$

+ Ước lượng không lệch

$$E(\Theta) = \theta$$

$$MSE = E(\Theta - \theta)^2 \ \rightarrow min$$

Phương phướng:

+ phương pháp ước lượng điểm

+ Khoảng ước lượng: μ → tìm [a,b]

$$P(\mu \in [a, b]) \rightarrow 1 - \alpha$$

Bài toán cụ thể:

1) Ước lượng cho trung bình μ của $N(\mu, \sigma^2)$ với σ^2

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Đặt $1 - \alpha$ độ tin cậy

$$\left[\overline{x} - z_{\alpha/2}\frac{\sigma}{n}, \overline{x} + z_{\alpha/2}\frac{\sigma}{n}\right]$$

Trong đó

$$\Phi\left(z_{\alpha/2}\right) = 1 - \alpha/2$$

Số $z_{\alpha/2}\frac{\sigma}{n}$ sai số của ước lượng, độ chính xác.

BT: Giả sử rằng thời gian mua sắm của khách hàng tại một trung tâm thương mại có phân phối chuẩn với độ lệch chuẩn tổng thể là 20 phút. Chọn ngẫu nhiên 64 người đã mua sắm ở trung tâm đó. Người ta thấy rằng thời gian mua sắm trung bình của 64 người này là 75 phút. Tìm thời gian mua sắm trung bình của khách hàng tại trung tâm này với độ tin cậy 95%.

Bài làm: Gọi X thời gian mua sắm của kh tại trung tâm

$$X \sim N(\mu, 20^2)$$

Khoảng ước lượng với độ tin cậy $1 - \alpha = 95\%$

$$\left[\overline{x} - z_{\alpha/2}\frac{\sigma}{n}, \overline{x} + z_{\alpha/2}\frac{\sigma}{n}\right]$$

Trong đó, $\overline{x} = 75,\ n = 64$ và

$$\Phi\left(z_{\alpha/2}\right) = 1 - \alpha/2 = 1 - 0.025 = 0.9725 \rightarrow z_{\alpha/2} = 1.96$$

Thế vào trên ta được kết quả.

2) Ước lượng trung bình μ của $N(\mu, \sigma^2)$ *với* σ^2

Th1: kích thước mẫu $n \geq 30$

Khoảng ước lượng với độ tin cậy $1 - \alpha$

$$\left[\overline{x} - z_{\alpha/2}\frac{s}{n}, \overline{x} + z_{\alpha/2}\frac{s}{n}\right]$$

Th2: mẫu bé $n < 30$

Khoảng ước lượng với độ tin cậy $1 - \alpha$

$$\left[\overline{x} - t_{n-1,\alpha/2}\frac{s}{n}, \overline{x} + t_{n-1,\alpha/2}\frac{s}{n}\right]$$

Trong đó

$$P\left(t > t_{n-1,\alpha/2}\right) = \frac{\alpha}{2} \rightarrow \ tra\ bảng\ pp\ Student$$

BT: Theo một thống kê cho thấy số thu nhập của 7 công nhân trong năm 2021 của một công ty được cho như sau (đơn vị triệu đồng)

109,2 118 121,8 126,2 143,2 168,8 198,6. Giả sử thu nhập trong năm 2021 của công ty có phân phối chuẩn. Tính khoảng thu nhập trung bình các công nhân của công ty này với độ tin cậy 99%

Bài làm: Gọi X thu nhập của nhân viên cty $X \in N(\mu, \sigma^2)$

Khoảng ước lượng với độ tin cậy $1-\alpha = 0.99$ ($\rightarrow \alpha = 0.01$)

$$\left[\overline{x} - t_{n-1,\alpha/2}\frac{s}{n}, \overline{x} + t_{n-1,\alpha/2}\frac{s}{n}\right]$$

Trong đó $\overline{x} = 140.826$, $s = 32.205$,

$$P(t > t_{n-1,\alpha/2}) = 0.005 \rightarrow t_{n-1,\alpha/2} = 3.7074$$

BT: Chọn ngẫu nhiên 30 người để kiểm tra thời gian sử dụng chiếc điện thoại di động đầu tiên. Người ta thấy rằng thời gian sử dụng trung bình của 30 người này là 5,6 năm. Giả sử thời gian sử dụng chiếc điện thoại di động đầu tiên có phân phối chuẩn với độ lệch chuẩn là 0,8 năm. Tính khoảng thời gian trung bình sử dụng chiếc điện thoại đầu tiên với độ tin cậy 99%. (Đáp số: [5,22; 5,98])

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Đặt $1-\alpha$ độ tin cậy

$$\left[\overline{x} - z_{\alpha/2}\frac{\sigma}{n}, \overline{x} + z_{\alpha/2}\frac{\sigma}{n}\right]$$

Trong đó

$$\Phi(z_{\alpha/2}) = 1 - \alpha/2$$

24410197, 24410186, 24410191, 24410218, 24410162, 24410156, 24410221, 24410131, 24410248, 24410253

Ví dụ 6.10 Kiểm tra tuổi thọ (tính bằng giờ) của 50 bóng đèn do nhà máy A sản xuất, người ta được bảng số liệu sau

Tuổi thọ	(3350;3450]	(3450;3550]	(3550;3650]	(3650;3750]
Số bóng đèn	10	20	12	8

a. Ước tính tuổi thọ trung bình của các bóng đèn do nhà máy A sản xuất với độ tin cậy 97%.

b. Dựa vào mẫu trên để ước lượng tuổi thọ trung bình của các bóng đèn do nhà máy A sản xuất có độ chính xác 29 giờ thì phải đảm bảo độ tin cậy là bao nhiêu?

24410196, 24410192, 24410238, 24410146

$$z_{\alpha/2}\frac{s}{n} = 29 \ \rightarrow \alpha =?$$

Giải. a. Ước tính tuổi thọ trung bình của các bóng đèn do nhà máy A sản xuất với độ tin cậy 97%. (Kích thước mẫu bằng 50 và chưa biết độ lệch chuẩn tổng thể)

- Trung bình mẫu: $\overline{x} = 3536$ (giờ)
- Độ lệch chuẩn mẫu: $s = 98,478$ (giờ)
- Độ tin cậy $1-\alpha = 0,97$. Suy ra $\Phi(z_{\alpha/2}) = 1-\alpha/2 = 0,985$. Do đó $z_{\alpha/2} = 2,17$.
- Độ chính xác:

$$z_{\alpha/2}\frac{s}{\sqrt{n}} = 2,17\frac{98,478}{\sqrt{50}} = 30,221.$$

- Khoảng tin cậy của tuổi thọ trung của bóng đèn với độ tin cậy 97% là

$$[3505,779; 3566,221].$$

Giải. b. Ta có độ chính xác bằng 29 giờ, tức là

$$z_{\alpha/2}\frac{s}{\sqrt{n}} = 29$$

Suy ra

$$z_{\alpha/2} = 29\frac{\sqrt{n}}{s} = 2,08$$

Do đó

$$\Phi(z_{\alpha/2}) = \Phi(2,08) = 1-\frac{\alpha}{2}.$$

Tìm trong phụ lục A4, ta có $\Phi(2,08) = 0,9812$ và do đó $\alpha = 0,0376$.
Như vậy, độ tin cậy là $1-\alpha = 0,9624$.

- p : tỉ lệ tổng thể (tỉ lệ phần tử có tính chất $\mathcal{P}$ trong tổng thể)
- $\hat{p}$: tỉ lệ mẫu (tỉ lệ phần tử có tính chất $\mathcal{P}$ trong mẫu)
- Nếu $np, n(1-p) \geq 5$ thì tỉ lệ mẫu ngẫu nhiên $\hat{p}$ có phân phối chuẩn $N(p; \sigma^2_{\hat{p}})$

- Với độ tin cậy $1-\alpha$, khoảng tin cậy chứa tỉ lệ tổng thể là

$$\left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}; \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

  trong đó $\Phi(z_{\alpha/2}) = 1 - \dfrac{\alpha}{2}$ (xem phụ lục A4).

- Độ chính xác (sai số) là $z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$.

Thăm dò ý kiến của 100 người dân được chọn ngẫu nhiên tại một thành phố cho thấy có 80% trong số này ủng hộ ứng viên A. Với độ tin cậy 98%, hãy ước lượng tỉ lệ của tất cả các người dân ủng hộ ứng viên A tại thành phố này.

Giải. Theo đề bài

- Tỉ lệ mẫu $\hat{p} = 0,8$
- Kích thước mẫu $n = 100$
- Độ tin cậy $1-\alpha = 0,98$ suy ra $1 - \dfrac{\alpha}{2} = 0,99$ Do đó $z_{\alpha/2} = 2,33$
- Độ chính xác (sai số) là

$$z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 2,33\frac{0,4}{10} = 0,0932.$$

- Khoảng tin cậy $[0,7068; 0,8932]$.

Như vậy có từ 70,68% đến 89,32% người dân ủng hộ ứng viên A.

Ví dụ 6.12 Lấy ngẫu nhiên 344 sản phẩm trong một kho hàng để kiểm tra thì thấy có 261 sản phẩm tốt.

a. Dựa vào mẫu trên, để ước tính tỉ lệ sản phẩm bị lỗi có độ chính xác là 0,059 thì độ tin cậy bằng bao nhiêu?

b. Nếu muốn ước lượng tỉ lệ sản phẩm bị lỗi với độ chính xác nhỏ hơn $0,02$ với độ tin cậy 95% thì cần kiểm tra ít nhất bao nhiêu sản phẩm?

24410221, 24410157, 24410157

---

a)

Tỉ lệ mẫu $\hat{p} = \frac{261}{344} = 0{,}759$

Kích thước mẫu $n = 344$

Độ chính xác là

$$z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = z_{\alpha/2}\sqrt{\frac{0{,}759(1-0{,}759)}{344}} = 0{,}059$$

Suy ra $z_{\alpha/2} = 2{,}58$. Do đó độ tin cậy là 99%.

b) Độ chính xác nhỏ hơn 0,02

$$\frac{z_{\alpha/2}}{2\sqrt{n}} < 0{,}02 \text{ hay } n > \frac{z_{\alpha/2}^2}{4.0{,}02^2}$$

Vì độ tin cậy $1-\alpha = 0{,}95$ nên $z_{\alpha/2} = 1{,}96$

Do đó

$$n > \frac{1{,}96^2}{4.0{,}02^2} = 2401$$

Như vậy, cần kiểm tra ít nhất 2402 sản phẩm.

Bài 6.3 Một mẫu ngẫu nhiên gồm 100 cửa hàng mua máy in hiệu T cho thấy rằng 59 cửa hàng có kế hoạch gia tăng việc mua hàng của mình trong năm tới. Hãy ước lượng tỷ lệ các cửa hàng trong tổng thể tất cả các cửa hàng mua máy in hiệu T mà có kế hoạch gia tăng việc mua hàng của mình trong năm tới với độ tin cậy 95%.

Bài 6.4 Chọn ngẫu nhiên 500 laptop trong một kho chứa thì thấy có 27 laptop hiệu UIT.

a. Dựa vào mẫu trên, để ước lượng tỉ lệ laptop hiệu UIT trong toàn bộ kho có độ chính xác 0,0177 thì phải đảm bảo độ tin cậy bằng bao nhiêu?

b. Dựa vào mẫu trên, nếu muốn có độ chính xác của ước lượng tỉ lệ laptop UIT nhỏ hơn 0,01 với độ tin cậy 95% thì cần chọn ít nhất bao nhiêu laptop nữa?

$n = 100$

$x = 59$

$$\hat{p} = \frac{59}{100} = 0.59$$

$1 - \alpha = 0.95 => a = 0.05$

$$z_{\frac{a}{2}} = 1 - \frac{0.05}{2} = 0.975 => 1.96$$

$$\left[\hat{p} - z_{\frac{a}{2}} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}; \hat{p} + z_{\frac{a}{2}} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right] => \left[0.59 - 1.96 \times \sqrt{\frac{0.59(1-0.59)}{100}}; 0.59\right.$$

$$\left.1.96 \times \sqrt{\frac{0.59(1-0.59)}{100}}\right] => [0.49, 0.68]$$

$$\hat{p} = \frac{27}{500} = 0.054$$

Công thức tính cỡ mẫu:

$$n = \frac{z^2 \times \hat{p} \times (1 - \hat{p})}{E^2}$$

Với độ tin cậy 95%, ta có $z = 1.96$ và sai số mong muốn $E = 0.01$.

Thay số:

$$n = \frac{(1.96)^2 \times 0.054 \times (1 - 0.054)}{(0.01)^2}$$

$$= \frac{3.8416 \times 0.054 \times 0.946}{0.0001}$$

$$= \frac{3.8416 \times 0.051084}{0.0001}$$

$$= \frac{0.1964}{0.0001}$$

$$= 1964$$

Cần ít nhất **1964 laptop** để đạt độ chính xác mong muốn.

24410162

$n = 500$

$x = 27$

$\hat{p} = \frac{27}{500} = 0.054$

$0.0177 = z_{\frac{a}{2}} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$0.0177 = z_{\frac{a}{2}} \times \sqrt{\frac{0.054(1-0.054)}{500}}$

$z_{\frac{a}{2}} \times 0.01 = 0.0177$

$z_{\frac{a}{2}} = 1.721$

24410069, 24410146, 24410174, 24410162,24410212, 24410162

24410191, 24410157, 24410212, 24410248, **24410218, 24410169, 24410162, 24410152, 24410157, 24410260, 24410246, 24410191, 24410186, 24410216**