

CHƯƠNG 5

1) CÁC KHÁI NIỆM CƠ BẢN

1) Các cách lấy mẫu thông dụng

C1: Lấy ngẫu nhiên n phần tử: *phân phối siêu bội*

C2: Lấy lần lượt n phần tử

C3: Lấy có hoàn lại n phần tử: *phân phối nhị thức*

* Về mặt xác suất: $C1 = C2$

* Khi $n \ll N$ thì C1 xấp xỉ C3

Ta quy ước là *mẫu được lấy theo cách có hoàn lại*.

Mẫu gồm có: *mẫu ngẫu nhiên* và *mẫu cụ thể*. Cần phân biệt rõ mẫu ngẫu nhiên và mẫu cụ thể.

2) Các đặc trưng số cơ bản của tổng thể và mẫu

Xét tổng thể về mặt định lượng:

Tổng thể được đặc trưng bởi dấu hiệu nghiên cứu X , X là đại lượng ngẫu nhiên.

$E(X) = \mu$ là *trung bình tổng thể*

$Var(X) = \sigma^2$ là *phương sai tổng thể*

σ là *độ lệch chuẩn tổng thể*

Xét tổng thể về mặt định tính:

Tổng thể có kích thước N , trong đó có M phần tử có *tính chất A quan tâm*.

$p = M/N$ gọi là *tỷ lệ tổng thể*

Tương tự, ta cũng có trung bình mẫu \bar{x} , phương sai mẫu (đã hiệu chỉnh) s^2 , tỷ lệ mẫu f .

3) Cách tính các đặc trưng số cơ bản của mẫu (dạng cụ thể):

* Định lượng:

Tính theo định nghĩa:

$$\text{Trung bình mẫu : } \bar{x} = \frac{1}{n} \sum x_i$$

$$\text{Phương sai mẫu (đã hiệu chỉnh) : } s^2 = \frac{1}{n-1} \left(\sum (x_i - \bar{x})^2 \right) = \frac{1}{n-1} \left(\sum x_i^2 - n(\bar{x})^2 \right)$$

$$\text{Độ lệch chuẩn mẫu (đã hiệu chỉnh) : } s = \sqrt{s^2}$$

$$\text{Sai số chuẩn mẫu (đã hiệu chỉnh) : } \frac{s}{\sqrt{n}}$$

Thực hành:

Mẫu dạng điểm					
x_i	x_1	...	x_i	...	x_k
n_i	n_1	...	n_i	...	n_k

* x_i là giá trị thu thập được từ thực tế

* n_i là số lần xuất hiện của x_i trong mẫu khảo sát

$$\bar{x} = \frac{1}{n} \sum n_i x_i$$

$$s^2 = \frac{1}{n-1} \left(\sum n_i (x_i - \bar{x})^2 \right) = \frac{1}{n-1} \left(\sum n_i x_i^2 - n(\bar{x})^2 \right)$$

Lưu ý 1:

Khi lấy kết quả trong máy tính Casio fx-570VN Plus cần phân biệt sx (đã hiệu chỉnh) và σx .

Lưu ý 2:

Với **mẫu dạng khoảng** thì đưa về dạng điểm bằng cách lấy trung bình cộng 2 giá trị đầu và cuối của từng khoảng, rồi dùng công thức dạng điểm để tính.

Lưu ý 3:

Tính trung bình mẫu và phương sai mẫu của mẫu gộp:

Mẫu 1 có cỡ mẫu n_x , trung bình mẫu \bar{x} , phương sai mẫu s_x^2

Mẫu 2 có cỡ mẫu n_y , trung bình mẫu \bar{y} , phương sai mẫu s_y^2

Mẫu gộp có cỡ mẫu n_z , trung bình mẫu \bar{z} , phương sai mẫu s_z^2

Mẫu 1:

$$\bar{x} = \frac{1}{n_x} \sum x \rightarrow \sum x = n_x \cdot \bar{x}$$

$$s_x^2 = \frac{1}{n_x - 1} \sum (x - \bar{x})^2 = \frac{1}{n_x - 1} \{ \sum x^2 - n_x (\bar{x})^2 \} \rightarrow \sum x^2 = (n_x - 1) \cdot s_x^2 + n_x (\bar{x})^2$$

Tương tự cho Mẫu 2

Mẫu gộp:

$$\bar{z} = \frac{1}{n_x + n_y} (\sum x + \sum y)$$

$$s_z^2 = \frac{1}{n_x + n_y - 1} \{ \sum x^2 + \sum y^2 - (n_x + n_y) \cdot (\bar{z})^2 \}$$

Lưu ý 4:

Mẫu cụ thể 2 chiều

Ta có bảng số liệu về 2 chỉ tiêu X, Y của 1 loại sản phẩm như sau:

X \ Y	Y				
	5	10	15	20	25
2	2	1			
4		2	2		
6		4	6	3	1
8			4	3	2

Xác định các đặc trưng số của mẫu về chỉ tiêu X, chỉ tiêu Y?

HD:

Ta có bảng tần số thực nghiệm của X và Y như sau:

x_i	2	4	6	8
n_i	3	4	14	9

y_i	5	10	15	20	25
n_i	2	7	12	6	3

$$\bar{x} = \frac{1}{n} \sum n_i x_i \quad ; \quad \bar{y} = \frac{1}{n} \sum n_i y_i$$

$$s_x^2 = \frac{1}{n-1} \left(\sum n_i x_i^2 - n(\bar{x})^2 \right) \quad ; \quad s_y^2 = \frac{1}{n-1} \left(\sum n_i y_i^2 - n(\bar{y})^2 \right)$$

*** Định tính:**

Trong thực hành ta xác định tỷ lệ mẫu:

$$f = m/n$$

Với: n : cỡ mẫu

m : số phần tử có tính chất A quan tâm trong mẫu

II) CÁC KẾT QUẢ VỀ QUY LUẬT PHÂN PHỐI MẪU

1) Định lý (định lượng):

Tổng thể có quy luật phân phối X với: $E(X) = \mu$ và $\text{var}(X) = \sigma^2$

Trung bình mẫu ngẫu nhiên \bar{X}

* Lấy mẫu có hoàn lại cỡ mẫu n :

$$E(\bar{X}) = \mu \text{ và } \text{var}(\bar{X}) = \frac{\sigma^2}{n}$$

* Lấy mẫu không hoàn lại cỡ mẫu n :

$$E(\bar{X}) = \mu \text{ và } \text{var}(\bar{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$$

Với N là kích thước tổng thể

$\frac{N-n}{N-1}$ gọi là hệ số hiệu chỉnh

2) Định lý (định lượng):

Lấy mẫu có hoàn lại cỡ mẫu n

Ta có $X \sim N(\mu, \sigma^2)$ thì $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

Do đó tất cả công thức của phân phối chuẩn ở chương 3 đều dùng lại được hết.

$$P(a < X < b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \Rightarrow P(a < \bar{X} < b) = \Phi\left(\frac{b-\mu}{\sigma/\sqrt{n}}\right) - \Phi\left(\frac{a-\mu}{\sigma/\sqrt{n}}\right)$$

$$P(|\bar{X} - \mu| < \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\sigma/\sqrt{n}}\right) \Rightarrow P(|\bar{X} - \mu| < \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\sigma/\sqrt{n}}\right)$$

3) Định lý (định tính):

Lấy mẫu có hoàn lại cỡ mẫu n

Tỷ lệ mẫu ngẫu nhiên $F = \frac{1}{n} \sum X_i$

Các X_i độc lập, có cùng quy luật phân phối xác suất 0-1 như sau:

X_i	0	1
P	$q = 1-p$	p

p là tỷ lệ tổng thể có tính chất A quan tâm

$X_i = 1$ có tính chất A mà ta quan tâm

$X_i = 0$ không có tính chất A mà ta quan tâm

$$E(F) = E(X_i) = p, \quad \text{var}(F) = \frac{\text{var}(X_i)}{n} = \frac{pq}{n}$$

4) Định lý (định lượng):

Tổng thể có phân phối X với $E(X) = \mu$ và $\text{var}(X) = \sigma^2$

Nếu lấy mẫu ngẫu nhiên có hoàn lại với cỡ mẫu n

$$\text{Phương sai mẫu ngẫu nhiên } S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

$$E(S^2) = \sigma^2$$

Lưu ý:

Nếu lấy mẫu ngẫu nhiên không hoàn lại với cỡ mẫu n thì $E(S^2) \neq \sigma^2$

<https://sites.google.com/a/ueh.edu.vn/phamtricao/>
<https://sites.google.com/site/phamtricao/>

