Chương 5
LÝ THUYẾT MẪU

5.1. MỘT SỐ KHÁI NIỆM VỀ MẪU

1. Tổng thể và mẫu

Tập hợp có phần tử là tất cả các đối tượng mà chúng ta nghiên cứu gọi là tổng thể. Tổng thể còn gọi là tập chính hay đám đông.

Số phần tử của tổng thể gọi là kích thước của tổng thể.

Khi nghiên cứu một tổng thể thường chúng ta không thể nghiên cứu hết tất cả các phần tử của nó vì nhiều lý do, chẳng hạn: do kích thước của tổng thể quá lớn; do không xác định được hết phần tử của tổng thể; do xem xét phần tử nào cũng là phá hủy luôn phần tử đó v.v... Vì vậy để nghiên cứu một tổng thể chúng ta chỉ có thông tin đầy đủ về một bộ phận của tổng thể.

Ta gọi một bộ phận gồm n phần tử, có thể phân biệt hoặc không, các phần tử của một tổng thể là một mẫu kích thước n.

Thống kê toán học cung cấp cơ sở toán học để, từ số liệu quan sát được trên một mẫu, cho ta phương pháp đánh giá theo xác suất về toàn bộ tổng thể.

2. Các loại mẫu

Mẫu mà chúng ta nghiên cứu được chọn theo một cách thức nào đó mang tính ngẫu nhiên, khách quan, gọi là mẫu ngẫu nhiên.

1- Phân loại mẫu theo phương pháp chọn mẫu

- *Mẫu không hoàn lại*: là mẫu được chọn bằng cách phần tử đã lấy ra quan sát thì loại khỏi tổng thể rồi mới lấy phần tử tiếp theo.

Mẫu không hoàn lại còn gọi là mẫu không lặp. Một mẫu không lặp kích thước n có các phần tử là một tập con gồm n phần tử của tổng thể.

- *Mẫu hoàn lại*: là mẫu được chọn bằng cách phần tử đã lấy ra quan sát được bỏ trở lại tổng thể rồi mới lấy phần tử tiếp theo.

Mẫu hoàn lại còn gọi là mẫu lặp. Trong mẫu lặp, một phần tử của tổng thể có thể được chọn nhiều lần.

Trường hợp kích thước tổng thể lớn thì mẫu không hoàn lại và mẫu hoàn lại có thể coi là như nhau. Trong lý thuyết tổng quát, chúng ta luôn giả thiết mẫu là mẫu hoàn lại.

2- Phân loại mẫu theo mục đích nghiên cứu

- *Mẫu định tính*: là mẫu mà ta chỉ quan tâm đến các phần tử của nó có một tính chất A nào đó hay không

Trường hợp này mẫu được cho dưới dạng:

+ Kích thước mẫu: n

+ Số phần tử có tính chất A: m.

- *Mẫu định lượng*: là mẫu mà ta quan tâm đến một yếu tố về lượng của các phần tử như khối lượng, chiều dài, nhiệt độ... Trường hợp này, một mẫu kích thước n được cho dưới dạng tổng quát:

X = (X1, X2, ..., Xn)

trong đó phần tử thứ i của mẫu nhận giá trị Xi (i =).

Trong mẫu tổng quát X1, X2, ..., Xn là các đại lượng ngẫu nhiên. Nếu Xi nhận giá trị cụ thể xi thì ta được mẫu cụ thể:

x = (x1, x2, ..., xn)

hay dưới dạng thu gọn:

x	x_1	x_2	...	x_k
Tần số	n_1	n_2	...	n_k

trong đó: x1 < x2 < ... < xk - là các giá trị khác nhau mà các phần tử trong mẫu nhận

ni - là số phần tử trong mẫu nhận giá trị xi (i = $\overline{1,k}$),

n1 + n2 + ... + nk = n.

Nếu ni là số phần tử trong mẫu nhận giá trị thuộc khoảng (ai, bi) thì ta đưa mẫu này về mẫu dạng đang xét bằng cách đặt:

$$x_i = \frac{a_i + b_i}{2}, \quad i = \overline{1,k}$$

5.2. CÁC ĐẶC TRƯNG MẪU

1- Tỷ lệ mẫu

Cho mẫu định tính kích thước n, trong đó có m phần tử có tính chất A. Khi đó ta gọi:

$$f = f_n = \frac{m}{n}$$

là tỷ lệ mẫu.

2- Trung bình mẫu và phương sai mẫu

Xét mẫu định lượng thu gọn:

X	x_1	x_2	...	x_k
Tần số	n_1	n_2	...	n_k

Ta gọi bảng:

X	x_1	x_2	...	x_k
P	p_1	p_2	...	p_k

trong đó: $p_i = \dfrac{n_i}{n} \quad i = \overline{1,k}$

là bảng phân phối xác suất mẫu.

Ta nhận xét rằng: bảng phân phối xác suất mẫu là bảng phân phối xác suất của đại lượng ngẫu nhiên.

$$x = \{x1, x2, ..., xk\}$$

với:

$$P(x = x_i) = p_i = \frac{n_i}{n}$$

là tần suất xuất hiện của xi trong mẫu.

Ta gọi kỳ vọng, phương sai của đại lượng ngẫu nhiên này là trung bình mẫu và phương sai mẫu. Cụ thể ta có:

- *Trung bình mẫu là*:-

$$\overline{x} = x_1 p_1 + x_2 p_2 + \ldots + x_k p_k = \frac{x_1 n_1 + x_2 n_2 + \ldots + x_k n_k}{n}$$

Phương sai mẫu là:

$$\hat{s}^2 = (x_1 - \overline{x})^2 p_1 + (x_2 - \overline{x})^2 p_2 + \ldots + (x_k - \overline{x})^2 p_k$$

$$= \frac{(x_1 - \overline{x})n_1 + (x_2 - \overline{x})^2 n_2 + \ldots + (x_k - \overline{x})^2 n_k}{n}$$

Đặt: $$\overline{x^2} = x_1^2 p_1 + x_2^2 p_2 + \ldots + x_k^2 p_k = \frac{x_1^2 n_1 + x_2^2 n_2 + \ldots + x_k^2 n_k}{n}$$

theo tính chất của phương sai ta có:

$$\widehat{s}^2 = \overline{x^2} - \overline{x}^2$$

- *Phương sai mẫu hiệu chỉnh là*:

$$s^2 = \frac{n\widehat{s}^2}{n-1} = \frac{(x_1 - \overline{x})^2 n_1 + (x_2 - \overline{x})^2 n_2 + \ldots + (x_k - \overline{x}) n_k}{n-1}$$

- Tương tự với đại lượng ngẫu nhiên, ta cũng gọi:

$\widehat{s} = \sqrt{\widehat{s}^2}$ - là *độ lệch mẫu*

$s = \sqrt{s^2}$ - là *độ lệch mẫu hiệu chỉnh.*

Để tính các đặc trưng này ta thường lập bảng sau:

x_i	n_i	$x_i n_i$	$x_i^2 n_i$
x_1	n_1	$x_1 n_1$	$x_1^2 n_1$
x_2	n_2	$x_2 n_2$	$x_1^2 n_1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_k	n_k	$x_k n_k$	$x_k^2 n_k$
Σ	n	$n\overline{x}$	$n\overline{x}^2$

Từ bảng này ta có:

$$\overline{x} = \frac{\Sigma x_i n_i}{n}, \quad \overline{x}^2 = \frac{\Sigma x_i^2 n_i}{n}$$

$$\widehat{s}^2 = \overline{x}^2 - \overline{x}^2, \quad s^2 = \frac{n\widehat{s}^2}{n-1}$$

Ví dụ 5.1. Cho mẫu:

x	5	10	15	20	25
n_i	15	25	30	20	10

Hãy tính các đặc trưng mẫu.

Giải. Ta có bảng tính:

X_i	n_i	$x_i n_i$	$x_i^2 n_i$
5 10 15 20 25	15 25 30 20 10	75 250 450 400 250	375 2500 6750 8000 6250
Σ	100	1425	23875

Từ bảng ta có:

$$\overline{x} = \frac{1425}{100} = 14{,}25; \quad \overline{x}^2 = \frac{23875}{100} = 238{,}75$$

$$\hat{s}^2 = 238{,}75 - (14{,}25)^2 = 35{,}6875$$

$$s^2 = \frac{100.35{,}6875}{99} = 36{,}0480$$

3- Phép đổi biến số

Với xo tùy ý và $h \neq 0$, đặt:

$$u_i = \frac{x_i - x_0}{h}, \quad i = \overline{1,k}$$

Vì xi = hui + xo, nên theo tính chất của kỳ vọng và phương sai ta có:

$$\overline{x} = \overline{hu + x_o} = h\overline{u} + x_o$$

$$\widehat{s}_x^2 = h^2 \widehat{s}_u^2 = h^2(\overline{u^2} - \overline{u}^2)$$

Nếu chọn được xo và h để tính $\overline{u}$ và s_u^2 một cách đơn giản thì ta sẽ có một cách đơn giản để tính $\overline{x}$ và $\widehat{s}_x^2$.

Trường hợp các giá trị xi cách đều nhau thì ta chọn h bằng khoảng cách đó và chọn xo = xi với:

ni = max(n1, n2, ..., nk)

Ví dụ 5.2. Đường kính của 100 chi tiết do một máy sản xuất có số liệu:

Đường kính (mm)	**Số chi tiết**
19,80 ÷ 19,85	3
19,85 ÷ 19,05	5
19,90 ÷ 19,95	16
19,95 ÷ 20,00	28
20,00 ÷ 20,05	23
20,05 ÷ 20,10	14
20,10 ÷ 20,15	7
20,15 ÷ 20,20	4

Hãy tính các tham số mẫu.

Giải. Chọn:

$$h = 0{,}05;\ x_o = \frac{19{,}90 + 19{,}95}{2} = 19{,}925$$

Ta có bảng tính sau đây:

x_i	n_i	$u_i = \frac{x_i - x_o}{h}$	$u_i n_i$	$u_i^2 u_i$
19,825	3	– 3	– 9	27
19,875	5	– 2	– 10	20
19,925	16	– 1	– 16	16
19,975	28	0	0	0
20,025	23	1	23	23
20,075	14	2	28	56
21,125	7	3	21	63
21,175	4	4	16	64
Σ	100		53	269
TB			$\overline{u} = 0{,}53$	$\overline{u}^2 = 2{,}69$

Từ bảng tính ta có:

$$\overline{x} = h\overline{u} + x_o = 0{,}05.0{,}53 + 19{,}975 = 20{,}0015$$

$$\hat{s}^2 = h^2(\overline{u}^2 - \overline{u}^2) = 0{,}05^2(2{,}69 - 0{,}53^2) = 0{,}00602$$

$$s^2 = \frac{n\hat{s}^2}{n-1} = 0{,}00608$$

5.3. TÍNH CHẤT CỦA ĐẶC TRƯNG MẪU

1. Kỳ vọng và phương sai của đặc trưng mẫu

1- Kỳ vọng và phương sai của tỷ lệ mẫu

Xét mẫu định tính kích thước n. Giả sử F là tỷ lệ mẫu tổng quát, đặt:

$$x_i = \begin{cases} 1 & \text{nếu phần tử thứ i có tính chất A} \\ 0 & \text{nếu trái lại} \end{cases}$$

khi đó: x1, x2, ..., xn - là các đại lượng ngẫu nhiên và:

$$F = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Định lý 5.1. Nếu tổng thể có tỷ lệ p thì:

$$E(F) = p \text{ và } D(F) = \frac{pq}{n}$$

Chứng minh. Phân phối của Xi là:

X_i	0	1
P	q	p

do đó: $E(X_i) = p; \quad D(X_i) = pq.$

Vì các Xi độc lập nên theo tính chất của kỳ vọng và phương sai:

$$E(F)=\frac{1}{n}E\left(\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{np}{n}=p$$

$$D(F)=\frac{1}{n^2}D\left(\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{npq}{n^2}=\frac{pq}{n}$$

2- Kỳ vọng và phương sai của trung bình mẫu

Xét mẫu định lượng tổng quát kích thước n:

X = (X1, X2, ..., Xn)

Khi đó X1, X2, ..., Xn là các đại lượng ngẫu nhiên và trung bình mẫu tổng quát là:

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

Định lý 5.2. Nếu tổng thể có kỳ vọng a, phương sai σ2, thì mọi mẫu kích thước n đều có:

$$E(\overline{X})=a,\quad D(\overline{X})=\frac{\sigma^2}{n}$$

Chứng minh. Vì các Xi độc lập và E(Xi) = a, D(Xi) = σ2 nên theo tính chất của kỳ vọng và phương sai:

$$E(\overline{X}) = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) = \frac{na}{n} = a$$

$$D(\overline{X}) = \frac{1}{n^2}D\left(\sum_{i=1}^{n} X_i\right) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

3- Kỳ vọng của phương sai mẫu

Tiếp tục xét mẫu tổng quát:

X = (X1, X2, ..., Xn)

Ta có phương sai mẫu tổng quát là:

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

và phương sai mẫu hiệu chỉnh tổng quát là:

$$S^2 = \frac{n\widehat{S}^2}{n-1}$$

Định lý 5.3. Nếu tổng thể có kỳ vọng a, phương sai σ2 thì:

$$E(\widehat{S}^2) = \frac{n-1}{n}\sigma^2, \quad E(S^2) = \sigma^2$$

Chứng minh. Ta có:

$$E(\widehat{S}^2) = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} X_i\right)^2\right]$$

$$= \frac{n-1}{n^2} E\left(\sum_{i=1}^{n} X_I^2\right) - \frac{2}{n^2} E\left(\sum_{i<j} X_i X_j\right)$$

Với i ≠ j, thì:

E(XiXj) = E(Xi)E(Xj) = a2

nên:

$$E(\widehat{S}^2) = \frac{n-1}{n^2} E\left(\sum_{i=1}^{n} X_i^2\right) - \frac{n(n-1)}{n^2} a^2$$

$$= \frac{n-1}{n^2}\sum_{i=1}^{n}\left[E(X_i^2) - a^2\right]$$

$$= \frac{n-1}{n^2}\sum_{i=1}^{n} D(X_i) = \frac{n-1}{n^2}.n\sigma^2 = \frac{n-1}{n}\sigma^2$$

$$E(S^2) = E\left(\frac{n}{n-1}\widehat{S}^2\right) = \frac{n}{n-1}E(\widehat{S}^2) = \frac{n}{n-1}.\frac{n-1}{n}\sigma^2 = \sigma^2$$

2. Phân phối xác suất của các đặc trưng mẫu

1- Phân phối xác suất của tỷ lệ mẫu

Vì $E(F) = p$ và $D(F) = \frac{pq}{n}$ nên theo định lý 4.5 chương 4, với $n \geq 30$ ta có thể coi:

$$F \sim N(p), \frac{pq}{n})$$

Với một mẫu cụ thể kích thước n, tỷ lệ mẫu f, ta có $p \approx f$, nên:

$$F \sim N(p, \frac{f(1-f)}{n})$$

hay: $$\frac{(F-p)\sqrt{n}}{f(1-f)} \sim N(0,1)$$

2- Phân phối xác suất của trung bình mẫu

Vì $D(\overline{X}) = a,\ D(\overline{X}) = \frac{\sigma^2}{n}$ nên nếu tổng thể có phân phối chuẩn thì:

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

hay: $$\frac{\overline{X}-a}{\sigma}\sqrt{n} \sim N(0,1)$$

Nếu $n \geq 30$ thì với một mẫu cụ thể kích thước n ta có:

$$\sigma^2 \approx s^2$$

do đó: $$\overline{X} \sim N(a, \frac{s^2}{n})$$

hay: $$\frac{\overline{X}-a}{s}\sqrt{n} \sim N(0,1)$$

trong đó: s2 - là phương sai mẫu hiệu chỉnh của một mẫu kích thước n bất kỳ.

Trường hợp n < 30, tổng thể có phân phối chuẩn, ta có:

$$\frac{\overline{X} - a}{s}\sqrt{n} \sim T(n-1)$$

3- Phân phối xác suất của phương sai mẫu

Nếu tổng thể có phân phối chuẩn thì ta có:

$$\frac{n\widehat{S}^2}{\sigma^2} = \frac{n-1}{\sigma^2}S^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2 \sim X^2(n-1)$$

5.4. ĐA GIÁC ĐỒ. TỔ CHỨC ĐỒ

Cho mẫu:

x_i	x_1	x_2	...	x_k
n_i	n_1	n_2	...	n_k

$$x_1 < x_2 < \ldots < x_k$$

1- Hàm phân phối mẫu

Đặt: $\mathbf{n_x} = \sum_{x_i < x} \mathbf{n_i}$

ta được hàm: $\mathbf{F(x)} = \dfrac{\mathbf{n_x}}{\mathbf{n}}$

xác định trên toàn trục số, gọi là hàm phân phối mẫu hay phân phối thực nghiệm.

Hàm phân phối mẫu là xấp xỉ của hàm phân phối của tổng thể.

Ví dụ 5.3. Một mẫu về X có số liệu:

x_i	1	2	3	4
n_i	10	15	20	5

Tìm hàm phân phối mẫu của X.

Giải. Ta có n = 50:

$$n_x = \begin{cases} 0 & \text{nếu} \quad x \le 1 \\ 10 & \text{nếu} \quad 1 < x \le 2 \\ 25 & \text{nếu} \quad 2 < x \le 3 \\ 45 & \text{nếu} \quad 3 < x \le 4 \\ 50 & \text{nếu} \quad x > 4 \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{nếu} \quad x \le 1 \\ 0,2 & \text{nếu} \quad 1 < x \le 2 \\ 0,5 & \text{nếu} \quad 2 < x \le 3 \\ 0,9 & \text{nếu} \quad 3 < x \le 4 \\ 1 & \text{nếu} \quad x > 4 \end{cases}$$

2- Đa giác đồ

Biểu diễn các điểm (xi, ni), $i = \overline{1, k}$, lên mặt phẳng tọa độ và nối các điểm (xi, ni) và (xi+1, ni+1), $i = \overline{1, k-1}$ bằng một đoạn thẳng, ta được một đường gấp khúc gọi là đa giác tần số hay đa giác đồ.

Ví dụ 5.4. Điểm môn toán của 20 học sinh có số liệu

x_i (điểm)	4	5	6	7	8
n_i (số học sinh)	2	5	4	6	3

Hãy lập đa giác tần số.

Giải. Ta có đa giác tần số như trong hình vẽ.

Nếu biểu diễn các điểm:

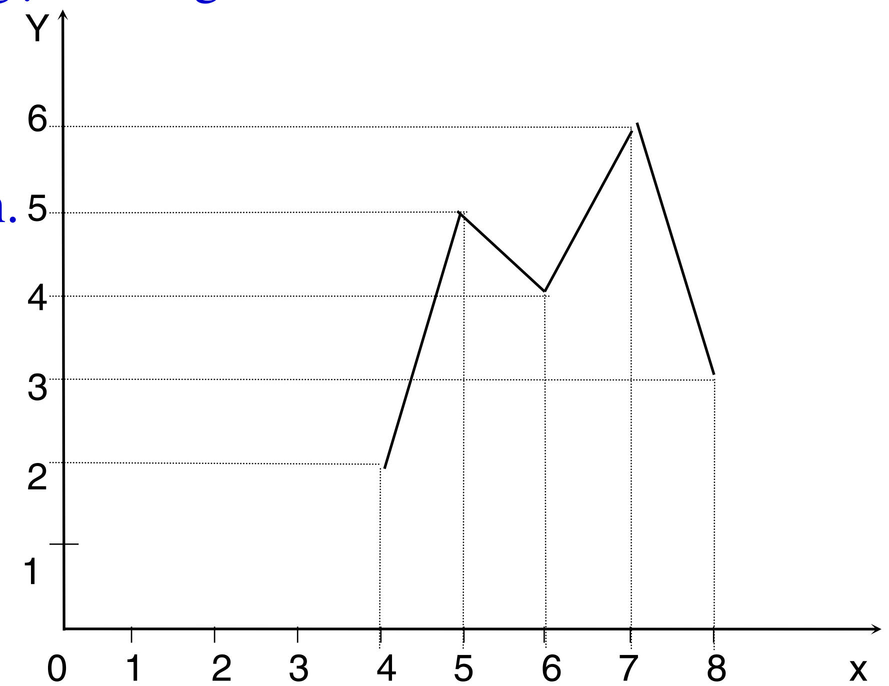
$$(xi, \frac{n_i}{n}), \quad i = \overline{1, k}$$

và nối lại ta được một đường khúc gọi là đa giác tần suất.

3- Tổ chức đồ

Chia đoạn [x1, xk] thành các khoảng bằng nhau có độ dài bằng h. Trên mỗi khoảng Jx ta tính tổng:

$$\gamma x = \sum_{x_i \in J_x} n_i$$



Dựng các hình chữ nhật đáy Jx, chiều cao $\frac{\gamma_x}{h}$. Hình vẽ nhận được gọi là biểu đồ tần số, hay tổ chức đồ của mẫu đã cho. Tổng diện tích của các hình chữ nhật bằng kích thước của mẫu.

Ví dụ 5.5. Điểm môn toán của 36 sinh viên có số liệu:

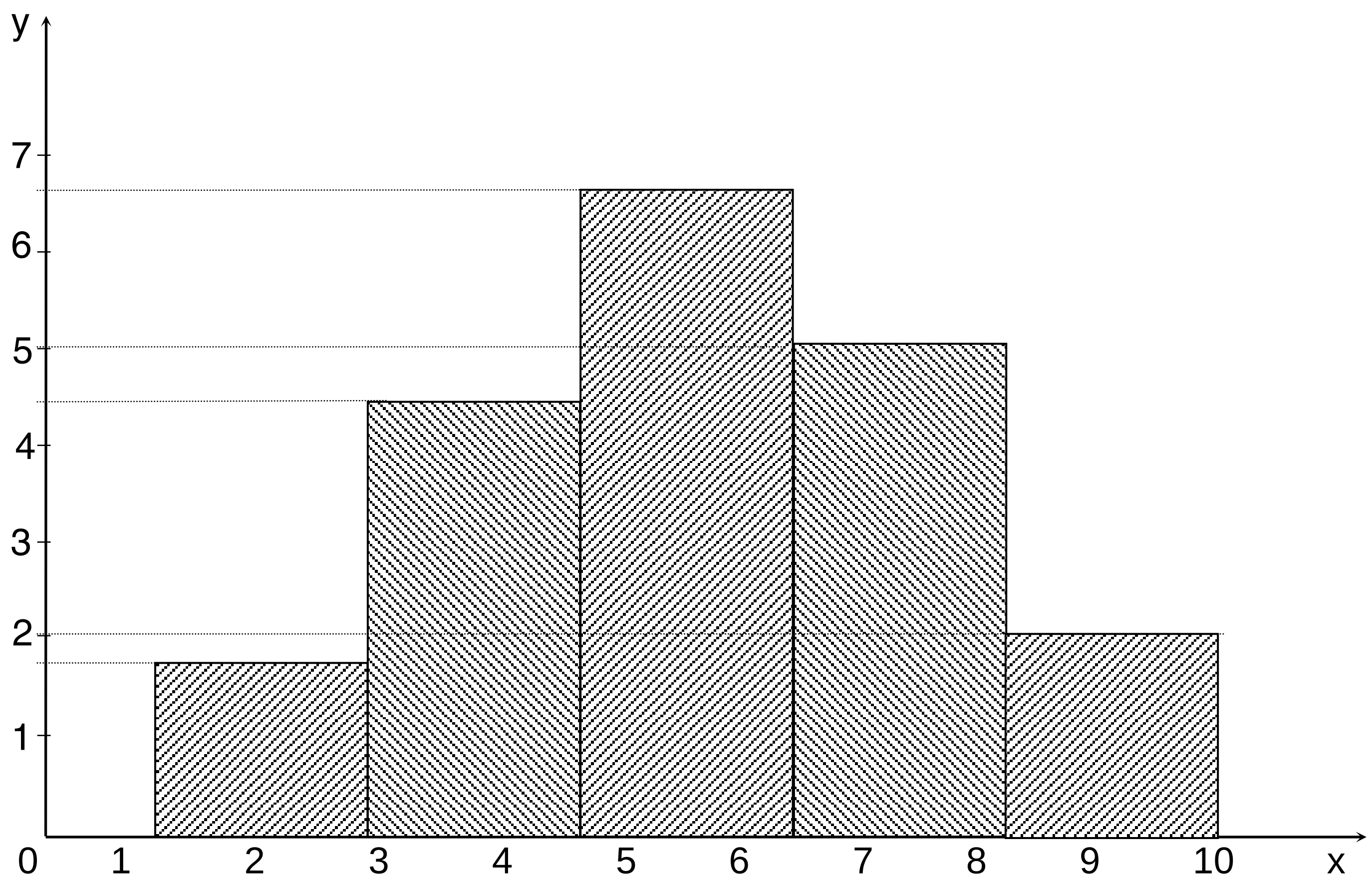
x_i (điểm)	1	2	3	4	5	6	7	8	9	10
n_i (số sinh viên)	1	2	4	4	5	7	5	4	3	1

Hãy lập đa giác tần số.

Giải. Ta chia đoạn [1,10] thành năm khoảng bằng nhau, h = 1,8.

Khoảng	1 ; 2,8	2,8 ; 4,6	4,6 ; 6,4	6,4 ; 8,2	8,2 ; 10
γ_x	3	8	12	9	4
$\dfrac{\gamma_x}{h}$	1,67	4,44	6,67	5,00	2,22

Ta có đa giác tần số như sau:



BÀI TẬP

5.1. Gặt ngẫu nhiên 360 thửa ruộng ở một tỉnh thu được bảng số liệu:

Năng suất x_i (tạ/ha)	125	130	133	134	135	136	137	138	139	140
Số thửa n_i	6	12	34	74	106	85	30	5	5	3

Hãy tính $\overline{x}, \widehat{s}^2, s^2, \widehat{s}, s$.

5.2. Lấy ngẫu nhiên chiều cao của 120 thanh niên ở một thành phố đem đo chiều cao, thu được bảng số liệu sau:

Chiều cao (cm)	154 -158	158 -162	162 -166	166 -170	170 -174	174 -178	178 -182
Số thanh niên (n_i)	10	14	26	38	22	8	2

Hãy tính các đặc trưng $\overline{x}, \widehat{s}^2, s^2, \widehat{s}, s$.

5.3. Cho 10 kết quả đo đạc một đại lượng X bởi cùng một máy không có sai số hệ thống: 369; 378; 315; 320; 325; 420; 385; 401; 372; 383. Hãy Tính $\overline{x}, s^2, s$.

5.4. Khi đo độ dài của 36 chi tiết chọn ngẫu nhiên của một loại sản phẩm có bảng số liệu:

x_i	39	40	41	42	43	44
n_i	4	5	10	12	4	1

a) Tìm hàm phân bố mẫu

b) Tính $\overline{x}$, $\widehat{s}^2$, s^2.

5.5. Kết quả về việc đo độ bền các sợi chỉ ta thu được bảng số liệu sau đây:

Độ bền x_i	120 -140	140 -160	160 -180	180 -200	200 -220	220 -240	240 -260	260 -280
Số các sợi	5	10	20	25	15	10	10	5

Hãy xác định độ bền trung bình, phương sai và phương sai hiệu chỉnh của mẫu trên.

5.6. Để nghiên cứu tuổi thọ trung bình (X) của một loại bóng đèn, người ta lấy một mẫu có số liệu:

Tuổi thọ x_i (giờ)	2900 -2920	2920 -2940	2940 -2960	2960 -2980	2980 -3000	3000 -3020	3020 -3040
Số bóng đèn (n_i)	10	15	25	30	20	15	5

Hãy tính tuổi thọ trung bình, phương sai và phương sai hiệu chỉnh của mẫu trên.

5.7. Điều tra doanh số hàng tháng của 100 hộ kinh doanh một loại hàng hóa, ta có bảng số liệu:

Doanh số (triệu đồng/tháng)	101	102	104	105	107	108	109	110	113	114
Số hộ (n_i)	2	3	8	13	25	20	12	10	6	1

a) Hãy tìm hàm phân bố mẫu, biểu đồ tần suất của bảng trên

b) Tìm doanh số trung bình và độ lệch chuẩn hiệu chỉnh của mẫu trên.

5.8. Tại một trại chăn nuôi heo, người ta áp dụng thử một loại thuốc tăng trọng bổ sung vào khẩu phần thức ăn. Sau khi nuôi ba tháng đạt được kết quả sau:

a)

x_i (kg)	75	77	78	79	80	81	83
n_i (số heo)	1	3	4	7	6	2	2

b) Lô đối chứng (không dùng thuốc tăng trọng) có kết quả:

x_i (kg)	60	61	62	63	65	67	69	70	72
n_i (số heo)	1	2	4	5	6	3	2	1	1

Hãy tìm trọng lượng trung bình, độ phân tán của trọng lượng của hai lô heo trên.

5.9. Bảng dưới đây chỉ kết quả thu hoạch Y (tạ/ha) và lượng phân bón X (kg/ha) của một loại hoa màu tại 100 thửa ruộng gieo loại hoa màu đó:

Y / X	14	15	16	17	18	n_i
1	10					10
2	8	12				20
3		7	28			35
4			6	8		14
5				9	12	21
m_j	18	19	34	17	22	n = 100

Hãy tính $\overline{x}$, $\overline{y}$, $\widehat{s}_X$, $\widehat{s}_Y$, $\overline{xy} = \dfrac{\Sigma x_i y_j n_{ij}}{n}$