

Bài 1. Khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu ghép nhóm

A. Kiến thức cần nhớ

1. Khoảng biến thiên

a) Định nghĩa

- Khoảng biến thiên, kí hiệu  $R$ , của mẫu số liệu ghép nhóm là hiệu số giữa đầu mút phải của nhóm cuối cùng và đầu mút trái của nhóm đầu tiên có chứa dữ liệu của mẫu số liệu.

**Chẳng hạn:** Xét mẫu số liệu ghép nhóm được cho ở bảng sau: (Bảng 1)

Nhóm	$[u_1; u_2)$	$[u_2; u_3)$	...	$[u_k; u_{k+1})$
Tần số	$n_1$	$n_2$	...	$n_k$

- Nếu  $n_1$  và  $n_k$  cùng lớn hơn 0 thì  $R = u_{k+1} - u_1$ .
- Khoảng biến thiên của mẫu số liệu ghép nhóm luôn lớn hơn hoặc bằng khoảng biến thiên của mẫu số liệu gốc.

b) Ý nghĩa của khoảng biến thiên của mẫu số liệu ghép nhóm

- Khoảng biến thiên của mẫu số liệu ghép nhóm là giá trị xấp xỉ khoảng biến thiên của mẫu số liệu gốc và có thể dùng để đo mức độ phân tán của mẫu số liệu. Khoảng biến thiên càng lớn thì mẫu số liệu càng phân tán.
- Khoảng biến thiên  $R = u_{k+1} - u_1$  chưa phản ánh được đầy đủ mức độ phân tán của phần lớn các số liệu. Hơn nữa, giá trị của  $R$  thường tăng vọt khi xuất hiện giá trị bất thường (còn gọi là giá trị ngoại lệ) trong mẫu số liệu. Do đó, để phản ánh mức độ phân tán của số liệu, người ta còn dùng các số đặc trưng khác.

2. Khoảng tứ phân vị

a) Định nghĩa

Xét mẫu số liệu ghép nhóm được cho ở bảng sau: (Bảng 1)

Nhóm	$[u_1; u_2)$	$[u_2; u_3)$	...	$[u_k; u_{k+1})$
Tần số	$n_1$	$n_2$	...	$n_k$

- Tứ phân vị thứ  $k$ , kí hiệu là  $Q_k$ , với  $k = 1, 2, 3$  của mẫu số liệu ghép nhóm (Bảng 1) được xác định

như sau:  $Q_k = u_m + \frac{4}{n_m} \left( \frac{kn}{4} - C \right) (u_{m+1} - u_m)$ , trong đó:

- $n = n_1 + n_2 + \dots + n_k$  là cỡ mẫu;
- $n_m$  là tần số của nhóm chứa tứ phân vị thứ  $k$ ;
- $[u_m; u_{m+1})$  là nhóm chứa tứ phân vị thứ  $k$ ;
- $C = n_1 + n_2 + \dots + n_{m-1}$ .

- Nếu tứ phân vị thứ  $k$  là  $\frac{1}{2}(x_m + x_{m+1})$ , trong đó  $x_m$  và  $x_{m+1}$  thuộc hai nhóm liên tiếp, ví dụ như

$x_m \in [u_{j-1}; u_j)$  và  $x_{m+1} \in [u_j; u_{j+1})$  thì ta lấy  $Q_k = u_j$ .

- Khoảng tứ phân vị của mẫu số liệu ghép nhóm cũng được xác định dựa trên tứ phân vị thứ nhất và tứ phân vị thứ ba như đối với mẫu số liệu không ghép nhóm.
- Khoảng tứ phân vị của mẫu số liệu ghép nhóm cho ở Bảng 1, kí hiệu  $\Delta_Q$ , là hiệu giữa tứ phân vị thứ ba  $Q_3$  và tứ phân vị thứ nhất  $Q_1$  của mẫu số liệu ghép nhóm đó, tức là

$$\Delta_Q = Q_3 - Q_1.$$

## b) Ý nghĩa của khoảng tứ phân vị của mẫu số liệu ghép nhóm

- Khoảng tứ phân vị của mẫu số liệu ghép nhóm là giá trị xấp xỉ cho khoảng tứ phân vị của mẫu số liệu gốc và có thể dùng để đo mức độ phân tán của nửa giữa của mẫu số liệu (tập hợp gồm 50% số liệu nằm chính giữa mẫu số liệu).
- Khoảng tứ phân vị của mẫu số liệu ghép nhóm càng nhỏ thì dữ liệu càng tập trung xung quanh trung vị.
- Khoảng tứ phân vị được dùng để xác định giá trị bất thường (giá trị ngoại lệ) trong mẫu số liệu. Giá trị  $x$  trong mẫu số liệu là giá trị ngoại lệ nếu  $x > Q_3 + 1,5\Delta_Q$  hoặc  $x < Q_1 - 1,5\Delta_Q$ .
- Khoảng tứ phân vị của mẫu số liệu ghép nhóm không bị ảnh hưởng nhiều bởi các giá trị bất thường (giá trị ngoại lệ) trong mẫu số liệu.

## 3. Một số kiến thức bổ trợ quan trọng

### ➤ Một số quy tắc ghép nhóm của mẫu số liệu

- Mỗi mẫu số liệu có thể được ghép nhóm theo nhiều cách khác nhau nhưng thường tuân theo một số quy tắc sau:
- Sử dụng từ  $k = 5$  đến  $k = 20$  nhóm. Cỡ mẫu càng lớn thì cần càng nhiều nhóm số liệu.
- Các nhóm có cùng độ dài bằng  $L$  thoả mãn  $R < k.L$ , trong đó  $R$  là khoảng biến thiên,  $k$  là số nhóm.
- Giá trị nhỏ nhất của mẫu số thuộc vào nhóm  $[u_1; u_2)$  và càng gần  $u_1$  càng tốt. Giá trị lớn nhất của mẫu thuộc nhóm  $[u_k; u_{k+1})$  và càng gần  $u_{k+1}$  càng tốt.

### ➤ Xét mẫu số liệu ghép nhóm được cho ở bảng sau: (Bảng 1)

Nhóm	$[u_1; u_2)$	$[u_2; u_3)$	...	$[u_k; u_{k+1})$
Tần số	$n_1$	$n_2$	...	$n_k$

- Giá trị chính giữa của mỗi nhóm được dùng làm **giá trị đại diện** cho nhóm ấy. Ví dụ nhóm  $[u_1; u_2)$  có giá trị đại diện là  $\frac{1}{2}(u_1 + u_2)$ .
- Hiệu  $u_{j+1} - u_j$  được gọi là **độ dài của nhóm**  $[u_j; u_{j+1})$ .
- **Số trung bình** của mẫu số liệu ghép nhóm:  

$$\bar{x} = \frac{n_1 c_1 + n_2 c_2 + \dots + n_k c_k}{n}$$
, trong đó  $n = n_1 + n_2 + \dots + n_k$  và  $c_i$  là giá trị đại diện của từng nhóm số liệu.

### ➤ Một

- **Nhóm chứa một** của mẫu số liệu ghép nhóm là nhóm có tần số lớn nhất.
- Giả sử nhóm chứa một là  $[u_m; u_{m+1})$ , khi đó **mốt của mẫu số liệu ghép nhóm**, kí hiệu là  $M_o$ , được xác định bởi công thức

$$M_o = u_m + \frac{n_m - n_{m-1}}{(n_m - n_{m-1}) + (n_m - n_{m+1})} (u_{m+1} - u_m)$$

- **Chú ý:** Nếu không có nhóm kế trước của nhóm chứa một thì  $n_{m-1} = 0$ . Nếu không có nhóm kế sau của nhóm chứa một thì  $n_{m+1} = 0$ .

### ➤ Trung vị

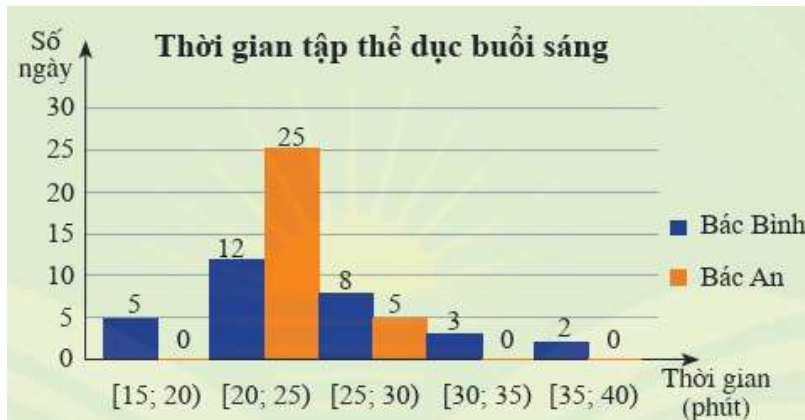
- Giả sử nhóm  $[u_m; u_{m+1})$  chứa trung vị;
- $n_m$  là tần số của nhóm chứa trung vị;
- $C = n_1 + n_2 + \dots + n_{m-1}$ .

$$\text{Khi đó: } M_e = u_m + \frac{\frac{n}{2} - C}{n_m} \cdot (u_{m+1} - u_m).$$

## B. Các dạng bài tập & phương pháp giải

### Dạng 1. Khoảng biến thiên

**Ví dụ 1.** Biểu đồ dưới đây thống kê thời gian tập thể dục buổi sáng mỗi ngày trong tháng 9/2022 của bác Bình và bác An



a) Em hãy chọn số thích hợp thay vào các vị trí được đánh dấu ? ở bảng sau:

Thời gian (phút)	[15; 20)	[20; 25)	[25; 30)	[30; 35)	[35; 40)
Bác Bình	?	12	8	3	2
Bác An	?	?	?	?	?

b) Hãy tìm khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng mỗi ngày của bác Bình và bác An.

### Giải

a) Ta có bảng sau:

Thời gian (phút)	[15; 20)	[20; 25)	[25; 30)	[30; 35)	[35; 40)
Bác Bình	5	12	8	3	2
Bác An	0	25	5	0	0

b) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bác Bình là  $40 - 15 = 25$  (phút).

Tuy nhiên, trong mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bác An, khoảng đầu tiên chứa dữ liệu là [20; 25) và khoảng cuối cùng chứa dữ liệu là [25; 30).

Do đó khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bác An là  $30 - 20 = 10$  (phút).

**Ví dụ 2.** Bảng sau thống kê cân nặng của 50 quả xoài Thanh Ca được lựa chọn ngẫu nhiên sau khi thu hoạch ở một nông trường.

Cân nặng (g)	[250; 290)	[290; 330)	[330; 370)	[370; 410)	[410; 450)
Số quả xoài	3	13	18	11	5

a) Hãy tìm khoảng biến thiên của mẫu số liệu ghép nhóm trên.

b) Có ý kiến cho rằng: “Trong 50 quả xoài trên, hiệu số cân nặng của hai quả bất kì không vượt quá 200 g”. Ý kiến đó đúng hay sai? Giải thích.

**Giải:** Khoảng biến thiên của mẫu số liệu ghép nhóm trên là:  $450 - 250 = 200$  (g).

**Ví dụ 3.** Cô Hà thống kê lại đường kính thân gỗ của một số cây xoan đào 6 năm tuổi được trồng ở một lâm trường ở bảng sau.

Đường kính (cm)	[40; 45)	[45; 50)	[50; 55)	[55; 60)	[60; 65)
Tần số	5	20	18	7	3

Hãy tìm khoảng biến thiên của mẫu số liệu ghép nhóm trên.

**Giải:** Khoảng biến thiên của mẫu số liệu ghép nhóm trên là:  $65 - 40 = 25$  (cm).

**Ví dụ 4.** Bạn Trang thống kê lại chiều cao (đơn vị: cm) của các bạn học sinh nữ lớp 12C và lớp 12D ở bảng sau:

Chiều cao (cm)	[155; 160)	[160; 165)	[165; 170)	[170; 175)	[175; 180)	[180; 185)
	Lớp					
12C	2	7	12	3	0	1
12D	5	9	8	2	1	0

Nếu so sánh theo khoảng biến thiên thì chiều cao của học sinh lớp nào có độ phân tán lớn hơn?

### Dạng 2. Khoảng tứ phân vị

**Ví dụ 5.** Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm trong bảng dưới đây.

Cân nặng (g)	[250; 290)	[290; 330)	[330; 370)	[370; 410)	[410; 450)
Số quả xoài	3	13	18	11	5

**Giải**

Cỡ mẫu  $n = 50$ .

Gọi  $x_1; x_2; \dots; x_{50}$  là mẫu số liệu gốc gồm cân nặng của 50 quả xoài được xếp theo thứ tự không giảm.

Ta có:  $x_1, x_2, x_3 \in [250; 290)$ ;  $x_4, \dots, x_{16} \in [290; 330)$ ;  $x_{17}, \dots, x_{34} \in [330; 370)$ ;  $x_{35}, \dots, x_{45} \in [370; 410)$ ;  $x_{46}, \dots, x_{50} \in [410; 450)$ .

Tứ phân vị thứ nhất của mẫu số liệu gốc là  $x_{13} \in [290; 330)$ . Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép

$$\text{nhóm là: } Q_1 = 290 + \frac{\frac{50}{4} - 3}{13} (330 - 290) = \frac{4150}{13}$$

Tứ phân vị thứ ba của mẫu số liệu gốc là  $x_{38} \in [370; 410)$ . Do đó, tứ phân vị thứ ba của mẫu số liệu ghép

$$\text{nhóm là: } Q_3 = 370 + \frac{\frac{3 \cdot 50}{4} - (3 + 13 + 18)}{11} (410 - 370) = \frac{4210}{11}$$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là:  $\Delta_Q = \frac{4210}{11} - \frac{4150}{13} = \frac{9080}{143}$

**Ví dụ 6.** Hằng ngày ông Thắng đều đi xe buýt từ nhà đến cơ quan. Dưới đây là bảng thống kê thời gian của 100 lần ông Thắng đi xe buýt từ nhà đến cơ quan.

Thời gian (phút)	[15; 18)	[18; 21)	[21; 24)	[24; 27)	[27; 30)	[30; 33)
Số lượt	22	38	27	8	4	1

- a) Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.
- b) Biết rằng trong 100 lần đi trên, chỉ có đúng một lần ông Thắng đi hết hơn 29 phút. Thời gian của lần đi đó có phải là giá trị ngoại lệ không?
- c) Hãy tìm khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu ghép nhóm ở trên sau khi đã loại bỏ các giá trị ngoại lệ. Em có nhận xét gì về khoảng biến thiên, khoảng tứ phân vị vừa tìm được và khoảng biến thiên, khoảng tứ phân vị ban đầu?

### Giải

Cỡ mẫu  $n = 100$ .

Gọi  $x_1; x_2; \dots; x_{100}$  là mẫu số liệu gốc gồm thời gian 100 lần đi xe buýt của ông Thắng.

Ta có:  $x_1, \dots, x_{22} \in [15; 18)$ ;  $x_{23}, \dots, x_{60} \in [18; 21)$ ;  $x_{61}, \dots, x_{87} \in [21; 24)$ ;  $x_{88}, \dots, x_{95} \in [24; 27)$ ;  $x_{96}, \dots, x_{99} \in [27; 30)$ ;  $x_{100} \in [30; 33)$ .

Tứ phân vị thứ nhất của mẫu số liệu gốc là  $\frac{1}{2}(x_{25} + x_{26}) \in [18; 21)$ . Do đó, tứ phân vị thứ nhất của mẫu số

$$\text{liệu ghép nhóm là: } Q_1 = 18 + \frac{\frac{100}{4} - 22}{38} (21 - 18) = \frac{693}{38}$$

Tứ phân vị thứ ba của mẫu số liệu gốc là  $\frac{1}{2}(x_{75} + x_{76}) \in [21; 24)$ . Do đó, tứ phân vị thứ ba của mẫu số liệu

$$\text{ghép nhóm là: } Q_3 = 21 + \frac{\frac{3 \cdot 100}{4} - (22 + 38)}{27} (24 - 21) = \frac{68}{3}$$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là:  $\Delta_Q = \frac{68}{3} - \frac{693}{38} = \frac{505}{114}$

Trong lần duy nhất ông Thắng đi hết hơn 29 phút, thời gian đi của ông thuộc nhóm [30; 33).

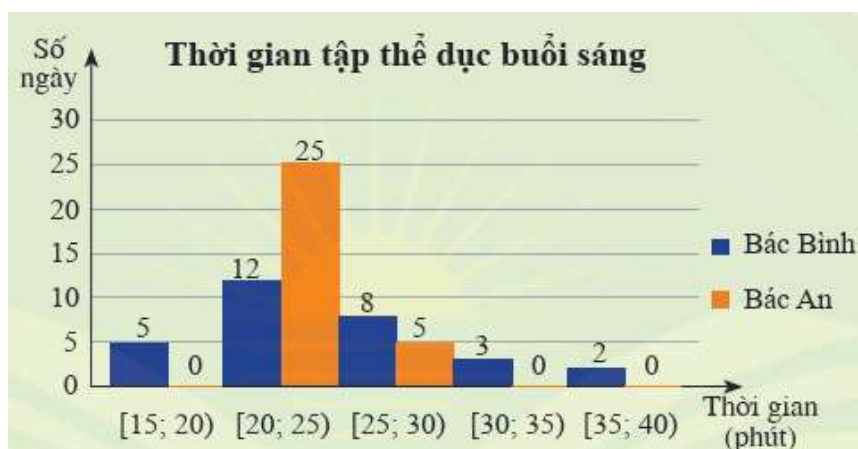
Vì  $Q_3 + 1,5\Delta_Q = \frac{6683}{228} < 30$  nên thời gian của lần ông Thắng đi hết hơn 29 phút là giá trị ngoại lệ của mẫu số liệu ghép nhóm.

**Ví dụ 7.** Kết quả điều tra tổng thu nhập trong năm 2022 của một số hộ gia đình trong một địa phương được ghi lại ở bảng sau:

Tổng thu nhập (triệu đồng)	[200; 250)	[250; 300)	[300; 350)	[350; 400)	[400; 450)
Số hộ gia đình	24	62	34	21	9

- a) Hãy tìm các tứ phân vị  $Q_1$  và  $Q_3$ .
- b) Một doanh nghiệp địa phương muốn hướng dịch vụ của mình đến các gia đình có mức thu nhập ở tầm trung, tức là 50% các hộ gia đình có mức thu nhập ở chính giữa so với tất cả các hộ gia đình của địa phương. Hỏi doanh nghiệp cần hướng đến các gia đình có mức thu nhập trong khoảng nào?

**Ví dụ 8.** Hãy so sánh khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng mỗi ngày của bác Bình và bác An mỗi ngày trong tháng 9/2022 được thống kê trong biểu đồ dưới đây:



**Ví dụ 9.** Giả sử kết quả khảo sát hai khu vực  $A$  và  $B$  về độ tuổi kết hôn của một số phụ nữ vừa lập gia đình được cho ở bảng sau:

Tuổi kết hôn	[19; 22)	[22; 25)	[25; 28)	[28; 31)	[31; 34)
Số phụ nữ khu vực $A$	10	27	31	25	7
Số phụ nữ khu vực $B$	47	40	11	2	0

- Hãy tìm khoảng biến thiên và khoảng tứ phân vị của từng mẫu số liệu ghép nhóm ứng với mỗi khu vực  $A$  và  $B$ .
- Nếu so sánh theo khoảng tứ phân vị thì phụ nữ ở khu vực nào có độ tuổi kết hôn đồng đều hơn?



## C. Bài tập tự luận rèn luyện

**Bài 1.** Bảng 3.3 biểu thị kết quả điều tra thời gian sử dụng Internet hằng ngày của một số người.

Bảng 3.3. Thời gian sử dụng Internet hằng ngày

Thời gian (phút)	[30; 60)	[60; 90)	[90; 120)	[120; 150)	[150; 180)
Số người	2	4	10	5	3

Tìm khoảng biến thiên của mẫu số liệu đã cho. Kết quả cho biết điều gì?

### Giải

Đầu mút phải của nhóm ghép cuối cùng là 180, đầu mút trái của nhóm ghép đầu tiên là 30. Vậy khoảng biến thiên của mẫu số liệu là  $R = 180 - 30 = 150$ .

Kết quả này cho biết thời gian sử dụng Internet hằng ngày của các thành viên thuộc nhóm người được điều tra chênh lệch nhau nhiều nhất là 150 phút.

**Bài 2.** Bảng 3.4 thống kê thành tích nhảy xa của một số học sinh lớp 12. Tìm khoảng biến thiên thành tích nhảy xa của số học sinh này.

Bảng 3.4. Thành tích nhảy xa của một số học sinh lớp 12

Thành tích (cm)	[150; 180)	[180; 210)	[210; 240)	[240; 270)	[270; 300)
Số học sinh	3	5	28	14	8

**Bài 3.** Để chuẩn bị mở một trung tâm thể dục thể thao, anh Dũng đã tiến hành điều tra tuổi thọ của máy chạy bộ do hai hãng X, Y sản xuất. Bảng 3.5 biểu thị hai mẫu số liệu mà anh thu thập được qua Internet.

Bảng 3.5. Tuổi thọ của máy chạy bộ (đơn vị: năm)

Tuổi thọ	[2; 4)	[4; 6)	[6; 8)	[8; 10)	[10; 12)
Số máy của hãng X	7	20	36	20	17
Số máy của hãng Y	0	20	35	35	10

Khoảng biến thiên của mẫu số liệu nào lớn hơn? Từ đó có thể nói là máy chạy bộ do hãng nào sản xuất có tuổi thọ phân tán hơn?

### Giải

Khoảng biến thiên của tuổi thọ máy chạy bộ do hãng X và hãng Y sản xuất tương ứng là  $R_X = 12 - 2 = 10$  và  $R_Y = 12 - 4 = 8$ . Vì  $R_X > R_Y$  nên có thể nói là máy do hãng X sản xuất có tuổi thọ phân tán hơn so với máy của hãng Y.

**Bài 4.** Cho mẫu số liệu ghép nhóm về chiều cao (đơn vị centimet) của 36 học sinh nam lớp 12 ở một trường THPT. Tìm khoảng biến thiên của mẫu số liệu ghép nhóm trên

Nhóm	Tần số
[160; 163)	6
[163; 166)	11
[166; 169)	9
[169; 172)	7
[172; 175)	3
	n = 36

**Bài 5.** Cho bảng tần số ghép nhóm biểu diễn mẫu số liệu ghi lại năng suất lúa (đơn vị: tạ/ha) của 60 địa phương. Tính khoảng biến thiên của mẫu số liệu trên

Nhóm	Tần số
[40; 47)	1
[47; 54)	6
[54; 61)	21
[61; 68)	21
[68; 75)	11
	n = 60

**Bài 6.** Cho bảng số liệu ghép nhóm như bảng bên. Tính khoảng biến thiên của mẫu số liệu đó

Nhóm	Tần số
[40; 45)	4
[45; 50)	11
[50; 55)	9
[55; 60)	8
[60; 65)	8
	n = 40

**Bài 7.** Thông kê thời gian sử dụng mạng xã hội trong ba ngày cả các bạn học sinh tổ 1, Tổ 2 lớp 12A được kết quả như bảng sau

Thời gian sử dụng (phút)	[0; 10)	[10; 30)	[30; 60)	[60; 90)
Số học sinh tổ 1	2	4	3	1
Số học sinh tổ 2	5	1	3	0

Tìm khoảng biến thiên thời gian sử dụng mạng xã hội của học sinh mỗi tổ và giải thích ý nghĩa.

**Bài 8.** Thời gian hoàn thành bài kiểm tra mô toán của các bạn học sinh trong lớp 12C được cho trong bảng sau.

Thời gian (phút)	[25; 30)	[30; 35)	[35; 40)	[40; 45)
Số học sinh	8	16	4	2

a) Tính khoảng biến thiên R của mẫu số liệu ghép nhóm trên

b) Nếu biết học sinh hoàn thành bài kiểm tra sớm nhất mất 27 phút và muộn nhất mất 43 phút thì khoảng biến thiên của mẫu số liệu gốc là bao nhiêu.

**Bài 9.** Người ta tiến hành phỏng vấn hai nhóm khán giả về một bộ phim mới công chiếu. Nhóm A gồm những khán giả thuộc lứa tuổi 20 – 30, nhóm B thuộc lứa tuổi trên 30. Người được hỏi ý kiến phải đánh giá bộ phim bằng cách cho điểm theo một số tiêu chí nêu trong phiếu điều tra và sau đó lấy tổng số điểm (thang điểm 100). Bảng dưới đây trình bày kết quả điều tra hai nhóm khán giả:

*Bảng 3.6. Điểm đánh giá của khán giả*

Điểm	[50; 60)	[60; 70)	[70; 80)	[80; 90)	[90; 100)
Số người của nhóm A	6	10	14	12	8
Số người của nhóm B	0	8	14	28	0

Ý kiến đánh giá của nhóm khán giả nào phân tán hơn?

**Bài 10.** Cho *Bảng 3.1* về khối lượng của 100 quả dưa giống E. Để tiện tính toán, ta biểu diễn dữ liệu bằng một bảng hai cột như bảng bên.

Khối lượng (gam)	Số quả
[700; 800)	6
[800; 900)	10
[900; 1 000)	14
[1 000; 1 100)	23
[1 100; 1 200)	15
[1 200; 1 300)	22
[1 300; 1 400)	6
[1 400; 1 500)	4

Hãy tính khoảng tứ phân vị của mẫu số liệu ghép nhóm đã cho.



**Bài 11.** Điểm kiểm tra cuối khoá môn Tiếng Anh của hai lớp ở một trung tâm ngoại ngữ được thống kê trong các Bảng 3.7a và 3.7b.

Bảng 3.7a. Điểm của lớp A

Điểm	Số học viên (tần số)
[50; 60)	8
[60; 70)	20
[70; 80)	50
[80; 90)	17
[90; 100)	5

Bảng 3.7b. Điểm của lớp B

Điểm	Số học viên (tần số)
[50; 60)	15
[60; 70)	20
[70; 80)	30
[80; 90)	20
[90; 100)	15

a) Tìm khoảng biến thiên của mỗi mẫu số liệu. Có thể dùng khoảng biến thiên để biết điểm của lớp nào đồng đều hơn không?

b) Tìm các tứ phân vị và khoảng tứ phân vị của mỗi mẫu số liệu. Mẫu số liệu nào có độ phân tán lớn hơn?

**Bài 12.** Ở một phòng điều trị nội trú của bệnh viện, dữ liệu thống kê thời gian ngủ hằng đêm của hai bệnh nhân trong suốt một tháng được tổng hợp bởi hai bảng dưới đây:

Bảng 3.8a. Thời gian ngủ của bệnh nhân A

Thời gian (phút)	Số đêm (tần số)
[180; 240)	5
[240; 300)	5
[300; 360)	10
[360; 420)	6
[420; 480)	4

Bảng 3.8b. Thời gian ngủ của bệnh nhân B

Thời gian (phút)	Số đêm (tần số)
[180; 240)	2
[240; 300)	9
[300; 360)	12
[360; 420)	5
[420; 480)	2

Bệnh nhân nào có thời gian ngủ ổn định hơn?

**Bài 13.** Thời gian chờ khám bệnh của các bệnh nhân tại phòng khám X được cho trong bảng sau:

Thời gian khám (phút)	[0; 5)	[5; 10)	[10; 15)	[15; 20)
Số bệnh nhân	3	12	15	8

a) Tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm này

b) Từ một mẫu số liệu về thời gian chờ khám bệnh của các bệnh nhân ở phòng khám Y người ta tính được khoảng tứ phân vị bằng 9,23. Hỏi thời gian chờ của bệnh nhân tại phòng khám nào phân tán hơn.

**Bài 14.** Thông kê số thẻ vàng của mỗi cầu lạc bộ trong giải ngoại hạng Anh mùa giải 2021 – 2022 cho kết quả như sau

101	79	79	78	75	73	68	67	67	63
63	61	60	59	57	55	55	50	47	42

a) Hãy ghép nhóm dãy số liệu trên thành các nhóm có độ dài bằng nhau với nhóm đầu tiên là [40; 50)

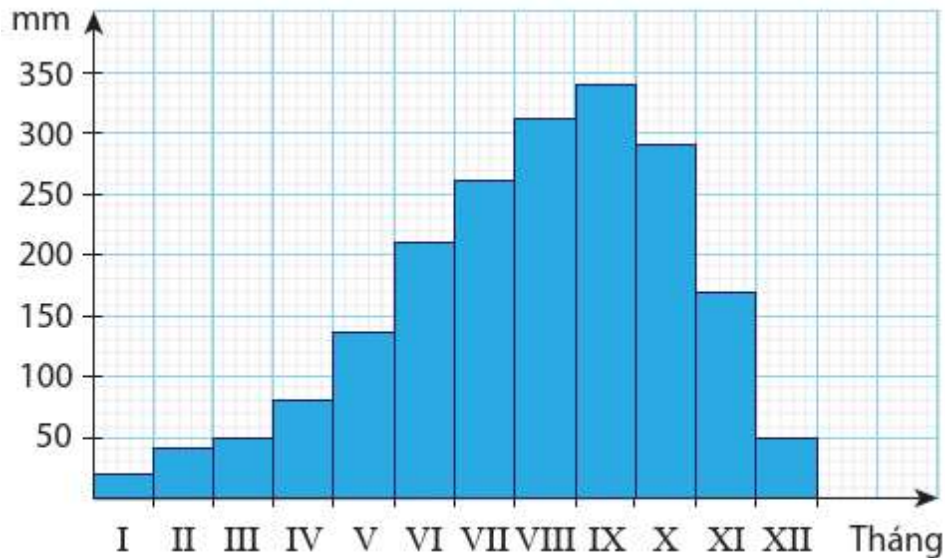
b) Tính khoảng biến thiên, khoảng tứ phân vị của mẫu số liệu gốc và mẫu số liệu ghép nhóm thu được ở câu a. Giá trị nào là giá trị chính xác, giá trị nào là giá trị xấp xỉ?

**Bài 15.** Thu nhập theo tháng (đơn vị: triệu đồng) của người lao động ở hai nhà máy như sau:

Thu nhập	[5; 8)	[8; 11)	[11; 14)	[14; 17)	[17; 20)
Số người của nhà máy A	20	35	45	35	20
Số người của nhà máy B	17	23	30	23	17

Tính mức thu nhập trung bình của người lao động ở hai nhà máy trên. Dựa vào khoảng tứ phân vị, hãy xác định xem mức thu nhập của người lao động ở nhà máy nào biến động nhiều hơn.

**Bài 16.** Hình 3.2 là biểu đồ biểu diễn lượng mưa trung bình của các tháng trong năm ở thành phố A.



Hình 3.2. Biểu đồ lượng mưa ở thành phố A

- a) Lập bảng số liệu ghép nhóm về lượng mưa của thành phố A, với độ dài các nhóm là 50 và đầu mút phải của nhóm cuối cùng là 350.
- b) Xác định khoảng tứ phân vị của mẫu số liệu (làm tròn kết quả đến hàng đơn vị). Nêu ý nghĩa của kết quả tìm được.

**Giải**

a) Bảng số liệu về lượng mưa của thành phố A

Lượng mưa (mm)	Tần số	Tần số tích lũy
[0; 50)	2	2
[50; 100)	3	5
[100; 150)	1	6
[150; 200)	1	7
[200; 250)	1	8
[250; 300)	2	10
[300; 350)	2	12

b)  $Q_1 \approx 67$ ;  $Q_3 = 275$ ;  $\Delta_Q = Q_3 - Q_1 = 208$

Kết quả tìm được cho thấy: Hàng năm, ở thành phố A có 3 tháng có lượng mưa trung bình không vượt quá 67 mm và 3 tháng có lượng mưa trung bình ít nhất là 275 mm. Trong 6 tháng còn lại, lượng mưa trung bình đạt từ 67 mm đến 275 mm và như vậy là lượng mưa của 6 tháng này có thể chênh lệch nhau đến 208 mm.

**Bài 17.** Bảng sau đây cho biết chiều cao của các học sinh lớp 12A và 12B

Chiều cao (cm)	[145; 150)	[150; 155)	[155; 160)	[160; 165)	[165; 170)	[170; 175)
Số hs lớp 12A	1	0	15	12	10	5
Số hs lớp 12B	0	0	17	10	9	6

- a) Tìm khoảng biến thiên, khoảng tứ phân vị cho các mẫu số liệu ghép nhóm về chiều cao của học sinh lớp 12A và 12B.
- b) Để so sánh độ phân tán về chiều cao của học sinh hai lớp này ta nên dùng khoảng biến thiên hay khoảng tứ phân vị? Vì sao