

CÁC SỐ ĐẶC TRƯNG ĐO MỨC ĐỘ PHÂN TÁN CỦA MẪU SỐ LIỆU GHÉP NHÓM

BÀI

01

KHOẢNG BIẾN THIÊN VÀ KHOẢNG TỨ PHÂN VỊ

A

LÝ THUYẾT CẦN NHỚ

1 Khoảng biến thiên

Định nghĩa: Cho mẫu số liệu ghép nhóm:

Nhóm	$[a_1; a_2)$	...	$[a_i; a_{i+1})$	...	$[a_k; a_{k+1})$
Tần số	m_1	...	m_i	...	m_k

trong đó các tần số $m_1 > 0, m_k > 0$ và $n = m_1 + \dots + m_k$ là cỡ mẫu.

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là $R = a_{k+1} - a_1$.

Ý nghĩa:

- Khoảng biến thiên của mẫu số liệu ghép nhóm xấp xỉ cho khoảng biến thiên của mẫu số liệu gốc. Khoảng biến thiên được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm. Khoảng biến thiên càng lớn thì mẫu số liệu càng phân tán.
- Trong các đại lượng đo mức độ phân tán của mẫu số liệu ghép nhóm, khoảng biến thiên là đại lượng dễ hiểu, dễ tính toán. Tuy nhiên, do khoảng biến thiên chỉ sử dụng hai giá trị là u_1 và u_{m+1} của mẫu số liệu nên đại lượng đó dễ bị ảnh hưởng bởi các giá trị bất thường.

2 Khoảng tứ phân vị

Định nghĩa: Cho mẫu số liệu ghép nhóm:

Nhóm	$[a_1; a_2)$	...	$[a_i; a_{i+1})$	...	$[a_k; a_{k+1})$
Tần số	m_1	...	m_i	...	m_k

Tứ phân vị thứ r là $Q_r = a_p + \frac{\frac{r \cdot n}{4} - (m_1 + \dots + m_{p-1})}{m_p} \cdot (a_{p+1} - a_p)$, trong đó $[a_p; a_{p+1})$ là nhóm chứa tứ phân vị thứ r với $r = 1, 2, 3$.

Khoảng tứ phân vị của mẫu số liệu ghép nhóm, kí hiệu là Δ_Q , là hiệu số giữa tứ phân vị thứ ba Q_3 và tứ phân vị thứ nhất Q_1 của mẫu số liệu đó, tức là $\Delta_Q = Q_3 - Q_1$.

**Ý nghĩa:**

- Khoảng tứ phân vị của mẫu số liệu ghép nhóm xấp xỉ cho khoảng tứ phân vị của mẫu số liệu gốc. Khoảng tứ phân vị cũng được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm. Khoảng tứ phân vị càng lớn thì mẫu số liệu càng phân tán.
- 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 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 $x < Q_1 - 1,5\Delta_Q$
- Do khoảng tứ phân vị của mẫu số liệu ghép nhóm chỉ phụ thuộc vào nửa giữa của mẫu số liệu, nên không bị ảnh hưởng bởi các giá trị bất thường và có thể dùng đại lượng này để loại giá trị bất thường.

B PHÂN LOẠI VÀ PHƯƠNG PHÁP GIẢI TOÁN**Dạng 1: Tìm khoảng biến thiên của mẫu số liệu ghép nhóm**

- Xác định u_1 là giá trị đầu mút trái của nhóm đầu tiên và u_{k+1} là giá trị đầu mút phải của nhóm cuối cùng có chứa dữ liệu (tần số khác 0).
- Khoảng biến thiên $R = u_{k+1} - u_1$

BÀI TẬP TỰ LUẬN

Bài tập 1: Thời gian hoàn thành bài kiểm tra môn Toán của các bạn 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 cho 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 tập 2: Bảng dưới biểu thị kết quả điều tra thời gian sử dụng Internet hằng ngày của một số người.

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?

Bài tập 3: Bảng dưới 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.

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



Bài tập 4: 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 dưới biểu thị hai mẫu số liệu mà anh thu thập được qua Internet. Bảng. 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?

Bài tập 5: 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. 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 tập 6: 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.

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

Bài tập 7: Thống kê thời gian sử dụng mạng xã hội trong ngày của các bạn 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 cho thời gian sử dụng mạng xã hội của học sinh mỗi tổ và giải thích ý nghĩa.



**BÀI TẬP TRẮC NGHIỆM**

PHẦN I. Câu trắc nghiệm nhiều phương án lựa chọn. Mỗi câu hỏi thí sinh chỉ chọn một phương án.

Câu 1. 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.

- A. 25. B. 30. C. 6. D. 69,8.

Câu 2. Bảng dưới biểu diễn mẫu số liệu ghép nhóm về số tiền (đơn vị: nghìn đồng) mà 60 khách hàng mua sách ở một cửa hàng trong một ngày.

Nhóm	Tần số
[40;50)	3
[50;60)	6
[60;70)	19
[70;80)	23
[80;90)	9
	$n = 60$

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là:

- A. 50. B. 30. C. 6. D. 69,8.

Câu 3. Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khỏe. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

Quãng đường (km)	[2, 7; 3, 0)	[3, 0; 3, 3)	[3, 3; 3, 6)	[3, 6; 3, 9)	[3, 9; 4, 2)
Số ngày	3	6	5	4	2

Khoảng biến thiên của mẫu số liệu ghép nhóm là

- A. 1,5. B. 0,9. C. 0,6. D. 0,3.

Câu 4. Bạn Chi rất thích nhảy hiện đại. Thời gian tập nhảy mỗi ngày trong thời gian gần đây của bạn Chi được thống kê lại ở bảng sau:

Thời gian (phút)	[20;25)	[25;30)	[30;35)	[35;40)	[40;45)
Số ngày	6	6	4	1	1

Khoảng biến thiên của mẫu số liệu ghép nhóm là

- A. 25. B. 20. C. 15. D. 30.

Câu 5. Dũng là học sinh rất giỏi chơi rubik, bạn có thể giải nhiều loại khối rubik khác nhau. Trong một lần tập luyện giải khối rubik 3×3 , bạn Dũng đã tự thống kê lại thời gian giải rubik trong 25 lần giải liên tiếp ở bảng sau:



Thời gian giải rubik (giây)	[8;10)	[10;12)	[12;14)	[14;16)	[16;18)
Số lần	4	6	8	4	3

Khoảng biến thiên của mẫu số liệu ghép nhóm nhận giá trị nào trong các giá trị dưới đây?

- A. 6. B. 8. C. 10. D. 12.

Câu 6. Một vườn thú ghi lại tuổi thọ (đơn vị: năm) của 20 con hổ và thu được kết quả như sau:

Tuổi thọ	[14;15)	[15;16)	[16;17)	[17;18)	[18;19)
Số con hổ	1	3	8	6	2

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A. 3. B. 4. C. 5. D. 6.

Câu 7. Khảo sát thời gian tập thể dục của một số học sinh khối 11 thu được mẫu số liệu ghép nhóm sau:

Thời gian	[0; 20)	[20; 40)	[40; 60)	[60; 80)	[80; 100)
Số học sinh	5	9	12	10	6

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A. 80. B. 60. C. 100. D. 12.

Câu 8. Mức thưởng tết (triệu đồng) cho các nhân viên của một công ty được thống kê trong bảng sau:

Mức thưởng tết	[5; 10)	[10; 15)	[15; 20)	[20; 25)	[25; 30)
Số nhân viên	13	35	47	25	10

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A. 20. B. 25. C. 47. D. 23

Câu 9. Cho bảng phân bố tần số ghép lớp sau. Chiều cao của 40 học sinh nam ở một trường THPT

Lớp chiều cao (cm)	[160; 163,5)	[164; 167,5)	[168; 171,5)	[172; 175,5)	Cộng
Tần số	9	20	7	4	40

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A. 31. B. 15,5. C. 74. D. 32.

Câu 10. Thời gian truy cập Internet mỗi buổi tối của một số học sinh được cho trong bảng sau:

Thời gian (phút)	[9,5; 12,5)	[12,5; 15,5)	[15,5; 18,5)	[18,5; 21,5)	[21,5; 24,5)
Số học sinh	0	12	15	24	26

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A. 26. B. 15. C. 20. D. 12.

Câu 11. Thời gian hoàn thành bài kiểm tra môn Toán của các bạn trong lớp 12C được cho trong bảng sau:

Thời gian (phút)	[25; 30)	[30; 35)	[35; 40)	[40; 45)
Số học sinh	8	16	12	2

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A. 20. B. 15. C. 27. D. 12.



Câu 12. Thời gian hoàn thành giải chạy của các vận động viên được cho như bảng sau:

Thời gian (phút)	[15; 20)	[20; 25)	[25; 30)	[30; 35)
Số vận động viên	7	11	4	6

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A. 15.
- B. 25.
- C. 37.
- D. 20.

-----HẾT-----



CÁC SỐ ĐẶC TRƯNG ĐO MỨC ĐỘ PHÂN TÁN CỦA MẪU SỐ LIỆU GHÉP NHÓM

BÀI

01

KHOẢNG BIẾN THIÊN VÀ KHOẢNG TỨ PHÂN VỊ

A

LÝ THUYẾT CẦN NHỚ

1 Khoảng biến thiên

Định nghĩa: Cho mẫu số liệu ghép nhóm:

Nhóm	$[a_1; a_2)$	...	$[a_i; a_{i+1})$	...	$[a_k; a_{k+1})$
Tần số	m_1	...	m_i	...	m_k

trong đó các tần số $m_1 > 0, m_k > 0$ và $n = m_1 + \dots + m_k$ là cỡ mẫu.

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là $R = a_{k+1} - a_1$.

Ý nghĩa:

- Khoảng biến thiên của mẫu số liệu ghép nhóm xấp xỉ cho khoảng biến thiên của mẫu số liệu gốc. Khoảng biến thiên được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm. Khoảng biến thiên càng lớn thì mẫu số liệu càng phân tán.
- Trong các đại lượng đo mức độ phân tán của mẫu số liệu ghép nhóm, khoảng biến thiên là đại lượng dễ hiểu, dễ tính toán. Tuy nhiên, do khoảng biến thiên chỉ sử dụng hai giá trị là u_1 và u_{m+1} của mẫu số liệu nên đại lượng đo dễ bị ảnh hưởng bởi các giá trị bất thường.

2 Khoảng tứ phân vị

Định nghĩa: Cho mẫu số liệu ghép nhóm:

Nhóm	$[a_1; a_2)$	...	$[a_i; a_{i+1})$	...	$[a_k; a_{k+1})$
Tần số	m_1	...	m_i	...	m_k

Tứ phân vị thứ r là $Q_r = a_p + \frac{\frac{r \cdot n}{4} - (m_1 + \dots + m_{p-1})}{m_p} \cdot (a_{p+1} - a_p)$, trong đó $[a_p; a_{p+1})$ là nhóm chứa tứ phân vị thứ r với $r = 1, 2, 3$.

Khoảng tứ phân vị của mẫu số liệu ghép nhóm, kí hiệu là Δ_Q , là hiệu số giữa tứ phân vị thứ ba Q_3 và tứ phân vị thứ nhất Q_1 của mẫu số liệu đó, tức là $\Delta_Q = Q_3 - Q_1$.



Ý nghĩa:

- Khoảng tứ phân vị của mẫu số liệu ghép nhóm xấp xỉ cho khoảng tứ phân vị của mẫu số liệu gốc. Khoảng tứ phân vị cũng được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm. Khoảng tứ phân vị càng lớn thì mẫu số liệu càng phân tán.
- 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 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 $x < Q_1 - 1,5\Delta_Q$
- Do khoảng tứ phân vị của mẫu số liệu ghép nhóm chỉ phụ thuộc vào nửa giữa của mẫu số liệu, nên không bị ảnh hưởng bởi các giá trị bất thường và có thể dùng đại lượng này để loại giá trị bất thường.

B PHÂN LOẠI VÀ PHƯƠNG PHÁP GIẢI TOÁN

Dạng 1: Tìm khoảng biến thiên của mẫu số liệu ghép nhóm

- Xác định u_1 là giá trị đầu mút trái của nhóm đầu tiên và u_{k+1} là giá trị đầu mút phải của nhóm cuối cùng có chứa dữ liệu (tần số khác 0).
- Khoảng biến thiên $R = u_{k+1} - u_1$

BÀI TẬP TỰ LUẬN

Bài tập 1: Thời gian hoàn thành bài kiểm tra môn Toán của các bạn 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 cho 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?

Lời giải

- a) Khoảng biến thiên R cho mẫu số liệu ghép nhóm trên là: $45 - 25 = 20$
- b) Khoảng biến thiên của mẫu số liệu gốc là: $43 - 27 = 16$.

Bài tập 2: Bảng dưới biểu thị kết quả điều tra thời gian sử dụng Internet hằng ngày của một số người.

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?

Lời 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 tập 3: Bảng dưới 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.

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

Lời giải

Khoảng biến thiên của mẫu số liệu là $R = 300 - 150 = 150$.

Bài tập 4: 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 dưới biểu thị hai mẫu số liệu mà anh thu thập được qua Internet. Bảng. 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?

Lời 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 tập 5: 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. 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?

Lời giải

$$R_A = 100 - 50 = 50$$

$$R_B = 90 - 60 = 30$$

Ta thấy $R_A > R_B$ nên nhóm khán giả A phân tán hơn



Bài tập 6: 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.

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

Lời giải

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).

Bài tập 7: Thống kê thời gian sử dụng mạng xã hội trong ngày của các bạn 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 cho thời gian sử dụng mạng xã hội của học sinh mỗi tổ và giải thích ý nghĩa.

Lời giải

Gọi R_1, R_2 tương ứng là khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian sử dụng mạng xã hội trong ngày của các bạn Tổ 1 và Tổ 2.

Ta có: $R_1 = 90 - 0 = 90$ và $R_2 = 60 - 0 = 60$.

Do $R_1 > R_2$ nên ta có thể kết luận rằng thời gian sử dụng mạng xã hội trong ngày của các bạn Tổ 1 phân tán hơn thời gian sử dụng mạng xã hội của các bạn Tổ 2.





BÀI TẬP TRẮC NGHIỆM

PHẦN I. Câu trắc nghiệm nhiều phương án lựa chọn. Mỗi câu hỏi thí sinh chỉ chọn một phương án.

Câu 1. 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.

- A. 25. B. 30. C. 6. D. 69,8.

Lời giải

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

Câu 2. Bảng dưới biểu diễn mẫu số liệu ghép nhóm về số tiền (đơn vị: nghìn đồng) mà 60 khách hàng mua sách ở một cửa hàng trong một ngày.

Nhóm	Tần số
[40;50)	3
[50;60)	6
[60;70)	19
[70;80)	23
[80;90)	9
	$n = 60$

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là:

- A. 50. B. 30. C. 6. D. 69,8.

Lời giải

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là: $R = 90 - 40 = 50$

Câu 3. Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khỏe. Quảng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

Quảng đường (km)	[2,7;3,0)	[3,0;3,3)	[3,3;3,6)	[3,6;3,9)	[3,9;4,2)
Số ngày	3	6	5	4	2

Khoảng biến thiên của mẫu số liệu ghép nhóm là

- A. 1,5. B. 0,9. C. 0,6. D. 0,3.

Lời giải

Khoảng biến thiên của mẫu số liệu ghép nhóm là: $4,2 - 2,7 = 1,5\text{ km}$

Câu 4. Bạn Chi rất thích nhảy hiện đại. Thời gian tập nhảy mỗi ngày trong thời gian gần đây của bạn Chi được thống kê lại ở bảng sau:



Thời gian (phút)	[20;25)	[25;30)	[30;35)	[35;40)	[40;45)
Số ngày	6	6	4	1	1

Khoảng biến thiên của mẫu số liệu ghép nhóm là

- A. 25. B. 20. C. 15. D. 30.

Lời giải

Khoảng biến thiên của mẫu số liệu ghép nhóm là: $45 - 20 = 25$ (phút)

- Câu 5.** Dũng là học sinh rất giỏi chơi rubik, bạn có thể giải nhiều loại khối rubik khác nhau. Trong một lần tập luyện giải khối rubik 3×3 , bạn Dũng đã tự thống kê lại thời gian giải rubik trong 25 lần giải liên tiếp ở bảng sau:

Thời gian giải rubik (giây)	[8;10)	[10;12)	[12;14)	[14;16)	[16;18)
Số lần	4	6	8	4	3

Khoảng biến thiên của mẫu số liệu ghép nhóm nhận giá trị nào trong các giá trị dưới đây?

- A. 6. B. 8. C. 10. D. 12.

Khoảng biến thiên của mẫu số liệu là: $18 - 8 = 10$ (giây)

- Câu 6.** Một vườn thú ghi lại tuổi thọ (đơn vị: năm) của 20 con hổ và thu được kết quả như sau:

Tuổi thọ	[14;15)	[15;16)	[16;17)	[17;18)	[18;19)
Số con hổ	1	3	8	6	2

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A. 3. B. 4. C. 5. D. 6.

Lời giải

Khoảng biến thiên: $19 - 14 = 5$

- Câu 7.** Khảo sát thời gian tập thể dục của một số học sinh khối 11 thu được mẫu số liệu ghép nhóm sau:

Thời gian	[0;20)	[20;40)	[40;60)	[60;80)	[80;100)
Số học sinh	5	9	12	10	6

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A. 80. B. 60. C. 100. D. 12.

Lời giải

Khoảng biến thiên: $100 - 0 = 100$

- Câu 8.** Mức thưởng tết (triệu đồng) cho các nhân viên của một công ty được thống kê trong bảng sau:

Mức thưởng tết	[5; 10)	[10; 15)	[15; 20)	[20; 25)	[25; 30)
Số nhân viên	13	35	47	25	10

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A. 20. B. 25. C. 47. D. 23

**Lời giải**

Khoảng biến thiên: $30 - 5 = 25$

Câu 9. Cho bảng phân bố tần số ghép lớp sau. Chiều cao của 40 học sinh nam ở một trường THPT

Lớp chiều cao (cm)	[160; 163,5)	[164; 167,5)	[168; 171,5)	[172; 175,5)	Cộng
Tần số	9	20	7	4	40

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A. 31. **B.** 15,5. C. 74. D. 32.

Lời giải

Khoảng biến thiên: $175,5 - 160 = 15,5$

Câu 10. Thời gian truy cập Internet mỗi buổi tối của một số học sinh được cho trong bảng sau:

Thời gian (phút)	[9,5; 12,5)	[12,5; 15,5)	[15,5; 18,5)	[18,5; 21,5)	[21,5; 24,5)
Số học sinh	0	12	15	24	26

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A. 26. **B.** 15. C. 20. D. 12.

Lời giải

Khoảng biến thiên: $24,5 - 9,5 = 15$.

Câu 11. Thời gian hoàn thành bài kiểm tra môn Toán của các bạn trong lớp 12C được cho trong bảng sau:

Thời gian (phút)	[25; 30)	[30; 35)	[35; 40)	[40; 45)
Số học sinh	8	16	12	2

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A.** 20. B. 15. C. 27. D. 12.

Lời giải

Khoảng biến thiên: $45 - 25 = 20$.

Câu 12. Thời gian hoàn thành giải chạy của các vận động viên được cho như bảng sau:

Thời gian (phút)	[15; 20)	[20; 25)	[25; 30)	[30; 35)
Số vận động viên	7	11	4	6

Khoảng biến thiên của mẫu số liệu ghép nhóm này là

- A. 15. B. 25. C. 37. **D.** 20.

Lời giải

Khoảng biến thiên: $35 - 15 = 20$.

-----**HẾT**-----



B PHÂN LOẠI VÀ PHƯƠNG PHÁP GIẢI TOÁN

Dạng 2: Tìm tứ phân vị của mẫu số liệu ghép nhóm

Với mẫu số liệu ghép nhóm:

Nhóm	$[a_1; a_2)$	...	$[a_i; a_{i+1})$	...	$[a_k; a_{k+1})$
Tần số	m_1	...	m_i	...	m_k

Các bước thực hiện:

- Tìm tứ phân vị Q_1 và Q_3 theo công thức:

$$Q_r = a_p + \frac{\frac{r \cdot n}{4} - (m_1 + \dots + m_{p-1})}{m_p} \cdot (a_{p+1} - a_p)$$

trong đó $[a_p; a_{p+1})$ là nhóm chứa tứ phân vị thứ r với $r = 1, 2, 3$ và n là cỡ mẫu.

- Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = Q_3 - Q_1$

BÀI TẬP TỰ LUẬN

Bài tập 1: Thời gian chờ khám bệnh của các bệnh nhân tại phòng khám X được cho trong bảng sau:

Thời gian (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 tại 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 tập 2: Một người ghi lại thời gian đàm thoại của một số cuộc gọi cho kết quả như bảng sau:

Thời gian t (phút)	Số cuộc gọi
$0 \leq t < 1$	8
$1 \leq t < 2$	17
$2 \leq t < 3$	25
$3 \leq t < 4$	20
$4 \leq t < 5$	10

Tính khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.

Bài tập 3: Thống kê số ngày trong tháng Sáu năm 2021 và năm 2022 theo nhiệt độ cao nhất trong ngày tại Hà Nội, người ta thu được bảng sau:

Nhiệt độ ($^{\circ}\text{C}$)	$[28; 30)$	$[30; 32)$	$[32; 34)$	$[34; 36)$	$[36; 38)$	$[38; 40)$
Số ngày trong tháng 6/2021	0	2	8	5	6	9
Số ngày trong tháng 6/2022	2	3	4	11	8	2

Hỏi tháng sáu năm nào ở Hà Nội nhiệt độ cao nhất trong ngày biến đổi nhiều hơn?



Bài tập 4: Thống 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 tập 5: 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 tập 6: Bảng sau đây cho biết chiều cao của các học sinh lớp 12A và 12B.

Thu nhập	$[145; 150)$	$[150; 155)$	$[155; 160)$	$[160; 165)$	$[165; 170)$	$[170; 175)$
Số học sinh lớp 12A	1	0	15	12	10	5
Số học sinh 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, 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?

Bài tập 7: Điểm kiểm tra cuối khoá môn Tiếng Anh của hai lớp ở một trung tâm ngoại ngữ được thống kê trong các bảng a) và b)

Bảng a). 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

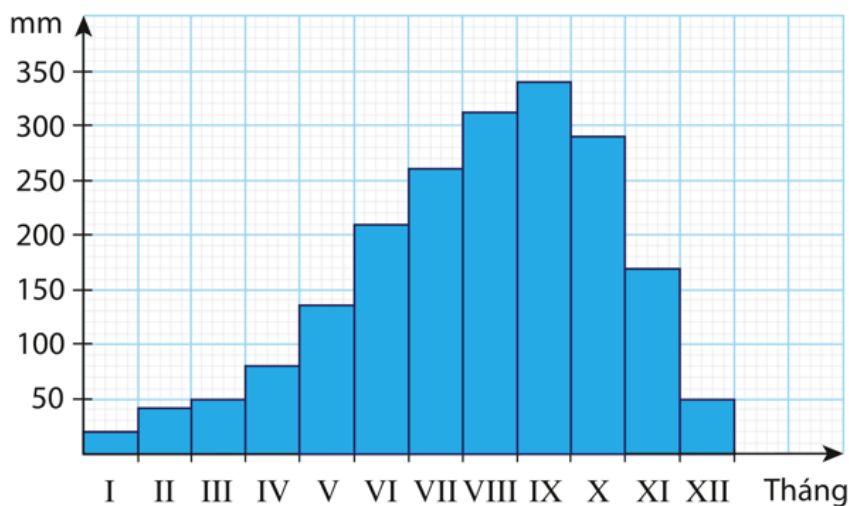
Bảng b). Điểm của lớp B

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

[80;90)	20
[90;100)	15

- 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?
- 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? Minh họa câu trả lời bằng cách biểu diễn các tứ phân vị và khoảng tứ phân vị của mỗi mẫu số liệu trên trục số.

Bài tập 8: Hình dưới 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

- 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.
- 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.

Bài tập 9: Hằng ngày ông Thắng đều đi xe buýt từ nhà đến cơ quan. Dưới đây là bản 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 lần	22	38	27	8	4	1

- Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm phía trên (Làm tròn kết quả đến hàng phần trăm)
- Biết rằng trong 100 lần đi trên, chỉ có đúng một lần ông Thắng đi hết 32 phút. Thời gian của lần đi đó có phải giá trị ngoại lệ không?



BÀI TẬP TRẮC NGHIỆM

PHẦN I. Câu trắc nghiệm nhiều phương án lựa chọn. Mỗi câu hỏi thí sinh chỉ chọn một phương án.

Câu 1. Một vườn thú ghi lại tuổi thọ (đơn vị: năm) của 20 con hổ và thu được kết quả như sau:

Tuổi thọ	[14;15)	[15;16)	[16;17)	[17;18)	[18;19)
Số con hổ	1	3	8	6	2

Nhóm chứa tứ phân vị thứ nhất là

- A. [14;15). B. [15;16). C. [16;17). D. [17;18).

Câu 2. Một vườn thú ghi lại tuổi thọ (đơn vị: năm) của 20 con hổ và thu được kết quả như sau:

Tuổi thọ	[14;15)	[15;16)	[16;17)	[17;18)	[18;19)
Số con hổ	1	3	8	6	2

Nhóm chứa tứ phân vị thứ ba là

- A. [15;16). B. [16;17). C. [17;18). D. [18;19).

Câu 3. Khảo sát thời gian tập thể dục trong ngày của một số học sinh khối 11 thu được mẫu số liệu ghép nhóm sau:

Thời gian (phút)	[0; 20)	[20; 40)	[40; 60)	[60; 80)	[80; 100)
Số học sinh	5	9	12	10	6

Nhóm chứa tứ phân vị thứ nhất là

- A. [0; 20). B. [20; 40). C. [40; 60). D. [60; 80).

Câu 4. Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một cửa hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

Doanh thu	[5; 7)	[7; 9)	[9; 11)	[11; 13)	[13; 15)
Số ngày	2	7	7	3	1

Tứ phân vị thứ nhất của mẫu số liệu trên gần nhất với giá trị nào trong các giá trị sau?

- A. 7. B. 7,6. C. 8. D. 8,6.

Câu 5. Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một cửa hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

Doanh thu	[5; 7)	[7; 9)	[9; 11)	[11; 13)	[13; 15)
Số ngày	2	7	7	3	1

Tứ phân vị thứ ba của mẫu số liệu trên gần nhất với giá trị nào trong các giá trị sau?

- A. 10. B. 11. C. 12. D. 13.



Câu 6. Cho mẫu số liệu ghép nhóm về tuổi thọ (đơn vị tính là năm) của một loại bóng đèn mới như sau.

Tuổi thọ	[2;3,5)	[3,5;5)	[5;6,5)	[6,5;8)
Số bóng đèn	8	22	35	15

Nhóm chứa tứ phân vị thứ nhất của mẫu số liệu là

- A. [2;3,5). B. [3,5;5). C. [5;6,5). D. [6,5;8).

Câu 7. Cho mẫu số liệu ghép nhóm về tuổi thọ (đơn vị tính là năm) của một loại bóng đèn mới như sau.

Tuổi thọ	[2;3,5)	[3,5;5)	[5;6,5)	[6,5;8)
Số bóng đèn	8	22	35	15

Nhóm chứa tứ phân vị thứ ba của mẫu số liệu là

- A. [2;3,5). B. [3,5;5). C. [5;6,5). D. [6,5;8).

Câu 8. Cho bảng số liệu về chiều cao của 100 học sinh một trường trung học phổ thông dưới đây.

Nhóm	Chiều cao (cm)	Số học sinh
1	[150;153)	7
2	[153;156)	13
3	[156;159)	40
4	[159;162)	21
5	[162;165)	13
6	[165;168)	6

Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm trên (làm tròn kết quả đến hàng phần trăm) là

- A. 156,25. B. 157,5. C. 156,38. D. 157,54.

Câu 9. Cho bảng số liệu về chiều cao của 100 học sinh một trường trung học phổ thông dưới đây.

Nhóm	Chiều cao (cm)	Số học sinh
1	[150;153)	7
2	[153;156)	13
3	[156;159)	40
4	[159;162)	21
5	[162;165)	13
6	[165;168)	6

Tứ phân vị thứ ba của mẫu số liệu ghép nhóm trên (làm tròn kết quả đến hàng phần trăm) là:

- A. 160,52. B. 161,52. C. 161,14. D. 162,25.

Câu 10. Bảng dưới biểu diễn mẫu số liệu ghép nhóm về số tiền (đơn vị: nghìn đồng) mà 60 khách hàng mua sách ở một cửa hàng trong một ngày.



Nhóm	Tần số
[40;50)	3
[50;60)	6
[60;70)	19
[70;80)	23
[80;90)	9
	$n = 60$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là:

- A. 50. B. 40. C. 14,23. D. 70,87.

Câu 11. Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khỏe. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

Quãng đường (km)	[2, 7; 3, 0)	[3, 0; 3, 3)	[3, 3; 3, 6)	[3, 6; 3, 9)	[3, 9; 4, 2)
Số ngày	3	6	5	4	2

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

- A. 0,9. B. 0,975. C. 0,5. D. 0,575.

Câu 12. Bạn Chi rất thích nhảy hiện đại. Thời gian tập nhảy mỗi ngày trong thời gian gần đây của bạn Chi được thống kê lại ở bảng sau:

Thời gian (phút)	[20;25)	[25;30)	[30;35)	[35;40)	[40;45)
Số ngày	6	6	4	1	1

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

- A. 23,75. B. 27,5. C. 31,88. D. 8,125.

Câu 13. Dũng là học sinh rất giỏi chơi rubik, bạn có thể giải nhiều loại khối rubik khác nhau. Trong một lần tập luyện giải khối rubik 3×3 , bạn Dũng đã tự thống kê lại thời gian giải rubik trong 25 lần giải liên tiếp ở bảng sau:

Thời gian giải rubik (giây)	[8;10)	[10;12)	[12;14)	[14;16)	[16;18)
Số lần	4	6	8	4	3

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

- A. 10,75. B. 1,75. C. 3,63. D. 14,38.

Câu 14. Thời gian (phút) truy cập Internet mỗi buổi tối của một số học sinh được cho trong bảng sau:

Thời gian (phút)	[9, 5; 12, 5)	[12, 5; 15, 5)	[15, 5; 18, 5)	[18, 5; 21, 5)	[21, 5; 24, 5)
Số học sinh	3	12	15	24	2

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

- A. 10,75. B. 4,75. C. 4,63. D. 4,38.



Câu 15. Ghi lại tốc độ bóng trong 200 lần giao bóng của một vận động viên môn quần vợt cho kết quả như bảng bên.

Tốc độ v (km/h)	Số lần
$150 \leq v < 155$	18
$155 \leq v < 160$	28
$160 \leq v < 165$	35
$165 \leq v < 170$	43
$170 \leq v < 175$	41
$175 \leq v < 180$	35

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

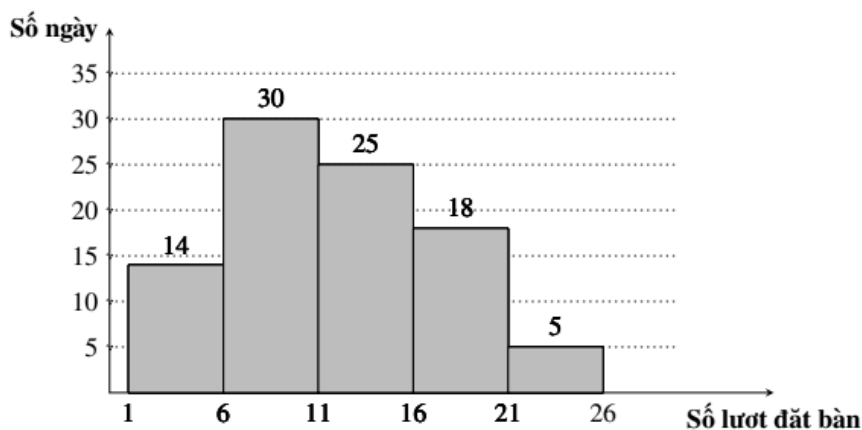
A. 12,6.

B. 11,5.

C. 14,3.

D. 16,8.

Câu 16. Biểu đồ dưới đây biểu diễn số lượng khách hàng đặt bàn qua hình thức trực tuyến mỗi ngày trong quý III năm 2022 của một nhà hàng. Cột thứ nhất biểu diễn số ngày có từ 1 đến 6 lượt đặt bàn, cột thứ hai biểu diễn số ngày có từ 6 đến 11 lượt đặt bàn;...



Khoảng tứ phân vị của mẫu số liệu ghép nhóm cho bởi biểu đồ trên

A. 9,5.

B. 8,5.

C. 10,5.

D. 7,5.

PHẦN II. Câu trắc nghiệm đúng sai. Trong mỗi ý a), b), c), d) ở mỗi câu, thí sinh chọn đúng hoặc sai.

Câu 1. 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) Lớp	[155;160)	[160;165)	[165;170)	[170;175)	[175;180)	[180;185)
12C	2	7	12	3	0	1
12D	5	9	8	2	1	0

Xét tính đúng sai của các khẳng định sau:

a) Chiều cao cao nhất của các bạn học sinh trong lớp 12D là 185 (cm)



- b) Khoảng biến thiên của chiều cao của các bạn học sinh nữ lớp 12C là: 30 (cm)
- c) Khoảng biến thiên của chiều cao của các bạn học sinh nữ lớp 12D là: 25 (cm)
- d) Chiều cao của học sinh lớp 12C có độ phân tán bé hơn

Câu 2. Bảng sau biểu diễn mẫu số liệu ghép nhóm thống kê mức lương của một công ty (đơn vị: triệu đồng).

Nhóm	Tần số
[10;15)	15
[15;20)	18
[20;25)	10
[25;30)	10
[30;35)	5
[35;40)	2
	$n = 60$

Xét tính đúng sai của các khẳng định sau:

- a) Khoảng biến thiên của mẫu số liệu ghép nhóm trên là: $R = 30$
- b) Số phần tử của mẫu là $n = 60$
- c) Tứ phân vị thứ nhất là: $Q_1 = 15$
- d) Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $\Delta_Q = 3$

Câu 3. Bảng sau biểu diễn mẫu số liệu ghép nhóm về độ tuổi của cư dân trong một khu phố.

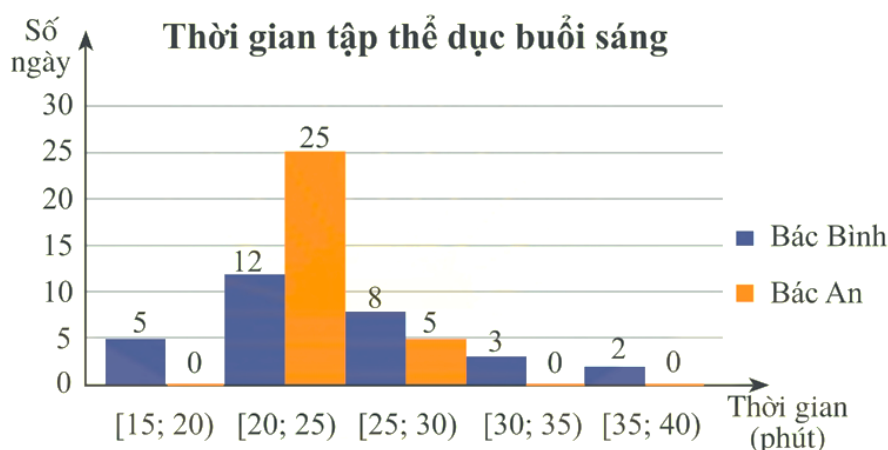
Nhóm	Tần số
[20;30)	25
[30;40)	20
[40;50)	20
[50;60)	15
[60;70)	14
[70;80)	6
	$n = 100$

Xét tính đúng sai của các khẳng định sau:

- a) Khoảng biến thiên của mẫu số liệu ghép nhóm trên là: $R = 60$
- b) Tứ phân vị thứ nhất là: $Q_1 = 35$
- c) Tứ phân vị thứ ba là: $Q_3 = \frac{160}{3}$
- d) Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $\Delta_Q = \frac{65}{3}$



Câu 4. 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.



Xét tính đúng sai của các khẳng định sau:

a) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bác Bình là 25 (phút).

b) Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bác An là: $\Delta_Q = 2$

c) Tứ phân vị thứ ba của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bác Bình là: $Q_3' = \frac{455}{16}$

d) Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng mỗi ngày của bác An lớn hơn bác Bình

Câu 5. 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

Xét tính đúng sai của các khẳng định sau:

a) Cỡ mẫu $n = 100$.

b) Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là $Q_1 = \frac{683}{38}$.

c) Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = \frac{515}{114}$

d) 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 là giá trị ngoại lệ của mẫu số liệu ghép nhóm.

Câu 6. 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

Xét tính đúng sai của các khẳng định sau:

- Khoảng biến thiên của mẫu số liệu ghép nhóm ứng với khu vực A là: 15 (tuổi)
- Khoảng biến thiên của mẫu số liệu ghép nhóm ứng với khu vực B là: 12 (tuổi)
- Khoảng tứ phân vị của mẫu số liệu ghép nhóm ứng với khu vực A là: $\frac{61}{3}$ (tuổi)
- Nếu so sánh theo khoảng tứ phân vị thì phụ nữ ở khu vực B có độ tuổi kết hôn đồng đều hơn

Câu 7. Bảng sau thống kê tổng lượng mưa (đơn vị: mm) đo được vào tháng 7 từ năm 2002 đến 2021 tại một trạm quan trắc đặt ở Cà Mau.

341,4	187,1	242,2	522,9	251,4	432,2	200,7	388,6	258,4	288,5
298,1	413,5	413,5	332	421	475	400	305	520	147

Xét tính đúng sai của các khẳng định sau:

- Khoảng biến thiên của mẫu số liệu: 375,9(mm)
- Khoảng tứ phân vị của mẫu số liệu là: $\Delta_Q = \frac{1827}{100}$
- Chia mẫu số liệu trên thành 4 nhóm như bảng:

Lượng mưa	[140;240)	[240;340)	[340;440)	[440;540)
Số tháng	3	7	7	3

Khoảng biến thiên của mẫu số liệu ghép nhóm: 400 (mm)

- Chia mẫu số liệu trên thành 4 nhóm như bảng:

Lượng mưa	[140;240)	[240;340)	[340;440)	[440;540)
Số tháng	3	7	7	3

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q' = \frac{1000}{7}$

Câu 8. Kết quả đo chiều cao của 100 cây keo 3 năm tuổi tại một nông trường được cho ở bảng sau:

Chiều cao (m)	[8,4;8,6)	[8,6;8,8)	[8,8;9,0)	[9,0;9,2)	[9,2;9,4)
Số cây	5	12	25	44	14

Xét tính đúng sai của các khẳng định sau:

- Khoảng biên thiên của mẫu số liệu ghép nhóm: 1m
- Tứ phân vị thứ ba của mẫu số liệu ghép nhóm là: $Q_3 = 10,5$



c) Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = 2,06$

d) Trong 100 cây keo trên có 1 cây cao 8,4 m. Thì chiều cao của cây keo này là giá trị ngoại lệ

Câu 9. Thống kê điểm trung bình môn Toán của một số học sinh lớp 11 được cho ở bảng sau:

Khoảng điểm	[6,5;7)	[7;7,5)	[7,5;8)	[8;8,5)	[8,5;9)	[9;9,5)	[9,5;10)
Tần số	8	10	16	24	13	7	4

Xét tính đúng sai của các khẳng định sau:

a) Khoảng biến thiên của mẫu số liệu là $R = 4$

b) Số trung bình của mẫu số liệu xấp xỉ bằng: 8,12

c) Mốt của mẫu số liệu là: 6,21

d) Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = 2,05$

Câu 10. Một công ty cung cấp nước sạch thống kê lượng nước các hộ gia đình trong một khu vực tiêu thụ trong một tháng ở bảng sau:

Lượng nước tiêu thụ (m^3)	[3;6)	[6;9)	[9;12)	[12;15)	[15;18)
Số hộ gia đình	24	57	42	29	8

Xét tính đúng sai của các khẳng định sau:

a) Số trung bình của mẫu số liệu ghép nhóm đã cho là 9,375.

b) Mốt của mẫu số liệu là $M_o = 8,0625$.

c) Khoảng biến thiên của mẫu số liệu ghép nhóm đã cho là 15

d) Công ty muốn gửi một thông báo khuyến nghị tiết kiệm nước đến 25% các hộ gia đình có lượng nước tiêu thụ cao nhất. Khi đó công ty nên gửi thông báo tiết kiệm nước đến các hộ gia đình có lượng nước tiêu thụ từ $14,79m^3$ nước trở lên.

PHẦN III. Câu trắc nghiệm trả lời ngắn

Câu 1. Bảng dưới biểu diễn mẫu số liệu ghép nhóm về chiều cao (đơn vị: centimét) của 36 học sinh nam lớp 12 ở một trường trung học phổ thông. Tìm khoảng biến thiên của mẫu số liệu ghép nhóm đó.

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

Câu 2. Bảng dưới biểu diễn mẫu số liệu ghép nhóm về chiều cao của 42 mẫu cây ở một vườn thực vật (đơn vị: centimét). Tính khoảng tứ phân vị của mẫu số liệu ghép nhóm đó (làm tròn kết quả đến hàng phần mười).



Nhóm	Tần số	Tần số tích lũy
[40; 45)	5	5
[45; 50)	10	15
[50; 55)	7	22
[55; 60)	9	31
[60; 65)	7	38
[65; 70)	4	42
	$n = 42$	

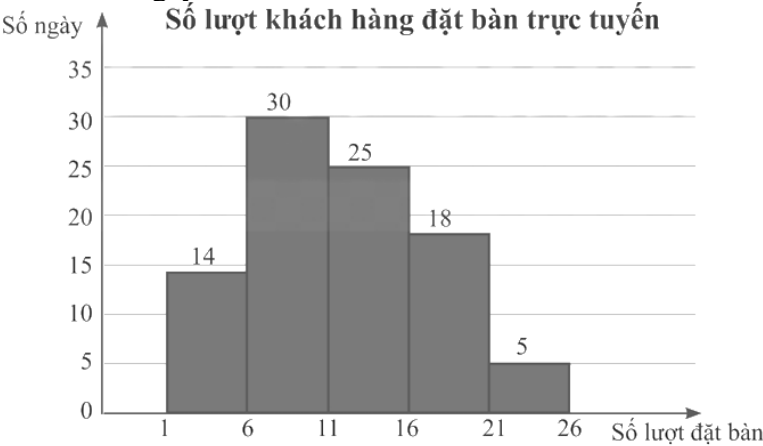
Câu 3. Bảng sau thống kê cân nặng của 50 quả xoài Thanh Ca được lựa chọn ngẫu nhiên sau khi thu hoạch ở một nông trường.

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

Khoảng tứ phân vị của mẫu số liệu ghép nhóm đó bằng $\frac{a}{b}$ với $\frac{a}{b}$ là phân số tối giản và $a, b \in \mathbb{Z}$

Tính giá trị biểu thức $T = a + b$

Câu 4. Biểu đồ dưới đây biểu diễn số lượt khách hàng đặt bàn qua hình thức trực tuyến mỗi ngày trong quý III năm 2022 của một nhà hàng. Cột thứ nhất biểu diễn số ngày có từ 1 đến dưới 6 lượt đặt bàn; cột thứ hai biểu diễn số ngày có từ 6 đến dưới 11 lượt đặt bàn;...



Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm cho bởi biểu đồ trên.

Câu 5. Thời gian luyện tập trong một ngày (tính theo giờ) của một số vận động viên được ghi lại ở bảng sau:

Thời gian luyện tập (giờ)	[0; 2)	[2; 4)	[4; 6)	[6; 8)	[8; 10)
Số vận động viên	3	8	12	12	4

Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.

Câu 6. Một người thống kê lại thời gian thực hiện các cuộc gọi điện thoại của người đó trong một tuần ở bảng sau:

Thời gian (đơn vị: giây)	[0; 60)	[60; 120)	[120; 180)	[180; 240)	[240; 300)	[300; 360)
Số cuộc gọi	8	10	7	5	2	1

Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.



Câu 7. Số điểm một cầu thủ bóng rổ ghi được trong 20 trận đấu được cho ở bảng sau:

Điểm số	$[5, 5; 10, 5)$	$[10, 5; 15, 5)$	$[15, 5; 20, 5)$	$[20, 5; 25, 5)$
Số trận	3	9	2	6

Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.

Câu 8. Thời gian sử dụng điện thoại trong một ngày của 30 sinh viên được ghi lại ở bảng sau (đơn vị: phút).

Thời gian (phút)	$[0; 60)$	$[60; 120)$	$[120; 180)$	$[180; 240)$	$[240; 300)$
Số sinh viên	2	7	7	10	4

Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.

Câu 9. Điểm thi môn Toán (thang điểm 100, điểm được làm tròn đến 1) của 60 thí sinh được cho mẫu số liệu ghép nhóm như sau:

Điểm	$[0; 9, 5)$	$[9, 5; 19, 5)$	$[19, 5; 29, 5)$	$[29, 5; 39, 5)$	$[39, 5; 49, 5)$
Số thí sinh	1	2	4	6	15
Điểm	$[49, 5; 59, 5)$	$[59, 5; 69, 5)$	$[69, 5; 79, 5)$	$[79, 5; 89, 5)$	$[89, 5; 99, 5)$
Số thí sinh	12	10	6	3	1

Tính khoảng tứ phân vị của mẫu số liệu ghép nhóm

Câu 10. Kết quả điều tra về số giờ làm thêm trong một tuần của 100 sinh viên được cho như bảng sau:

Số giờ làm thêm	$[2; 4)$	$[4; 6)$	$[6; 8)$	$[8; 10)$	$[10; 12)$
Số giờ làm thêm đại diện	3	5	7	9	11
Số sinh viên	12	20	37	21	10

Tìm khoảng tứ phân vị của số liệu đó.

-----HẾT-----

**B //** PHÂN LOẠI VÀ PHƯƠNG PHÁP GIẢI TOÁN**Dạng 2: Tìm tứ phân vị của mẫu số liệu ghép nhóm**

Với mẫu số liệu ghép nhóm:

Nhóm	$[a_1; a_2)$	...	$[a_i; a_{i+1})$	...	$[a_k; a_{k+1})$
Tần số	m_1	...	m_i	...	m_k

Các bước thực hiện:

- Tìm tứ phân vị Q_1 và Q_3 theo công thức:

$$Q_r = a_p + \frac{\frac{r \cdot n}{4} - (m_1 + \dots + m_{p-1})}{m_p} \cdot (a_{p+1} - a_p)$$

trong đó $[a_p; a_{p+1})$ là nhóm chứa tứ phân vị thứ r với $r = 1, 2, 3$ và n là cỡ mẫu.

- Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = Q_3 - Q_1$

BÀI TẬP TỰ LUẬN**Bài tập 1:** Thời gian chờ khám bệnh của các bệnh nhân tại phòng khám X được cho trong bảng sau:

Thời gian (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 tại 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?

Lời giải

- a) Cỡ mẫu là $n = 3 + 12 + 15 + 8 = 38$. Gọi $x_1, \dots, x_{38}$ là thời gian chờ khám bệnh của 38 bệnh nhân này và giả sử rằng dãy số liệu gốc này đã được sắp xếp theo thứ tự tăng dần. Tứ phân vị thứ nhất của mẫu số liệu gốc là x_{10} nên nhóm chứa tứ phân vị thứ nhất là nhóm $[5; 10)$ và ta có:

$$Q_1 = 5 + \left[\frac{\frac{38}{4} - 3}{12} \right] \cdot 5 \approx 7,71.$$

Tứ phân vị thứ ba của mẫu số liệu gốc là x_{29} nên nhóm chứa tứ phân vị thứ ba là nhóm $[10; 15)$

$$\text{và ta có: } Q_3 = 10 + \left[\frac{\frac{3 \cdot 38}{4} - 15}{15} \right] \cdot 5 = 14,5.$$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 \approx 14,5 - 7,71 = 6,79$.

- b) Do $\Delta_Q = 6,79 < 9,23$ nên thời gian chờ của bệnh nhân tại phòng khám Y phân tán hơn thời gian chờ của bệnh nhân tại phòng khám X .



Bài tập 2: Một người ghi lại thời gian đàm thoại của một số cuộc gọi cho kết quả như bảng sau:

Thời gian t (phút)	Số cuộc gọi
$0 \leq t < 1$	8
$1 \leq t < 2$	17
$2 \leq t < 3$	25
$3 \leq t < 4$	20
$4 \leq t < 5$	10

Tính khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.

Lời giải

Hiệu chỉnh lại bảng số liệu ta có:

Thời gian t (phút)	[0;1)	[1;2)	[2;3)	[3;4)	[4;5)
Số cuộc gọi	8	17	25	20	10

Cỡ mẫu $n = 80$. Giả sử $x_1, x_2, \dots, x_{80}$ là thời gian đàm thoại của 80 cuộc gọi và giả sử dãy số liệu gốc này đã được sắp xếp theo thứ tự tăng dần.

Vì $\frac{n}{4} = 20$ và $8 < 20 < 8 + 17$ nên nhóm chứa tứ phân vị thứ nhất là nhóm [1; 2) và tứ phân vị thứ

nhất là: $Q_1 = 1 + \frac{\frac{80}{4} - 8}{17} \cdot 1 = \frac{29}{17}$

vì $\frac{3n}{4} = 60$ và $8 + 17 + 25 < 60 < 8 + 17 + 25 + 20$ nên nhóm chứa tứ phân vị thứ ba là nhóm [3; 4)

và tứ phân vị thứ ba là: $Q_3 = 3 + \frac{\frac{3 \cdot 80}{4} - (8 + 17 + 25)}{20} \cdot 1 = 3,5$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $3,5 - \frac{29}{17} = \frac{61}{34}$

Bài tập 3: Thống kê số ngày trong tháng Sáu năm 2021 và năm 2022 theo nhiệt độ cao nhất trong ngày tại Hà Nội, người ta thu được bảng sau:

Nhiệt độ ($^{\circ}\text{C}$)	[28;30)	[30;32)	[32;34)	[34;36)	[36;38)	[38;40)
Số ngày trong tháng 6/2021	0	2	8	5	6	9
Số ngày trong tháng 6/2022	2	3	4	11	8	2

Hỏi tháng sáu năm nào ở Hà Nội nhiệt độ cao nhất trong ngày biến đổi nhiều hơn?

Lời giải

Năm 2021: Khoảng biến thiên của nhiệt độ là: $R_1 = 40 - 30 = 10$

Cỡ mẫu $n = 30$. Giả sử $y_1, y_2, \dots, y_{30}$ là nhiệt độ cao nhất trong ngày của 30 ngày tháng Sáu năm 2021 và giả sử dãy số liệu gốc này đã được sắp xếp theo thứ tự không giảm.



Vì $\frac{n}{4} = \frac{30}{4} = 7,5$ và $2 < 7,5 < 2+8$ nên tứ phân vị thứ nhất thuộc nhóm $[32;34)$ và tứ phân vị

thứ nhất là: $Q_1 = 32 + \frac{\frac{30}{4} - 2}{8} \cdot 2 = 33,375$

vì $\frac{3n}{4} = \frac{3 \cdot 30}{4} = 22,5$ và $2+8+5+6 < 22,5 < 2+8+5+6+9$ nên tứ phân vị thứ ba thuộc nhóm

$[38;40)$ và tứ phân vị thứ ba là: $Q_3 = 38 + \frac{\frac{3 \cdot 30}{4} - (2+8+5+6)}{9} \cdot 2 = \frac{115}{3}$

Khoảng biến thiên của mẫu số liệu ghép nhóm là: $\Delta_{Q_1} = \frac{115}{3} - 33,375 = \frac{119}{24}$

Năm 2022: Khoảng biến thiên của nhiệt độ là: $R_2 = 40 - 28 = 12$

Cỡ mẫu $n = 30$. Giả sử $z_1, z_2, \dots, z_{30}$ là nhiệt độ cao nhất trong ngày của 30 ngày tháng Sáu năm 2022 và giả sử dãy số liệu gốc này đã được sắp xếp theo thứ tự tăng dần.

vì $\frac{n}{4} = \frac{30}{4} = 7,5$ và $2+3 < 7,5 < 2+3+4$ nên nhóm chứa tứ phân vị thứ nhất là nhóm $[32;34)$

và tứ phân vị thứ nhất là: $Q'_1 = 32 + \frac{\frac{30}{4} - (2+3)}{4} \cdot (34 - 32) = 33,25$

vì $\frac{3n}{4} = \frac{3 \cdot 30}{4} = 22,5$ và $2+3+4+11 < 22,5 < 2+3+4+11+8$ nên nhóm chứa tứ phân vị thứ ba

là nhóm $[36;38)$ và tứ phân vị thứ ba là: $Q'_3 = 36 + \frac{\frac{3 \cdot 30}{4} - (2+3+4+11)}{8} \cdot (38 - 36) = 36,625$

Khoảng biến thiên của mẫu số liệu ghép nhóm là: $\Delta_{Q_2} = 36,625 - 33,25 = 3,375$

Theo khoảng biến thiên: vì $R_2 > R_1$ nên nhiệt độ cao nhất trong ngày vào tháng 6 năm 2022 biến đổi nhiều hơn nhiệt độ cao nhất trong ngày vào tháng 6 năm 2021.

Theo khoảng tứ phân vị: vì $\Delta_{Q_1} > \Delta_{Q_2}$ nên nhiệt độ cao nhất trong ngày vào tháng 6 năm 2021 biến đổi nhiều hơn nhiệt độ cao nhất trong ngày vào tháng 6 năm 2022.

Bài tập 4: Thống kê số thẻ vàng của mỗi câu 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?



Lời giải

a) Bảng số liệu ghép nhóm:

Số thẻ	[40;50)	[50;60)	[60;70)	[70;80)	[80;90)	[90;100)	[100;110)
Tần số	2	5	7	5	0	0	1

b) Với mẫu số liệu gốc: Khoảng biến thiên là: $R_1 = 101 - 42 = 59$

Sắp xếp mẫu số liệu theo thứ tự không giảm là:

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

Vì $n = 20$ nên tứ phân vị thứ nhất là trung vị của dãy số liệu: 42; 47; 50; 55; 55; 57; 59; 60; 61; 63.

$$\text{Do đó, } Q_1 = \frac{55 + 57}{2} = 56$$

Tứ phân vị thứ ba là trung vị của dãy số liệu: 63; 67; 67; 68; 73; 75; 78; 79; 79; 101.

$$\text{Do đó, } Q_3 = \frac{73 + 75}{2} = 74.$$

Khoảng tứ phân vị là: $\Delta_{Q_1} = 74 - 56 = 18$ Với mẫu số liệu ghép nhóm: Khoảng biến thiên là: $R_2 = 110 - 40 = 70$ Cỡ mẫu $n = 20$. Giả sử $x_1, x_2, \dots, x_{20}$ là số thẻ vàng mà mỗi câu lạc bộ ngoại hạng Anh nhận được mùa giải 2021-2022, các giá trị này đã được sắp xếp theo thứ tự không giảm.Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{x_5 + x_6}{2}$ nên nhóm chứa tứ phân vị thứ nhất là nhóm

$$[50; 60) \text{ và ta có: } Q'_1 = 50 + \frac{\frac{20}{4} - 2}{5} \cdot 10 = 56$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{x_{15} + x_{16}}{2}$ nên nhóm chứa tứ phân vị thứ ba là nhóm

$$[70; 80) \text{ và ta có: } Q'_3 = 70 + \frac{\frac{3 \cdot 20}{4} - (2 + 5 + 7)}{5} \cdot 10 = 72$$

Khoảng biến thiên của mẫu số liệu ghép nhóm là: $\Delta_{Q_2} = 72 - 56 = 16$ Giá trị chính xác là $R_1; \Delta Q_1$, giá trị xấp xỉ là $R_2; \Delta Q_2$ **Bài tập 5:** 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.

Lời giải

Ta có bảng số liệu với giá trị đại diện của nhóm là:

Thu nhập	[5; 8)	[8; 11)	[11; 14)	[14; 17)	[17; 20)
Giá trị đại diện	6,5	9,5	12,5	15,5	18,5
Số người của nhà máy A	20	35	45	35	20
Số người của nhà máy B	17	23	30	23	17

Mức thu nhập trung bình của người lao động nhà máy A là:

$$\frac{6,5 \cdot 20 + 9,5 \cdot 35 + 12,5 \cdot 45 + 15,5 \cdot 35 + 18,5 \cdot 20}{20 + 35 + 45 + 35 + 20} = \frac{25}{2} \text{ (triệu đồng)}$$

Mức thu nhập trung bình của người lao động nhà máy B là:

$$\frac{6,5 \cdot 17 + 9,5 \cdot 23 + 12,5 \cdot 30 + 15,5 \cdot 23 + 18,5 \cdot 17}{17 + 23 + 30 + 23 + 17} = \frac{25}{2} \text{ (triệu đồng)}$$

Nhà máy A: Ta có cỡ mẫu $n = 155$. Giả sử $x_1, x_2, \dots, x_{155}$ là mức thu nhập của người lao động nhà máy A và giả sử dãy số liệu gốc này đã được sắp xếp theo thứ tự không giảm.

Vì $\frac{n}{4} = 38,75$ và $20 < 38,75 < 20 + 35$ nên nhóm chứa tứ phân vị thứ nhất là nhóm [8;11) và tứ

$$\text{phân vị thứ nhất là: } Q_1 = 8 + \frac{\frac{155}{4} - 20}{35} \cdot 3 = \frac{269}{28}$$

Vì $\frac{3n}{4} = 116,25$ và $20 + 35 + 45 < 116,25 < 20 + 35 + 45 + 35$ nên nhóm chứa tứ phân vị thứ ba là

$$\text{nhóm [14;17) và tứ phân vị thứ ba là: } Q_3 = 14 + \frac{\frac{3 \cdot 155}{4} - (20 + 35 + 45)}{35} \cdot 3 = \frac{431}{28}$$

$$\text{Khoảng biến thiên của mẫu số liệu ghép nhóm là: } \Delta_{Q_1} = \frac{431}{28} - \frac{269}{28} = \frac{81}{14}$$

Nhà máy B: Ta có cỡ mẫu $n = 110$. Giả sử $x_1, x_2, \dots, x_{110}$ là mức thu nhập của người lao động nhà máy B và giả sử dãy số liệu gốc này đã được sắp xếp theo thứ tự không giảm.

Vì $\frac{n}{4} = 27,5$ và $17 < 27,5 < 17 + 23$ nên nhóm chứa tứ phân vị thứ nhất là nhóm [8;11) và tứ

$$\text{phân vị thứ nhất là: } Q_1' = 8 + \frac{\frac{110}{4} - 17}{23} \cdot 3 = \frac{431}{46}$$



Vì $\frac{3n}{4} = 82,5$ và $17 + 23 + 30 < 82,5 < 17 + 23 + 30 + 23$ nên nhóm chứa tứ phân vị thứ ba là nhóm

$$[14;17) \text{ và tứ phân vị thứ ba là: } Q'_3 = 14 + \frac{\frac{3 \cdot 110}{4} - (17 + 23 + 30)}{23} \cdot 3 = \frac{719}{46}$$

$$\text{Khoảng biến thiên của mẫu số liệu ghép nhóm là: } \Delta_{Q_2} = \frac{719}{46} - \frac{431}{46} = \frac{144}{23}$$

Vì $\Delta_{Q_1} < \Delta_{Q_2}$ nên mức thu nhập của người lao động nhà máy B biến động nhiều hơn.

Bài tập 6: Bảng sau đây cho biết chiều cao của các học sinh lớp 12A và 12B.

Thu nhập	[145;150)	[150;155)	[155;160)	[160;165)	[165;170)	[170;175)
Số học sinh lớp 12A	1	0	15	12	10	5
Số học sinh 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, 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?

Lời giải

a) Lớp 12A: Khoảng biến thiên: $R = 175 - 145 = 30$

Ta có cỡ mẫu $n = 43$. Giả sử $x_1, x_2, \dots, x_{43}$ là chiều cao của các học sinh lớp 12A và giả sử dãy số liệu gốc này đã được sắp xếp theo thứ tự không giảm.

vì $\frac{n}{4} = 10,75$ và $1 < 10,75 < 1 + 15$ nên nhóm chứa tứ phân vị thứ nhất là nhóm $[150;160)$ và tứ

$$\text{phân vị thứ nhất là: } Q_1 = 155 + \frac{\frac{43}{4} - 1}{15} \cdot 5 = 158,25$$

Vì $\frac{3n}{4} = 32,25$ và $1 + 15 + 12 < 32,25 < 1 + 15 + 12 + 10$ nên nhóm chứa tứ phân vị thứ ba là nhóm

$$[165;170) \text{ và tứ phân vị thứ ba là: } Q_3 = 165 + \frac{\frac{3 \cdot 43}{4} - (1 + 15 + 12)}{10} \cdot 5 = 167,125$$

$$\text{Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: } \Delta_{Q_1} = 167,125 - 158,25 = 8,875$$

Lớp 12B: Khoảng biến thiên: $R = 175 - 155 = 20$

Ta có cỡ mẫu $n = 42$. Giả sử $x_1, x_2, \dots, x_{42}$ là chiều cao của các học sinh lớp 12B và giả sử dãy số liệu gốc này đã được sắp xếp theo thứ tự không giảm.



Vì $\frac{n}{4} = 10,5$ và $0 < 10,5 < 17$ nên nhóm chứa tứ phân vị thứ nhất là nhóm $[155; 160)$ và ta có:

$$Q'_1 = 155 + \frac{\frac{42}{4} - 0}{17} \cdot 5 = \frac{5375}{34}$$

vì $\frac{3n}{4} = 31,5$ và $17 + 10 < 31,5 < 17 + 10 + 9$ nên nhóm chứa tứ phân vị thứ ba là nhóm $[165; 170)$

$$\text{và tứ phân vị thứ ba là } Q'_3 = 165 + \frac{\frac{3 \cdot 42}{4} - (17 + 10)}{9} \cdot 5 = \frac{335}{2}$$

$$\text{Khoảng biến thiên của mẫu số liệu ghép nhóm là: } \Delta_{Q_2} = \frac{335}{2} - \frac{5375}{34} = \frac{160}{17}$$

b) Để so sánh độ phân tán về chiều cao của học sinh hai lớp này, ta nên dùng khoảng tứ phân vị vì khoảng tứ phân vị chỉ phụ thuộc vào nửa giữa của mẫu số liệu, không bị ảnh hưởng bởi các giá trị bất thường.

Bài tập 7: Điểm kiểm tra cuối khoá môn Tiếng Anh của hai lớp ở một trung tâm ngoại ngữ được thống kê trong các bảng a) và b)

Bảng a). 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 b). 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.

c) Mẫu số liệu nào có độ phân tán lớn hơn? Minh hoạ câu trả lời bằng cách biểu diễn các tứ phân vị và khoảng tứ phân vị của mỗi mẫu số liệu trên trục số.

Lời giải

a) Hai mẫu số liệu đều có khoảng biến thiên là $R = 100 - 50 = 50$ nên không thể căn cứ vào đó để nói điểm của lớp nào đồng đều hơn.

b) Kích thước của hai mẫu số liệu đều là $N = 100$. Ta có $\frac{N}{4} = 25$; $\frac{N}{2} = 50$; $\frac{3N}{4} = 75$.



Đối với mẫu số liệu về điểm của lớp A, ta tìm các tứ phân vị Q_1^A, Q_2^A, Q_3^A và khoảng tứ phân vị Δ_Q^A qua bảng tần số tích lũy dưới đây:

Điểm	Tần số	Tần số tích lũy
[50;60)	8	8
[60;70)	20	28
[70;80)	50	78
[80;90)	17	95
[90;100)	5	100

Nhóm chứa Q_1^A là [60;70). $Q_1^A = 60 + \frac{25-8}{20} \cdot 10 = 68,5$.

Nhóm chứa Q_2^A là [70;80). $Q_2^A = 70 + \frac{50-28}{50} \cdot 10 = 74,4$.

Nhóm chứa Q_3^A là [70;80). $Q_3^A = 70 + \frac{75-28}{50} \cdot 10 = 79,4$.

Vậy $\Delta_Q^A = 79,4 - 68,5 = 10,9$.

Gọi Q_1^B, Q_2^B, Q_3^B là các tứ phân vị và Δ_Q^B là khoảng tứ phân vị của mẫu số liệu về điểm của lớp B. Ta lập bảng tần số tích lũy và tính được:

Điểm	Tần số	Tần số tích lũy
[50;60)	15	15
[60;70)	20	35
[70;80)	30	65
[80;90)	20	85
[90;100)	15	100

Nhóm chứa Q_1^B là [60;70). $Q_1^B = 60 + \frac{25-15}{20} \cdot 10 = 65$.

Nhóm chứa Q_2^B là [70;80). $Q_2^B = 70 + \frac{50-35}{30} \cdot 10 = 75$.

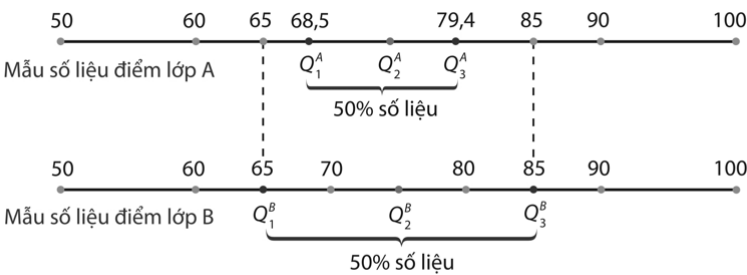
Nhóm chứa Q_3^B là [80;90). $Q_3^B = 80 + \frac{75-65}{20} \cdot 10 = 85$.

Vậy $\Delta_Q^B = 85 - 65 = 20$.

c) $\Delta_Q^B > \Delta_Q^A$ nên điểm của lớp B phân tán hơn điểm của lớp A. Minh họa trên trục số: Mỗi mẫu đều có 100 số liệu thuộc khoảng [50; 100). Có 50% số liệu ở giữa của bảng điểm lớp B thuộc

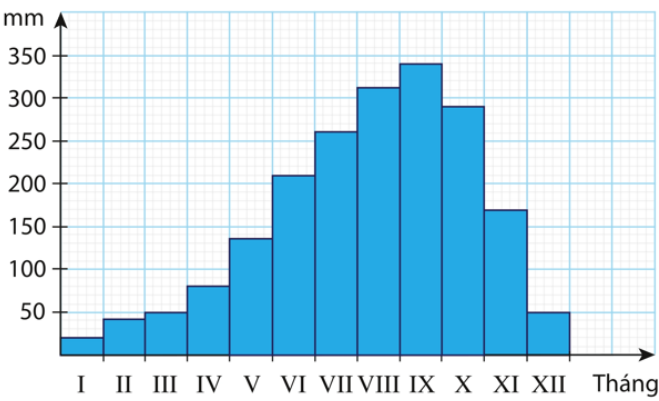


khoảng (Q_1^B, Q_3^B) . Bảng điểm lớp A cũng có 50% số liệu ở giữa thuộc khoảng (Q_1^A, Q_3^A) . Vì $(Q_1^A, Q_3^A) \subset (Q_1^B, Q_3^B)$ nên ta có thể nói là điểm của lớp B phân tán hơn so với điểm lớp A.



Hình 3.1

Bài tập 8: Hình dưới 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.

Lời 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) $N = 12$ nên $\frac{N}{4} = 3; \frac{3N}{4} = 9$.



Nhóm chứa Q_1 là $[50;100)$. $Q_1 = 50 + \frac{3-2}{3} \cdot 50 \approx 67$.

Nhóm chứa Q_3 là $[250;300)$. $Q_3 = 250 + \frac{9-8}{2} \cdot 50 = 275$. Vậy $\Delta_Q = 275 - 67 = 208$.

Kết quả tìm được cho thấy: Hằng 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 tập 9: Hằng ngày ông Thắng đều đi xe buýt từ nhà đến cơ quan. Dưới đây là bản 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 lần	22	38	27	8	4	1

a) Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm phía trên (Làm tròn kết quả đến hàng phần trăm)

b) Biết rằng trong 100 lần đi trên, chỉ có đúng một lần ông Thắng đi hết 32 phút. Thời gian của lần đi đó có phải giá trị ngoại lệ không?

Lời giải

a) Cỡ mẫu $n = 100$. Gọi $x_1, x_2, x_3, \dots, x_{100}$ là mẫu số liệu gốc gồm 100 lần đi xe buýt của ông Thắng. Khi đó ta có:

$$\begin{aligned} x_1, \dots, x_{22} &\in [15;18) & x_{23}, \dots, x_{60} &\in [18;21) & x_{61}, \dots, x_{87} &\in [21;24) \\ x_{88}, \dots, x_{95} &\in [21;27) & x_{96}, \dots, x_{99} &\in [27;30) & x_{100} &\in [30;33) \end{aligned}$$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_{75} + x_{76}) \in [18;21)$

Do đó tứ phân vị thứ nhất của mẫu số 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 ghép nhóm là:

$$Q_3 = 18 + \frac{\frac{3 \cdot 100}{4} - (22 + 38)}{27}(24 - 21) = \frac{68}{3}$$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = \frac{68}{3} - \frac{693}{38} = \frac{505}{114} \approx 4,43$.

b) Trong lần duy nhất ông Thắng đi hết 32 phút, thời gian đi của ông thuộc nhóm $[30;33)$

Vì $Q_3 + 1,5\Delta_Q \approx 29,31 < 30$ nên đây là giá trị ngoại lệ của mẫu số liệu ghép nhóm.

**BÀI TẬP TRẮC NGHIỆM**

PHẦN I. Câu trắc nghiệm nhiều phương án lựa chọn. Mỗi câu hỏi thí sinh chỉ chọn một phương án.

Câu 1. Một vườn thú ghi lại tuổi thọ (đơn vị: năm) của 20 con hổ và thu được kết quả như sau:

Tuổi thọ	[14;15)	[15;16)	[16;17)	[17;18)	[18;19)
Số con hổ	1	3	8	6	2

Nhóm chứa tứ phân vị thứ nhất là

- A. [14;15). B. [15;16). C. [16;17). D. [17;18).

Lời giải

Ta có: $\frac{n}{4} = \frac{20}{4} = 5$ và $1+3 < 5 < 1+3+8$ nên tứ phân vị thứ nhất của mẫu số liệu thuộc nhóm [16;17)

Câu 2. Một vườn thú ghi lại tuổi thọ (đơn vị: năm) của 20 con hổ và thu được kết quả như sau:

Tuổi thọ	[14;15)	[15;16)	[16;17)	[17;18)	[18;19)
Số con hổ	1	3	8	6	2

Nhóm chứa tứ phân vị thứ ba là

- A. [15;16). B. [16;17). C. [17;18). D. [18;19).

Lời giải

Ta có: $\frac{3.20}{4} = 15$ và $1+3+8 < 15 < 1+3+8+6$ tứ phân vị thứ ba thuộc nhóm [17;18)

Câu 3. Khảo sát thời gian tập thể dục trong ngày của một số học sinh khối 11 thu được mẫu số liệu ghép nhóm sau:

Thời gian (phút)	[0; 20)	[20; 40)	[40; 60)	[60; 80)	[80; 100)
Số học sinh	5	9	12	10	6

Nhóm chứa tứ phân vị thứ nhất là

- A. [0; 20). B. [20; 40). C. [40; 60). D. [60; 80).

Lời giải

Ta có cỡ mẫu là $n = 5 + 9 + 12 + 10 + 6 = 42$.

Gọi $x_1, x_2, \dots, x_{42}$ là thời gian tập thể dục trong ngày của 42 học sinh khối 11 và giả sử dãy này đã sắp xếp theo thứ tự tăng dần.

Khi đó tứ phân vị thứ nhất Q_1 là trung vị của dãy gồm 21 số liệu đầu nên $Q_1 = x_{11}$. Do x_{11} thuộc nhóm [20; 40) nên nhóm này chứa Q_1 .



- Câu 4.** Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một cửa hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

Doanh thu	[5; 7)	[7; 9)	[9; 11)	[11; 13)	[13; 15)
Số ngày	2	7	7	3	1

Tứ phân vị thứ nhất của mẫu số liệu trên gần nhất với giá trị nào trong các giá trị sau?

- A. 7. **B. 7,6.** C. 8. D. 8,6.

Lời giải

Tứ phân vị thứ nhất của dãy số liệu là $\frac{1}{2}(x_4 + x_5)$ thuộc nhóm [7; 9) nên tứ phân vị thứ nhất của

$$\text{mẫu số liệu là } Q_1 = 7 + \frac{\frac{20}{4} - 2}{7}(9 - 7) = 7,86$$

- Câu 5.** Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một cửa hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

Doanh thu	[5; 7)	[7; 9)	[9; 11)	[11; 13)	[13; 15)
Số ngày	2	7	7	3	1

Tứ phân vị thứ ba của mẫu số liệu trên gần nhất với giá trị nào trong các giá trị sau?

- A. 10. B. 11. C. 12. D. 13.

Lời giải

Tứ phân vị thứ ba của dãy số liệu là $\frac{1}{2}(x_{15} + x_{16})$ thuộc nhóm [9; 11) nên tứ phân vị thứ ba của

$$\text{mẫu số liệu là } Q_3 = 9 + \frac{\frac{3 \cdot 20}{4} - 9}{7}(11 - 9) = 10,7$$

- Câu 6.** Cho mẫu số liệu ghép nhóm về tuổi thọ (đơn vị tính là năm) của một loại bóng đèn mới như sau.

Tuổi thọ	[2; 3, 5)	[3, 5; 5)	[5; 6, 5)	[6, 5; 8)
Số bóng đèn	8	22	35	15

Nhóm chứa tứ phân vị thứ nhất của mẫu số liệu là

- A. [2; 3, 5). **B. [3, 5; 5).** C. [5; 6, 5). D. [6, 5; 8).

Lời giải

Nhóm chứa tứ phân vị thứ nhất của mẫu số liệu là [3, 5; 5)

- Câu 7.** Cho mẫu số liệu ghép nhóm về tuổi thọ (đơn vị tính là năm) của một loại bóng đèn mới như sau.

Tuổi thọ	[2; 3, 5)	[3, 5; 5)	[5; 6, 5)	[6, 5; 8)
Số bóng đèn	8	22	35	15

Nhóm chứa tứ phân vị thứ ba của mẫu số liệu là

- A. [2; 3, 5). B. [3, 5; 5). **C. [5; 6, 5).** D. [6, 5; 8).



Lời giải

Nhóm chứa tứ phân vị thứ ba của mẫu số liệu là [5;6,5)

Câu 8. Cho bảng số liệu về chiều cao của 100 học sinh một trường trung học phổ thông dưới đây.

Nhóm	Chiều cao (cm)	Số học sinh
1	[150;153)	7
2	[153;156)	13
3	[156;159)	40
4	[159;162)	21
5	[162;165)	13
6	[165;168)	6

Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm trên (làm tròn kết quả đến hàng phần trăm) là

- A. 156,25. B. 157,5. C. 156,38. D. 157,54.

Lời giải

Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm trên là 156,38.

Câu 9. Cho bảng số liệu về chiều cao của 100 học sinh một trường trung học phổ thông dưới đây.

Nhóm	Chiều cao (cm)	Số học sinh
1	[150;153)	7
2	[153;156)	13
3	[156;159)	40
4	[159;162)	21
5	[162;165)	13
6	[165;168)	6

Tứ phân vị thứ ba của mẫu số liệu ghép nhóm trên (làm tròn kết quả đến hàng phần trăm) là:

- A. 160,52. B. 161,52. C. 161,14. D. 162,25.

Lời giải

Tứ phân vị thứ ba của mẫu số liệu ghép nhóm trên là 161,14

Câu 10. Bảng dưới biểu diễn mẫu số liệu ghép nhóm về số tiền (đơn vị: nghìn đồng) mà 60 khách hàng mua sách ở một cửa hàng trong một ngày.

Nhóm	Tần số
[40;50)	3
[50;60)	6

$[60; 70)$	19
$[70; 80)$	23
$[80; 90)$	9
	$n = 60$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là:

- A. 50. B. 40. C. 14,23. D. 70,87.

Lời giải

Số phân tử của mẫu là $n = 42$

Tần số tích lũy của các nhóm lần lượt là $cf_1 = 3, cf_2 = 9, cf_3 = 28, cf_4 = 51, cf_5 = 60$

Ta có: $\frac{n}{4} = \frac{60}{4} = 15$ mà $9 < 15 < 28$ suy ra nhóm 3 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 15. Xét nhóm 3 là nhóm $[60; 70]$ có $s = 60, h = 10, n_3 = 19$ và nhóm 2 là nhóm $[50; 60]$ có $cf_2 = 9$

Ta có tứ phân vị thứ nhất là: $Q_1 = s + \left(\frac{15 - cf_2}{n_3} \right) \cdot h = 60 + \left(\frac{15 - 9}{19} \right) \cdot 10 = \frac{1200}{19}$

Ta có: $\frac{3n}{4} = \frac{3 \cdot 60}{4} = 45$ mà $28 < 45 < 51$ suy ra nhóm 4 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 45. Xét nhóm 4 là nhóm $[70; 80]$ có $t = 70, l = 10, n_4 = 23$ và nhóm 3 là nhóm $[60; 70]$ có $cf_3 = 28$

Ta có tứ phân vị thứ ba là: $Q_3 = t + \left(\frac{45 - cf_3}{n_4} \right) \cdot l = 70 + \left(\frac{45 - 28}{23} \right) \cdot 10 = \frac{1780}{23}$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $Q_3 - Q_1 = \frac{1780}{23} - \frac{1200}{19} \approx 14,23$

Câu 11. Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khỏe. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

Quãng đường (km)	$[2, 7; 3, 0)$	$[3, 0; 3, 3)$	$[3, 3; 3, 6)$	$[3, 6; 3, 9)$	$[3, 9; 4, 2)$
Số ngày	3	6	5	4	2

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

- A. 0,9. B. 0,975. C. 0,5. D. 0,575.

Lời giải

Cỡ mẫu $n = 20$. Gọi $x_1; x_2; \dots; x_{20}$ là mẫu số liệu gốc về quãng đường đi bộ mỗi ngày của bác Hương trong 20 ngày được xếp theo thứ tự không giảm.

Ta có: $x_1; \dots; x_3 \in [2, 7; 3, 0); x_4; \dots; x_9 \in [3, 0; 3, 3); x_{10}; \dots; x_{14} \in [3, 3; 3, 6)$



$x_{15}; \dots; x_{18} \in [3, 6; 3, 9); \quad x_{19}; x_{20} \in [3, 9; 4, 2).$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_5 + x_6) \in [3, 0; 3, 3).$

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là: $Q_1 = 3, 0 + \frac{\frac{20}{4} - 3}{6}(3, 3 - 3, 0) = 3, 1$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{15} + x_{16}) \in [3, 6; 3, 9).$

Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là:

$$Q_3 = 3, 6 + \frac{\frac{3.20}{4} - (3 + 6 + 5)}{4}(3, 9 - 3, 6) = 3, 675$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 0, 575$

Câu 12. Bạn Chi rất thích nhảy hiện đại. Thời gian tập nhảy mỗi ngày trong thời gian gần đây của bạn Chi được thống kê lại ở bảng sau:

Thời gian (phút)	[20;25)	[25;30)	[30;35)	[35;40)	[40;45)
Số ngày	6	6	4	1	1

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

- A. 23,75. B. 27,5. C. 31,88. **D. 8,125.**

Lời giải

Cỡ mẫu $n = 18$

Gọi $x_1; x_2; \dots; x_{18}$ là mẫu số liệu gốc về thời gian tập nhảy mỗi ngày của bạn Chi được xếp theo thứ tự không giảm.

Ta có: $x_1; \dots; x_6 \in [20; 25); x_7; \dots; x_{12} \in [25; 30); x_{13}; \dots; x_{16} \in [30; 35); x_{17}; \dots; x_{18} \in [35; 40); x_{18} \in [40; 45)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $x_5 \in [20; 25)$. Do đó, tứ phân vị thứ nhất của mẫu số

liệu ghép nhóm là: $Q_1 = 20 + \frac{\frac{18}{4} - 6}{6}(25 - 20) = 23, 75$

Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{14} \in [30; 35)$. Do đó, tứ phân vị thứ ba của mẫu số liệu

ghép nhóm là: $Q_3 = 30 + \frac{\frac{3.18}{4} - (6 + 6)}{4}(35 - 30) = 31, 875$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 8, 125$



Câu 13. Dũng là học sinh rất giỏi chơi rubik, bạn có thể giải nhiều loại khối rubik khác nhau. Trong một lần tập luyện giải khối rubik 3×3 , bạn Dũng đã tự thống kê lại thời gian giải rubik trong 25 lần giải liên tiếp ở bảng sau:

Thời gian giải rubik (giây)	[8;10)	[10;12)	[12;14)	[14;16)	[16;18)
Số lần	4	6	8	4	3

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

- A. 10,75. B. 1,75. C. 3,63. D. 14,38.

Lời giải

Cỡ mẫu $n = 25$

Gọi $x_1; x_2; \dots; x_{25}$ là mẫu số liệu gốc về thời gian giải rubik trong 25 lần của bạn Dũng được xếp theo thứ tự không giảm.

Ta có: $x_1; \dots; x_4 \in [8;10); x_5; \dots; x_{10} \in [10;12); x_{11}; \dots; x_{18} \in [12;14); x_{19}; \dots; x_{22} \in [14;16);$

$x_{23}; \dots; x_{25} \in [16;18)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_6 + x_7) \in [10;12)$. Do đó, tứ phân vị thứ nhất của

mẫu số liệu ghép nhóm là: $Q_1 = 10 + \frac{\frac{25}{2} - 4}{6}(12 - 10) = 10,75$

Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{19} \in [14;16)$. Do đó, tứ phân vị thứ ba của mẫu số liệu

ghép nhóm là: $Q_3 = 14 + \frac{\frac{3 \cdot 25}{4} - (4 + 6 + 8)}{4}(16 - 14) = 14,375$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 3,63$

Câu 14. Thời gian (phút) truy cập Internet mỗi buổi tối của một số học sinh được cho trong bảng sau:

Thời gian (phút)	[9, 5; 12, 5)	[12, 5; 15, 5)	[15, 5; 18, 5)	[18, 5; 21, 5)	[21, 5; 24, 5)
Số học sinh	3	12	15	24	2

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

- A. 10,75. B. 4,75. C. 4,63. D. 4,38.

Lời giải

Cỡ mẫu là $n = 56$.

Tứ phân vị thứ nhất Q_1 là $\frac{x_{14} + x_{15}}{2}$. Do x_{14}, x_{15} đều thuộc nhóm $[12,5;15,5)$ nên nhóm này chứa

Q_1 . Do đó, $p = 2; a_2 = 12,5; m_2 = 12; m_1 = 3; a_3 - a_2 = 3$ và ta có $Q_1 = 12,5 + \frac{\frac{56}{4} - 3}{12} \cdot 3 = 15,25$



Với tứ phân vị thứ ba Q_3 là $\frac{x_{42} + x_{43}}{2}$. Do x_{42}, x_{43} đều thuộc nhóm $[18,5;21,5)$ nên nhóm này chứa Q_3 . Do đó, $p = 4; a_4 = 18,5; m_4 = 24; m_1 + m_2 + m_3 = 3 + 12 + 15 = 30; a_5 - a_4 = 3$ và ta có

$$Q_3 = 18,5 + \frac{\frac{3.56}{4} - 30}{24} \cdot 3 = 20.$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 4,75$.

Câu 15. Ghi lại tốc độ bóng trong 200 lần giao bóng của một vận động viên môn quần vợt cho kết quả như bảng bên.

Tốc độ v (km/h)	Số lần
$150 \leq v < 155$	18
$155 \leq v < 160$	28
$160 \leq v < 165$	35
$165 \leq v < 170$	43
$170 \leq v < 175$	41
$175 \leq v < 180$	35

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

- A. 12,6. B. 11,5. C. 14,3. D. 16,8.

Lời giải

Cỡ mẫu là $n = 200$.

Tứ phân vị thứ nhất Q_1 là $\frac{x_{50} + x_{51}}{2}$. Do x_{50}, x_{51} đều thuộc nhóm $[160;165)$ nên nhóm này chứa Q_1 . Do đó, $p = 3; a_3 = 160; m_3 = 35; m_1 + m_2 = 18 + 28 = 46; a_4 - a_3 = 165 - 160 = 5$ và ta có

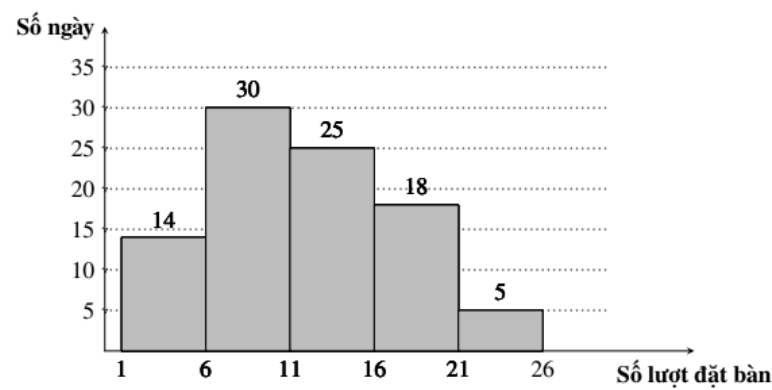
$$Q_1 = 160 + \frac{\frac{200}{4} - 46}{35} \cdot 5 \approx 160,57$$

Tứ phân vị thứ ba Q_3 là $\frac{x_{150} + x_{151}}{2}$. Do x_{150}, x_{151} đều thuộc nhóm $[170;175)$ nên nhóm này chứa Q_3 . Do đó, $p = 5; a_5 = 170;$
 $m_5 = 41; m_1 + m_2 + m_3 + m_4 = 18 + 28 + 35 + 43 = 124; a_6 - a_5 = 175 - 170 = 5$ và ta có

$$Q_3 = 170 + \frac{\frac{3.200}{4} - 124}{41} \cdot 5 \approx 173,17$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 \approx 12,6$

Câu 16. Biểu đồ dưới đây biểu diễn số lượng khách hàng đặt bàn qua hình thức trực tuyến mỗi ngày trong quý III năm 2022 của một nhà hàng. Cột thứ nhất biểu diễn số ngày có từ 1 đến 6 lượt đặt bàn, cột thứ hai biểu diễn số ngày có từ 6 đến 11 lượt đặt bàn;...



Khoảng tứ phân vị của mẫu số liệu ghép nhóm cho bởi biểu đồ trên

A. 9,5. B. 8,5. C. 10,5. D. 7,5.

Lời giải

Dựa vào biểu đồ ta lập được bảng ghép nhóm như sau:

Lượt đặt bàn	[1; 6)	[6; 11)	[11; 16)	[16; 21)	[21; 26)
Số ngày	14	30	25	18	5

Cỡ mẫu $n = 92$ và gọi $x_1, x_2, \dots, x_{92}$ là mẫu số liệu đã cho

Ta có:

$$x_1, \dots, x_{14} \in [1; 6)$$

$$x_{45}, \dots, x_{69} \in [11; 16)$$

$$x_{70}, \dots, x_{87} \in [16; 21)$$

$$x_{88}, \dots, x_{92} \in [21; 26)$$

Tứ phân vị thứ nhất của mẫu số liệu là $\frac{x_{23} + x_{24}}{2} \in [6; 11)$. Do đó tứ phân vị thứ nhất của mẫu số liệu là $Q_1 = 6 + \frac{\frac{92}{4} - 14}{30}(11 - 6) = 7,5$

Tứ phân vị thứ ba của mẫu số liệu là $\frac{x_{69} + x_{70}}{2} \in [11; 16)$ và $x_{70} \in [16; 21)$. Do đó tứ phân vị thứ ba của mẫu số liệu là $Q_3 = 16$.

Vậy khoảng tứ phân vị của mẫu số liệu là: $\Delta_Q = Q_3 - Q_1 = 16 - 7,5 = 8,5$.

PHẦN II. Câu trắc nghiệm đúng sai. Trong mỗi ý a), b), c), d) ở mỗi câu, thí sinh chọn đúng hoặc sai.

Câu 1. 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) \ Lớp	[155;160)	[160;165)	[165;170)	[170;175)	[175;180)	[180;185)
12C	2	7	12	3	0	1
12D	5	9	8	2	1	0



Xét tính đúng sai của các khẳng định sau:

- a) Chiều cao cao nhất của các bạn học sinh trong lớp 12D là 185 (cm)
- b) Khoảng biến thiên của chiều cao của các bạn học sinh nữ lớp 12C là: 30 (cm)
- c) Khoảng biến thiên của chiều cao của các bạn học sinh nữ lớp 12D là: 25 (cm)
- d) Chiều cao của học sinh lớp 12C có độ phân tán bé hơn

Lời giải

Khoảng biến thiên của chiều cao của các bạn học sinh nữ lớp 12C là: $185 - 155 = 30$ (cm)

Khoảng biến thiên của chiều cao của các bạn học sinh nữ lớp 12D là: $180 - 155 = 25$ (cm)

Vậy nếu so sánh theo khoảng biến thiên thì chiều cao của học sinh lớp 12C có độ phân tán lớn hơn

- a) Sai: Chiều cao cao nhất của các bạn học sinh trong lớp 12D là 180 (cm)
- b) Đúng: Khoảng biến thiên của chiều cao của các bạn học sinh nữ lớp 12C là: 30 (cm)
- c) Đúng: Khoảng biến thiên của chiều cao của các bạn học sinh nữ lớp 12D là: 25 (cm)
- d) Sai: Chiều cao của học sinh lớp 12D có độ phân tán bé hơn

Câu 2. Bảng sau biểu diễn mẫu số liệu ghép nhóm thống kê mức lương của một công ty (đơn vị: triệu đồng).

Nhóm	Tần số
[10;15)	15
[15;20)	18
[20;25)	10
[25;30)	10
[30;35)	5
[35;40)	2
	$n = 60$

Xét tính đúng sai của các khẳng định sau:

- a) Khoảng biến thiên của mẫu số liệu ghép nhóm trên là: $R = 30$
- b) Số phần tử của mẫu là $n = 60$
- c) Tứ phân vị thứ nhất là: $Q_1 = 15$
- d) Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $\Delta_Q = 3$

Lời giải

- a) Đúng: Khoảng biến thiên của mẫu số liệu ghép nhóm trên là: $R = 40 - 10 = 30$
- b) Đúng: Số phần tử của mẫu là $n = 60$
- c) Đúng: Tần số tích lũy của các nhóm lần lượt là



$$cf_1 = 15, cf_2 = 33, cf_3 = 43, cf_4 = 53, cf_5 = 58, cf_6 = 60$$

Ta có: $\frac{n}{4} = \frac{60}{4} = 15$ suy ra nhóm 1 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 15. Xét nhóm 1 là nhóm $[10;15]$ có $s = 10, h = 5, n_1 = 15$

Ta có tứ phân vị thứ nhất là: $Q_1 = s + \left(\frac{15 - cf_0}{n_1} \right) \cdot h = 10 + \left(\frac{15 - 0}{15} \right) \cdot 5 = 15$

d) Sai: Ta có: $\frac{3n}{4} = \frac{3 \cdot 60}{4} = 45$ mà $43 < 45 < 53$ suy ra nhóm 4 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 45. Xét nhóm 4 là nhóm $[25;30]$ có $t = 25, l = 5, n_4 = 10$ và nhóm 3 là nhóm $[20;25]$ có $cf_3 = 43$

Ta có tứ phân vị thứ ba là: $Q_3 = t + \left(\frac{45 - cf_3}{n_4} \right) \cdot l = 25 + \left(\frac{45 - 43}{10} \right) \cdot 5 = 26$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $Q_3 - Q_1 = 26 - 15 = 9$

Câu 3. Bảng sau biểu diễn mẫu số liệu ghép nhóm về độ tuổi của cư dân trong một khu phố.

Nhóm	Tần số
$[20;30)$	25
$[30;40)$	20
$[40;50)$	20
$[50;60)$	15
$[60;70)$	14
$[70;80)$	6
	$n = 100$

Xét tính đúng sai của các khẳng định sau:

a) Khoảng biến thiên của mẫu số liệu ghép nhóm trên là: $R = 60$

b) Tứ phân vị thứ nhất là: $Q_1 = 35$

c) Tứ phân vị thứ ba là: $Q_3 = \frac{160}{3}$

d) Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $\Delta_Q = \frac{65}{3}$

Lời giải

a) Đúng: Khoảng biến thiên của mẫu số liệu ghép nhóm trên là: $R = 80 - 20 = 60$

b) Đúng: Số phần tử của mẫu là $n = 100$

Tần số tích lũy của các nhóm lần lượt là $cf_1 = 25, cf_2 = 45, cf_3 = 65, cf_4 = 80, cf_5 = 94, cf_6 = 100$



Ta có: $\frac{n}{4} = \frac{100}{4} = 25$ suy ra nhóm 1 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 25.

Xét nhóm 1 là nhóm $[20;30]$ có $s = 20, h = 10, n_1 = 25$

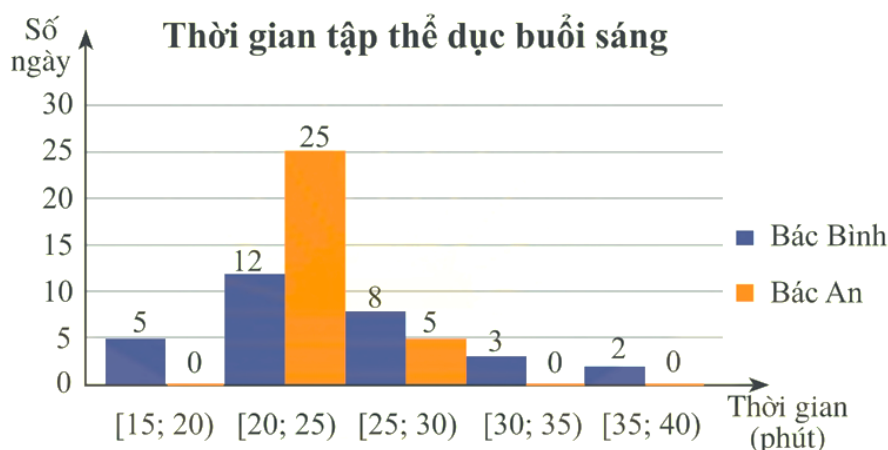
Ta có tứ phân vị thứ nhất là: $Q_1 = s + \left(\frac{25 - cf_0}{n_1} \right) \cdot h = 25 + \left(\frac{25 - 0}{25} \right) \cdot 10 = 35$

c) Sai: Ta có: $\frac{3n}{4} = \frac{3 \cdot 100}{4} = 75$ mà $65 < 75 < 80$ suy ra nhóm 4 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 80. Xét nhóm 4 là nhóm $[50;60]$ có $t = 50, l = 10, n_4 = 15$ và nhóm 3 là nhóm $[40;50]$ có $cf_3 = 65$

Ta có tứ phân vị thứ ba là: $Q_3 = t + \left(\frac{75 - cf_3}{n_4} \right) \cdot l = 50 + \left(\frac{75 - 65}{15} \right) \cdot 10 = \frac{170}{3}$

d) Đúng: Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $Q_3 - Q_1 = \frac{170}{3} - 35 = \frac{65}{3}$

Câu 4. 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.



Xét tính đúng sai của các khẳng định sau:

a) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bác Bình là 25 (phút).

b) Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bác An là: $\Delta_Q = 2$

c) Tứ phân vị thứ ba của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bác Bình là: $Q_3' = \frac{455}{16}$

d) Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng mỗi ngày của bác An lớn hơn bác Bình

Lời giải

a) Đúng: 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) Sai: Cỡ mẫu $n = 30$;

Gọi $x_1; x_2; \dots; x_{30}$ là mẫu số liệu gốc về thời gian tập thể dục buổi sáng mỗi ngày của bác An được xếp theo thứ tự không giảm.

Ta có: $x_1; x_2; \dots; x_{25} \in [20; 25); x_{26}; \dots; x_{30} \in [25; 30)$;

Tứ phân vị thứ nhất của mẫu số liệu gốc là $x_8 \in [20; 25)$. Do đó, tứ phân vị thứ nhất của mẫu số

$$\text{liệu ghép nhóm là: } Q_1 = 20 + \frac{\frac{30}{4}}{25} (25 - 20) = \frac{43}{2}$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{23} \in [20; 25)$. Do đó, tứ phân vị thứ ba của mẫu số liệu

$$\text{ghép nhóm là: } Q_3 = 20 + \frac{\frac{3.30}{4}}{25} (25 - 20) = \frac{49}{2}$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 3$

Gọi $y_1; y_2; \dots; y_{30}$ là mẫu số liệu gốc về thời gian tập thể dục buổi sáng mỗi ngày của bác Bình được xếp theo thứ tự không giảm.

Ta có: $y_1; y_2; \dots; y_5 \in [15; 20); y_6; \dots; y_{17} \in [20; 25); y_{18}; \dots; y_{25} \in [25; 30); y_{26}; y_{27}; y_{28} \in [30; 35)$;

$y_{29}; y_{30} \in [35; 40)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $y_8 \in [20; 25)$. Do đó, tứ phân vị thứ nhất của mẫu số

$$\text{liệu ghép nhóm là: } Q_1' = 20 + \frac{\frac{30}{4}}{12} (25 - 20) = \frac{185}{8}$$

c) Đúng: Tứ phân vị thứ ba của mẫu số liệu gốc là $y_{23} \in [25; 30)$. Do đó, tứ phân vị thứ ba của

$$\text{mẫu số liệu ghép nhóm là: } Q_3' = 25 + \frac{\frac{3.30}{4} - (5 + 12)}{8} (30 - 25) = \frac{455}{16}$$

d) Sai: Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng mỗi ngày của bác Bình lớn hơn bác An

Câu 5. 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



Xét tính đúng sai của các khẳng định sau:

a) Cỡ mẫu $n = 100$.

b) Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là $Q_1 = \frac{683}{38}$.

c) Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = \frac{515}{114}$

d) 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 là giá trị ngoại lệ của mẫu số liệu ghép nhóm.

Lời giải

a) Đúng: 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: $x_1, \dots, x_{22} \in [15; 18); x_{23}, \dots, x_{60} \in [18; 21); x_{61}, \dots, x_{87} \in [21; 24); x_{88}, \dots, x_{95} \in [24; 27);$

$x_{96}, \dots, x_{99} \in [27; 30); x_{100} \in [30; 33)$.

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

c) Sai: 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 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 $\Delta_Q = \frac{68}{3} - \frac{693}{38} = \frac{505}{114}$

d) Đúng: 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.

Câu 6. 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

Xét tính đúng sai của các khẳng định sau:

a) Khoảng biến thiên của mẫu số liệu ghép nhóm ứng với khu vực A là: 15 (tuổi)

b) Khoảng biến thiên của mẫu số liệu ghép nhóm ứng với khu vực B là: 12 (tuổi)



c) Khoảng tứ phân vị của mẫu số liệu ghép nhóm ứng với khu vực A là: $\frac{61}{3}$ (tuổi)

d) Nếu so sánh theo khoảng tứ phân vị thì phụ nữ ở khu vực B có độ tuổi kết hôn đồng đều hơn

Lời giải

a) Đúng: Khoảng biến thiên của mẫu số liệu ghép nhóm ứng với khu vực A là: $34 - 19 = 15$ (tuổi)

b) Đúng: Khoảng biến thiên của mẫu số liệu ghép nhóm ứng với khu vực B là: $31 - 19 = 12$ (tuổi)

c) Sai: Cỡ mẫu $n = 100$

Gọi $x_1; x_2; \dots; x_{100}$ là mẫu số liệu gốc về độ tuổi kết hôn của phụ nữ ở khu vực A được xếp theo thứ tự không giảm.

Ta có: $x_1; x_2; \dots; x_{10} \in [19; 22); x_{11}; \dots; x_{37} \in [22; 25); x_{38}; \dots; x_{68} \in [25; 28); x_{69}; \dots; x_{93} \in [28; 31);$

$x_{94}; \dots; x_{100} \in [31; 34)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_{25} + x_{26}) \in [22; 25)$. Do đó, tứ phân vị thứ nhất

của mẫu số liệu ghép nhóm là: $Q_1 = 22 + \frac{\frac{100}{4} - 10}{27}(25 - 22) = \frac{71}{3}$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{75} + x_{76}) \in [28; 31)$. Do đó, tứ phân vị thứ ba của

mẫu số liệu ghép nhóm là: $Q_3 = 28 + \frac{\frac{3 \cdot 100}{4} - (10 + 27 + 31)}{25}(31 - 28) = \frac{721}{25}$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = \frac{388}{75}$

Gọi $y_1; y_2; \dots; y_{100}$ là mẫu số liệu gốc về độ tuổi kết hôn của phụ nữ ở khu vực B được xếp theo thứ tự không giảm.

Ta có: $y_1; y_2; \dots; y_{47} \in [19; 22); y_{48}; \dots; y_{87} \in [22; 25); y_{88}; \dots; y_{98} \in [25; 30); y_{99}; y_{100} \in [28; 31)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(y_{25} + y_{26}) \in [19; 22)$. Do đó, tứ phân vị thứ nhất

của mẫu số liệu ghép nhóm là: $Q'_1 = 19 + \frac{\frac{100}{4}}{47}(22 - 19) = \frac{968}{47}$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(y_{75} + y_{76}) \in [22; 25)$. Do đó, tứ phân vị thứ ba của

mẫu số liệu ghép nhóm là: $Q'_3 = 22 + \frac{\frac{3 \cdot 100}{4} - 47}{40}(25 - 22) = \frac{241}{10}$

d) Đúng: Có $\Delta_Q' < \Delta_Q$ nên phụ nữ ở khu vực B có độ tuổi kết hôn đồng đều hơn



Câu 7. Bảng sau thống kê tổng lượng mưa (đơn vị: mm) đo được vào tháng 7 từ năm 2002 đến 2021 tại một trạm quan trắc đặt ở Cà Mau.

341,4	187,1	242,2	522,9	251,4	432,2	200,7	388,6	258,4	288,5
298,1	413,5	413,5	332	421	475	400	305	520	147

Xét tính đúng sai của các khẳng định sau:

a) Khoảng biến thiên của mẫu số liệu: $375,9(mm)$

b) Khoảng tứ phân vị của mẫu số liệu là: $\Delta_Q = \frac{1827}{100}$

c) Chia mẫu số liệu trên thành 4 nhóm như bảng:

Lượng mưa	[140;240)	[240;340)	[340;440)	[440;540)
Số tháng	3	7	7	3

Khoảng biến thiên của mẫu số liệu ghép nhóm: $400(mm)$

d) Chia mẫu số liệu trên thành 4 nhóm như bảng:

Lượng mưa	[140;240)	[240;340)	[340;440)	[440;540)
Số tháng	3	7	7	3

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q' = \frac{1000}{7}$

Lời giải

a) Đúng: Sắp xếp lại mẫu số liệu theo thứ tự tăng dần:

147; 187,1; 200,7; 242,2; 251,4; 258,4; 288,5; 298,1; 305; 332; 341,4; 388,6; 400; 413,5; 413,5; 421; 432,2; 475; 520; 522,9.

Khoảng biến thiên của mẫu số liệu: $522,9 - 147 = 375,9(mm)$

Tứ phân vị thứ nhất của mẫu số liệu là trung vị của 147; 187,1; 200,7; 242,2; 251,4; 258,4; 288,5;

298,1; 305; 332 nên: $Q_1 = \frac{6276}{25}$

Tứ phân vị thứ nhất của mẫu số liệu là trung vị của 341,4; 388,6; 400; 413,5; 413,5; 421; 432,2;

475; 520; 522,9 nên: $Q_3 = \frac{43281}{100}$

Khoảng tứ phân vị của mẫu số liệu là: $\Delta_Q = Q_3 - Q_1 = \frac{18177}{100}$

b) Sai:

Lượng mưa	[140;240)	[240;340)	[340;440)	[440;540)
Số tháng	3	7	7	3



c) Đúng: Khoảng biến thiên của mẫu số liệu ghép nhóm: $540 - 140 = 400(mm)$

Cỡ mẫu $n = 20$;

Gọi $x_1; x_2; \dots; x_{20}$ là mẫu số liệu gốc về lượng mưa đo được vào tháng 7 từ năm 2002 đến 2021 tại một trạm quan trắc đặt ở Cà Mau được xếp theo thứ tự không giảm.

Ta có: $x_1; \dots; x_3 \in [140; 240); x_4; \dots; x_{10} \in [240; 340); x_{11}; \dots; x_{17} \in [340; 440); x_{18}; \dots; x_{20} \in [440; 540)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_5 + x_6) \in [240; 340)$. Do đó, tứ phân vị thứ nhất

của mẫu số liệu ghép nhóm là: $Q_1' = 240 + \frac{\frac{20}{2} - 3}{7}(340 - 240) = \frac{1880}{7}$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{15} + x_{16}) \in [340; 440)$. Do đó, tứ phân vị thứ ba của

mẫu số liệu ghép nhóm là: $Q_3' = 340 + \frac{\frac{3 \cdot 20}{2} - (3 + 7)}{7}(440 - 340) = \frac{2880}{7}$

d) Đúng: Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q' = Q_3' - Q_1' = \frac{1000}{7}$

Khoảng biến thiên của mẫu số liệu ghép nhóm lớn hơn mẫu số liệu; khoảng tứ phân vị của mẫu số liệu ghép nhóm nhỏ hơn mẫu số liệu

Câu 8. Kết quả đo chiều cao của 100 cây keo 3 năm tuổi tại một nông trường được cho ở bảng sau:

Chiều cao (m)	[8, 4; 8, 6)	[8, 6; 8, 8)	[8, 8; 9, 0)	[9, 0; 9, 2)	[9, 2; 9, 4)
Số cây	5	12	25	44	14

Xét tính đúng sai của các khẳng định sau:

a) Khoảng biên thiên của mẫu số liệu ghép nhóm: 1 m

b) Tứ phân vị thứ ba của mẫu số liệu ghép nhóm là: $Q_3 = 10,5$

c) Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = 2,06$

d) Trong 100 cây keo trên có 1 cây cao 8,4 m. Thì chiều cao của cây keo này là giá trị ngoại lệ

Lời giải

a) Đúng: Khoảng biên thiên của mẫu số liệu ghép nhóm: $9,4 - 8,4 = 1(m)$

b) Đúng: Cỡ mẫu $n = 100$;

Gọi $x_1; x_2; \dots; x_{100}$ là mẫu số liệu gốc được xếp theo thứ tự không giảm.

Ta có:

$x_1; \dots; x_5 \in [8, 4; 8, 6); x_6; \dots; x_{17} \in [8, 6; 8, 8); x_{18}; \dots; x_{42} \in [8, 8; 9, 0); x_{43}; \dots; x_{86} \in [9, 0; 9, 2); x_{87}; \dots; x_{100} \in [9, 2; 9, 4)$



Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_{25} + x_{26}) \in [8,8;9,0)$. Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là: $Q_1 = 8,8 + \frac{\frac{100}{4} - (5 + 12)}{25}(9,0 - 8,8) = 9,44$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{75} + x_{76}) \in [9,0;9,2)$. Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là: $Q_3 = 9,0 + \frac{\frac{3 \cdot 100}{4} - (5 + 12 + 25)}{44}(9,2 - 9,0) = 10,5$

c) Sai: Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 1,06$

d) Sai: Giá trị x trong mẫu số liệu là giá trị ngoại lệ nếu $x > Q_3 + 1,5\Delta_Q$ hoặc $x < Q_1 - 1,5\Delta_Q$

Hay $x > 10,5 + 1,5 \cdot 1,06 = 12,09$ hoặc $x < 9,44 - 1,5 \cdot 1,06 = 7,85$

Vậy cây cao 8,4m không phải là giá trị ngoại lệ.

Câu 9. Thống kê điểm trung bình môn Toán của một số học sinh lớp 11 được cho ở bảng sau:

Khoảng điểm	[6,5;7)	[7;7,5)	[7,5;8)	[8;8,5)	[8,5;9)	[9;9,5)	[9,5;10)
Tần số	8	10	16	24	13	7	4

Xét tính đúng sai của các khẳng định sau:

a) Khoảng biến thiên của mẫu số liệu là $R = 4$

b) Số trung bình của mẫu số liệu xấp xỉ bằng: 8,12

c) Mốt của mẫu số liệu là: 6,21

d) Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = 2,05$

Lời giải

Khoảng điểm	[6,5;7)	[7;7,5)	[7,5;8)	[8;8,5)	[8,5;9)	[9;9,5)	[9,5;10)
Giá trị đại diện	6,75	7,25	7,75	8,25	8,75	9,25	9,75
Tần số	8	10	16	24	13	7	4

a) Đúng: Khoảng biến thiên của mẫu số liệu là $R = 4$

b) Đúng: Số trung bình của mẫu số liệu xấp xỉ bằng:

$$(6,75 \cdot 8 + 7,25 \cdot 10 + 7,75 \cdot 16 + 8,25 \cdot 24 + 8,75 \cdot 13 + 9,25 \cdot 7 + 9,75 \cdot 4) : 82 = 8,12$$

Nhóm chứa mốt của mẫu số liệu là: [8;8,5)

c) Sai: Mốt của mẫu số liệu là: $M_0 = 8 + \frac{24 - 16}{(24 - 16) + (24 - 13)} \cdot (8,5 - 8) = 8,21$



d) Sai: Gọi $x_1; x_2; x_3; \dots; x_{85}$ lần lượt là tần số theo thứ tự không gian

Do $x_1, \dots, x_8 \in [6,5;7); x_9, \dots, x_{18} \in [7;7,5); x_{19}, \dots, x_{34} \in [7,5;8);$

$x_{35}, \dots, x_{58} \in [8;8,5); x_{59}, \dots, x_{71} \in [8,5;9); \dots$

Tứ phân vị thứ hai của dãy số liệu là $\frac{1}{2}(x_{41} + x_{42})$ thuộc nhóm $[8;8,5)$ nên tứ phân vị thứ hai của

$$\text{mẫu số liệu là } Q_2 = 8 + \frac{\frac{82}{2} - 34}{24} (8,5 - 8) = 8,15$$

Tứ phân vị thứ nhất của dãy số liệu là $\frac{1}{2}(x_{20} + x_{21})$ thuộc nhóm $[7,5;8)$ nên tứ phân vị thứ nhất

$$\text{của mẫu số liệu là } Q_1 = 7,5 + \frac{\frac{82}{2} - 18}{16} (8 - 7,5) = 7,58$$

Tứ phân vị thứ ba của dãy số liệu là $\frac{1}{2}(x_{61} + x_{62})$ thuộc nhóm $[8,5;9)$ nên tứ phân vị thứ ba của

$$\text{mẫu số liệu là } Q_3 = 8,5 + \frac{\frac{3.82}{2} - 58}{13} (9 - 8,5) = 8,63$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = 1,05$

Câu 10. Một công ty cung cấp nước sạch thống kê lượng nước các hộ gia đình trong một khu vực tiêu thụ trong một tháng ở bảng sau:

Lượng nước tiêu thụ (m^3)	[3;6)	[6;9)	[9;12)	[12;15)	[15;18)
Số hộ gia đình	24	57	42	29	8

Xét tính đúng sai của các khẳng định sau:

a) Số trung bình của mẫu số liệu ghép nhóm đã cho là 9,375.

b) Mốt của mẫu số liệu là $M_o = 8,0625$.

c) Khoảng biến thiên của mẫu số liệu ghép nhóm đã cho là 15

d) Công ty muốn gửi một thông báo khuyến nghị tiết kiệm nước đến 25% các hộ gia đình có lượng nước tiêu thụ cao nhất. Khi đó công ty nên gửi thông báo tiết kiệm nước đến các hộ gia đình có lượng nước tiêu thụ từ $14,79m^3$ nước trở lên.

Lời giải

a) Đúng: Cỡ mẫu $n = 160$. Bảng tần số ghép nhóm của mẫu số liệu trên như sau:

Lượng nước tiêu thụ (m^3)	[3;6)	[6;9)	[9;12)	[12;15)	[15;18)
Giá trị đại diện	4,5	7,5	10,5	13,5	16,5
Số hộ gia đình	24	57	42	29	8



Số trung bình của mẫu số liệu ghép nhóm đã cho là

$$\bar{x} = \frac{1}{160}(24.4,5 + 57.7,5 + 42.10,5 + 29.13,5 + 8.16,5) = 9,375.$$

Nhóm chứa một của mẫu số liệu ghép nhóm trên là nhóm $[6;9)$.

Do đó: $u_m = 6; n_m = 57; n_{m-1} = 24; n_{m+1} = 42; u_{m+1} = 9$.

b) Đúng: Một của mẫu số liệu là $M_o = 6 + \frac{(57 - 24)}{(57 - 24) + (57 - 42)} \cdot (9 - 6) = 8,0625$.

Gọi $x_1; x_2; \dots; x_{160}$ là mẫu số liệu được xếp theo thứ tự không giảm.

Ta có: $x_1, \dots, x_{24} \in [3;6); x_{25}, \dots, x_{81} \in [6;9); x_{82}, \dots, x_{123} \in [9;12);$
 $x_{124}, \dots, x_{152} \in [12;15); x_{153}, \dots, x_{160} \in [15;18).$

Cỡ mẫu $n = 160$ là số chẵn nên trung vị là $M_e = \frac{1}{2}(x_{80} + x_{81})$.

Trung vị của mẫu số liệu là $M_e = 6 + \frac{\frac{160}{2} - 24}{57}(9 - 6) \approx 8,95$.

c) Đúng: Khoảng biến thiên của mẫu số liệu là: $R = 18 - 3 = 15$

d) Sai: 25% các hộ gia đình có lượng nước tiêu thụ cao nhất có lượng nước tiêu thụ không nhỏ hơn Q_3 , với Q_3 là tứ phân vị thứ ba của mẫu số liệu.

Tứ phân vị thứ ba của dãy số liệu $x_1; x_2; \dots; x_{160}$ là $\frac{1}{2}(x_{120} + x_{121})$.

Do x_{120} và x_{121} thuộc nhóm $[9;12)$ nên tứ phân vị thứ ba của mẫu số liệu là

$$Q_3 = 9 + \frac{\frac{160.3}{4} - (24 + 57)}{42} \cdot (12 - 9) \approx 11,79.$$

Vậy công ty nên gửi thông báo tiết kiệm nước đến các hộ gia đình có lượng nước tiêu thụ từ $11,79m^3$ nước trở lên.

PHẦN III. Câu trắc nghiệm trả lời ngắn

Câu 11. Bảng dưới biểu diễn mẫu số liệu ghép nhóm về chiều cao (đơn vị: centimét) của 36 học sinh nam lớp 12 ở một trường trung học phổ thông. Tìm khoảng biến thiên của mẫu số liệu ghép nhóm đó.

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

**Lời giải**

Trong mẫu số liệu ghép nhóm đó, ta có: đầu mút trái của nhóm 1 là $a_1 = 160$, đầu mút phải của nhóm 5 là $a_6 = 175$. Vậy khoảng biến thiên của mẫu số liệu ghép nhóm đó là:

$$R = a_6 - a_1 = 175 - 160 = 15 \text{ (cm)}$$

Câu 12. Bảng dưới biểu diễn mẫu số liệu ghép nhóm về chiều cao của 42 mẫu cây ở một vườn thực vật (đơn vị: centimét). Tính khoảng tứ phân vị của mẫu số liệu ghép nhóm đó (làm tròn kết quả đến hàng phần mười).

Nhóm	Tần số	Tần số tích lũy
[40; 45)	5	5
[45; 50)	10	15
[50; 55)	7	22
[55; 60)	9	31
[60; 65)	7	38
[65; 70)	4	42
	$n = 42$	

Lời giải

Số phần tử của mẫu là $n = 42$.

Ta có: $\frac{n}{4} = \frac{42}{4} = 10,5$ mà $5 < 10,5 < 15$. Suy ra nhóm 2 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 10,5. Xét nhóm 2 là nhóm $[45; 50)$ có $s = 45; h = 5; n_2 = 10$ và nhóm 1 là nhóm $[40; 45)$ có $cf_1 = 5$.

Áp dụng công thức, ta có tứ phân vị thứ nhất là: $Q_1 = 45 + \left(\frac{10,5 - 5}{10} \right) \cdot 5 = 47,75 \text{ (cm)}$

Ta có: $\frac{3n}{4} = \frac{3 \cdot 42}{4} = 31,5$ mà $31 < 31,5 < 38$. Suy ra nhóm 5 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 31,5.

Xét nhóm 5 là nhóm $[60; 65)$ có $t = 60; l = 5; n_5 = 7$ và nhóm 4 là nhóm $[55; 60)$ có $cf_4 = 31$.

Áp dụng công thức, ta có tứ phân vị thứ ba là: $Q_3 = 60 + \left(\frac{31,5 - 31}{7} \right) \cdot 5 \approx 60,4 \text{ (cm)}$.

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm đã cho là:

$$\Delta_Q = Q_3 - Q_1 \approx 60,4 - 47,75 = 12,65 \text{ (cm)}$$

Câu 13. 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 (gam)	[250; 290)	[290; 330)	[330; 370)	[370; 410)	[410; 450)
Số quả xoài	3	13	18	11	5

Khoảng tứ phân vị của mẫu số liệu ghép nhóm đó bằng $\frac{a}{b}$ với $\frac{a}{b}$ là phân số tối giản và $a, b \in \mathbb{Z}$

Tính giá trị biểu thức $T = a + b$

Lời 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 nhóm là

$$Q_1 = 290 + \frac{\frac{50}{4} - 3}{13} (330 - 290) = \frac{4150}{13}.$$

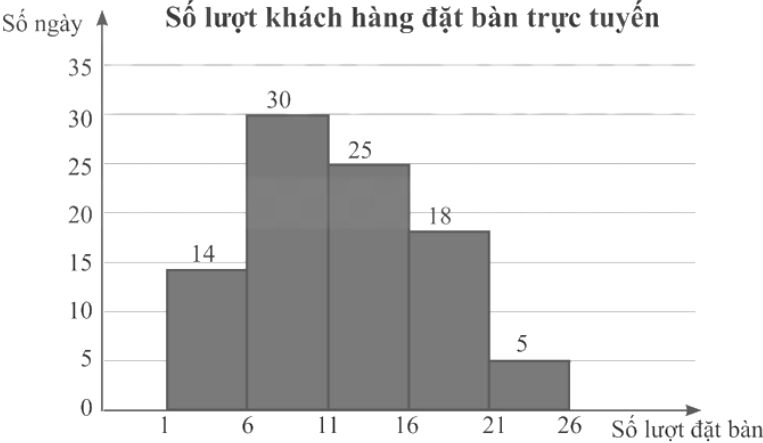
Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{38} \in [370; 410)$. Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là

$$Q_3 = 370 + \frac{\frac{3.50}{4} - (3 + 13 + 18)}{11} (410 - 370) = \frac{4210}{11}.$$

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

Vậy giá trị biểu thức $T = a + b = 9080 + 143 = 9223$.

Câu 14. Biểu đồ dưới đây biểu diễn số lượt khách hàng đặt bàn qua hình thức trực tuyến mỗi ngày trong quý III năm 2022 của một nhà hàng. Cột thứ nhất biểu diễn số ngày có từ 1 đến dưới 6 lượt đặt bàn; cột thứ hai biểu diễn số ngày có từ 6 đến dưới 11 lượt đặt bàn;...



Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm cho bởi biểu đồ trên.

Lời giải



Cỡ mẫu $n = 92$;

Gọi $x_1; x_2; \dots; x_{92}$ là mẫu số liệu gốc về số lượt khách hàng đặt bàn qua hình thức trực tuyến mỗi ngày trong quý III năm 2022 của nhà hàng được xếp theo thứ tự không giảm.

Ta có:

$$x_1; \dots; x_{14} \in [1; 6); x_{15}; \dots; x_{44} \in [6; 11); x_{45}; \dots; x_{69} \in [11; 16); x_{70}; \dots; x_{87} \in [16; 21); x_{88}; \dots; x_{92} \in [21; 26)$$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_{23} + x_{24}) \in [6; 11)$. Do đó, tứ phân vị thứ nhất của

$$\text{mẫu số liệu ghép nhóm là: } Q_1 = 6 + \frac{\frac{92}{4} - 14}{30}(11 - 6) = 7,5$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{69} + x_{70}) \in [11; 16)$ và $[16; 21)$.

Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là: $Q_3 = 16$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 8,5$

Câu 15. Thời gian luyện tập trong một ngày (tính theo giờ) của một số vận động viên được ghi lại ở bảng sau:

Thời gian luyện tập (giờ)	[0; 2)	[2; 4)	[4; 6)	[6; 8)	[8; 10)
Số vận động viên	3	8	12	12	4

Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.

Lời giải

Nhắc lại: $x_1 \leq x_2 \leq \dots \leq x_{39}$ là thời gian luyện tập của 39 vận động viên.

Tứ phân vị thứ hai của dãy số liệu $x_1; x_2; \dots; x_{39}$ là $x_{20} \in [4; 6)$. Do đó tứ phân vị thứ hai của mẫu

$$\text{số liệu ghép nhóm là } Q_2 = 4 + \frac{\frac{2.39}{4} - (3 + 8)}{12}(6 - 4) = \frac{65}{12} \approx 5,417.$$

Tứ phân vị thứ nhất của dãy số liệu $x_1; x_2; \dots; x_{39}$ là $x_{10} \in [2; 4)$.

$$\text{Do đó tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là } Q_1 = 2 + \frac{\frac{1.39}{4} - 3}{8}(4 - 2) = \frac{59}{16} = 3,6875.$$

Tứ phân vị thứ ba của dãy số liệu $x_1; x_2; \dots; x_{39}$ là $x_{30} \in [6; 8)$.

Do đó tứ phân vị thứ ba của mẫu số liệu ghép nhóm là

$$Q_3 = 6 + \frac{\frac{3.39}{4} - (3 + 8 + 12)}{12}(8 - 6) = \frac{169}{24} \approx 7,042.$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = 1,625$



Câu 16. Một người thống kê lại thời gian thực hiện các cuộc gọi điện thoại của người đó trong một tuần ở bảng sau:

Thời gian (đơn vị: giây)	[0; 60)	[60; 120)	[120; 180)	[180; 240)	[240; 300)	[300; 360)
Số cuộc gọi	8	10	7	5	2	1

Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.

Lời giải

Số lần thực hiện cuộc gọi là: $n = 8 + 10 + 7 + 5 + 2 + 1 = 33$

Gọi $x_1; x_2; x_3; \dots; x_{33}$ lần lượt là thời gian thực hiện cuộc gọi theo thứ tự không gian

Do $x_1, \dots, x_8 \in [0; 60); x_9, \dots, x_{18} \in [60; 120); x_{19}, \dots, x_{25} \in [120; 180); x_{26}, \dots, x_{30} \in [180; 240); \dots$

Tứ phân vị thứ hai của dãy số liệu là $\frac{1}{2}(x_{16} + x_{17})$ thuộc nhóm $[60; 120)$ nên tứ phân vị thứ hai

$$\text{của mẫu số liệu là } Q_2 = 60 + \frac{\frac{33}{2} - 8}{10}(120 - 60) = 111$$

Tứ phân vị thứ nhất của dãy số liệu là $\frac{1}{2}(x_8 + x_9)$ thuộc nhóm $[60; 120)$ nên tứ phân vị thứ nhất

$$\text{của mẫu số liệu là } Q_1 = 60 + \frac{\frac{33}{4} - 8}{10}(120 - 60) = 61,5$$

Tứ phân vị thứ ba của dãy số liệu là $\frac{1}{2}(x_{24} + x_{25})$ thuộc nhóm $[120; 180)$ nên tứ phân vị thứ ba

$$\text{của mẫu số liệu là } Q_3 = 120 + \frac{\frac{3.33}{4} - 18}{7}(180 - 120) = 177,8$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = 116,3$

Câu 17. Số điểm một cầu thủ bóng rổ ghi được trong 20 trận đấu được cho ở bảng sau:

Điểm số	[5, 5; 10, 5)	[10, 5; 15, 5)	[15, 5; 20, 5)	[20, 5; 25, 5)
Số trận	3	9	2	6

Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.

Lời giải

Gọi $x_1; x_2; x_3; \dots; x_{20}$ lần lượt là số trận theo thứ tự không gian

Do $x_1, \dots, x_3 \in [5, 5; 10, 5); x_4, \dots, x_{12} \in [10, 5; 15, 5); x_{13}, x_{14} \in [15, 5; 20, 5); x_{15}, \dots, x_{20} \in [20, 5; 25, 5)$

Tứ phân vị thứ hai của dãy số liệu là $\frac{1}{2}(x_{10} + x_{11})$ thuộc nhóm $[10, 5; 15, 5)$ nên tứ phân vị thứ hai

$$\text{của mẫu số liệu là } Q_2 = 10,5 + \frac{\frac{20}{2} - 3}{9}(15,5 - 10,5) = 14,4$$



Tứ phân vị thứ nhất của dãy số liệu là $\frac{1}{2}(x_5 + x_6)$ thuộc nhóm $[10,5;15,5)$ nên tứ phân vị thứ nhất của mẫu số liệu là $Q_1 = 10,5 + \frac{\frac{20}{2} - 3}{9}(15,5 - 10,5) = 11,6$

Tứ phân vị thứ ba của dãy số liệu là $\frac{1}{2}(x_{15} + x_{16})$ thuộc nhóm $[20,5;25,5)$ nên tứ phân vị thứ ba của mẫu số liệu là $Q_3 = 20,5 + \frac{\frac{3 \cdot 20}{2} - 14}{6}(25,5 - 20,5) = 21,3$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = 9,7$.

Câu 18. Thời gian sử dụng điện thoại trong một ngày của 30 sinh viên được ghi lại ở bảng sau (đơn vị: phút).

Thời gian (phút)	$[0;60)$	$[60;120)$	$[120;180)$	$[180;240)$	$[240;300)$
Số sinh viên	2	7	7	10	4

Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.

Lời giải

Gọi $x_1; x_2; \dots; x_{30}$ là mẫu số liệu được xếp theo thứ tự không giảm.

Ta có: $x_1, x_2 \in [0;60); x_3, \dots, x_9 \in [60;120); x_{10}, \dots, x_{16} \in [120;180);$
 $x_{17}, \dots, x_{26} \in [180;240); x_{27}, \dots, x_{30} \in [240;300).$

Tứ phân vị thứ hai của mẫu số liệu ghép nhóm là $\frac{1}{2}(x_{15} + x_{16}).$

Do x_{15} và x_{16} thuộc nhóm $[120;180)$ nên tứ phân vị thứ hai của mẫu số liệu ghép nhóm là

$$Q_2 = 120 + \frac{\frac{30}{2} - (2 + 7)}{7}(180 - 120) = \frac{1200}{7} \approx 171,43.$$

Tứ phân vị thứ nhất của mẫu là $x_8 \in [60;120)$. Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép

$$\text{nhóm là } Q_1 = 60 + \frac{\frac{30}{4} - 2}{7}(120 - 60) = \frac{750}{7} \approx 107,14.$$

Tứ phân vị thứ ba của mẫu số liệu là $x_{23} \in [180;240).$

Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là

$$Q_3 = 180 + \frac{\frac{3 \cdot 30}{4} - (2 + 7 + 7)}{10}(240 - 180) = 219.$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = 111,86$



Câu 19. Điểm thi môn Toán (thang điểm 100, điểm được làm tròn đến 1) của 60 thí sinh được cho mẫu số liệu ghép nhóm như sau:

Điểm	[0; 9, 5)	[9, 5; 19, 5)	[19, 5; 29, 5)	[29, 5; 39, 5)	[39, 5; 49, 5)
Số thí sinh	1	2	4	6	15
Điểm	[49, 5; 59, 5)	[59, 5; 69, 5)	[69, 5; 79, 5)	[79, 5; 89, 5)	[89, 5; 99, 5)
Số thí sinh	12	10	6	3	1

Tính khoảng tứ phân vị của mẫu số liệu ghép nhóm

Lời giải

Cỡ mẫu là $n = 60$.

Gọi $x_1, x_2, \dots, x_{60}$ là điểm thi môn Toán của 60 thí sinh và giả sử dãy này đã được sắp xếp theo thứ tự tăng dần.

Tứ phân vị thứ nhất Q_1 là $\frac{x_{15} + x_{16}}{2}$. Do x_{15} và x_{16} đều thuộc nhóm $[39, 5; 49, 5)$ nên nhóm này chứa Q_1 . Do đó, $p = 5$; $a_5 = 39, 5$; $m_5 = 15$; $m_1 + m_2 + m_3 + m_4 = 13$; $a_6 - a_5 = 10$ và ta có:

$$Q_1 = 39, 5 + \frac{\frac{60}{4} - 13}{15} \cdot 10 \approx 40, 83$$

Tứ phân vị thứ ba Q_3 là $\frac{x_{45} + x_{46}}{2}$. Do x_{45} và x_{46} đều thuộc nhóm $[59, 5; 69, 5)$ nên nhóm này chứa Q_3 .

Do đó, $p = 7$; $a_7 = 59, 5$; $m_7 = 10$; $m_1 + m_2 + m_3 + m_4 + m_5 + m_6 = 40$; $a_6 - a_5 = 10$ và ta có:

$$Q_3 = 59, 5 + \frac{\frac{3 \cdot 60}{4} - 40}{10} \cdot 10 = 64, 5$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = 23, 67$

Câu 20. Kết quả điều tra về số giờ làm thêm trong một tuần của 100 sinh viên được cho cho như bảng sau:

Số giờ làm thêm	[2; 4)	[4; 6)	[6; 8)	[8; 10)	[10; 12)
Số giờ làm thêm đại diện	3	5	7	9	11
Số sinh viên	12	20	37	21	10

Tìm khoảng tứ phân vị của số liệu đó.

Lời giải

Cỡ mẫu $n = 100$.

Gọi $x_1; x_2; \dots; x_{100}$ là mẫu số liệu được xếp theo thứ tự không giảm.



Tứ phân vị thứ hai của dãy số liệu $x_1; x_2; \dots; x_{100}$ là $\frac{1}{2}(x_{50} + x_{51})$.

Do x_{50} và x_{51} thuộc nhóm $[6;8)$ nên $Q_2 = 6 + \frac{\frac{2 \cdot 100}{4} - 32}{37} \cdot (8 - 6) \approx 6,97$.

Tứ phân vị thứ nhất của dãy số liệu $x_1; x_2; \dots; x_{100}$ là $\frac{1}{2}(x_{25} + x_{26})$.

Do x_{25} và x_{26} thuộc nhóm $[4;6)$ nên $Q_1 = 4 + \frac{\frac{1 \cdot 100}{4} - 12}{20} \cdot (6 - 4) = 5,3$.

Tứ phân vị thứ ba của dãy số liệu $x_1; x_2; \dots; x_{100}$ là $\frac{1}{2}(x_{75} + x_{76})$.

Do x_{75} và x_{76} thuộc nhóm $[8;10)$ nên $Q_3 = 8 + \frac{\frac{3 \cdot 100}{4} - 69}{21} \cdot (10 - 8) \approx 8,57$.

Khoảng tứ phân vị là $\Delta Q = 3,27$.

-----HẾT-----

**BÀI 02 PHƯƠNG SAI VÀ ĐỘ LỆCH CHUẨN CỦA MSLGN****A LÝ THUYẾT CẦN NHỚ**

Xét mẫu số liệu ghép nhóm cho bởi công thức sau:

Nhóm	$[u_1; u_2)$	$[u_1; u_2)$	...	$[u_k; u_{k+1})$
Giá trị đại diện	c_1	c_2	...	c_k
Tần số	n_1	n_2	...	n_k

Phương sai: Phương sai của mẫu số liệu ghép nhóm, kí hiệu S^2 , được tính bởi công thức

$$S^2 = \frac{1}{n} \left[n_1 (c_1 - \bar{x})^2 + n_2 (c_2 - \bar{x})^2 + \dots + n_k (c_k - \bar{x})^2 \right]$$

Trong đó: $n = n_1 + n_2 + \dots + n_k$ là cỡ mẫu; $\bar{x} = \frac{1}{n} (n_1 c_1 + n_2 c_2 + \dots + n_k c_k)$

Độ lệch chuẩn: Độ lệch chuẩn của mẫu số liệu ghép nhóm, kí hiệu S , là căn bậc hai số học của phương sai, nghĩa là $S = \sqrt{S^2}$

Ý nghĩa:

- Phương sai (độ lệch chuẩn) của mẫu số liệu ghép nhóm là giá trị xấp xỉ cho phương sai (độ lệch chuẩn) của mẫu số liệu gốc. Chúng được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm xung quanh số trung bình của mẫu số liệu. Phương sai và độ lệch chuẩn càng lớn thì dữ liệu càng phân tán.
- Độ lệch chuẩn có cùng đơn vị với đơn vị của mẫu số liệu.

Chú ý:

- a) Phương sai của mẫu số liệu ghép nhóm có thể được tính theo công thức sau:

$$S^2 = \frac{1}{n} (n_1 c_1^2 + n_2 c_2^2 + \dots + n_k c_k^2) - \bar{x}^2$$

- b) Trong thống kê, người ta còn dùng đại lượng sau để đo mức độ phân tán của mẫu số liệu ghép nhóm:

$$\hat{s}^2 = \frac{1}{n-1} \left[n_1 (c_1 - \bar{x})^2 + n_2 (c_2 - \bar{x})^2 + \dots + n_k (c_k - \bar{x})^2 \right]$$

**B PHÂN LOẠI VÀ PHƯƠNG PHÁP GIẢI TOÁN****Dạng 1: Tính phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm**

Xét mẫu số liệu ghép nhóm cho bởi công thức sau:

Nhóm	$[u_1; u_2)$	$[u_1; u_2)$	...	$[u_k; u_{k+1})$
Giá trị đại diện	c_1	c_2	...	c_k
Tần số	n_1	n_2	...	n_k

Số trung bình cộng của mẫu số liệu ghép nhóm trên được tính bằng công thức:

$$\bar{x} = \frac{1}{n}(n_1c_1 + n_2c_2 + \dots + n_kc_k)$$

Phương sai: Phương sai của mẫu số liệu ghép nhóm, kí hiệu S^2 , được tính bởi công thức

$$S^2 = \frac{1}{n} \left[n_1(c_1 - \bar{x})^2 + n_2(c_2 - \bar{x})^2 + \dots + n_k(c_k - \bar{x})^2 \right]$$

Trong đó: $n = n_1 + n_2 + \dots + n_k$ là cỡ mẫu; $\bar{x} = \frac{1}{n}(n_1c_1 + n_2c_2 + \dots + n_kc_k)$

Độ lệch chuẩn: Độ lệch chuẩn của mẫu số liệu ghép nhóm, kí hiệu S , là căn bậc hai số học của phương sai, nghĩa là $S = \sqrt{S^2}$

BÀI TẬP TỰ LUẬN

Bài tập 1: Tìm cân nặng trung bình của học sinh lớp 11D cho trong bảng sau

Cân nặng	$[40,5; 45,5)$	$[45,5; 50,5)$	$[50,5; 55,5)$	$[55,5; 60,5)$	$[60,5; 65,5)$	$[65,5; 70,5)$
Số học sinh	10	7	16	4	2	3

Bài tập 2: Kết quả khảo sát cân nặng của 25 quả cam ở mỗi lô hàng A và B được cho bởi bảng sau:

Cân nặng (g)	$[150; 155)$	$[155; 160)$	$[160; 165)$	$[165; 170)$	$[170; 175)$
Số quả cam ở lô hàng A	2	6	12	4	1
Số quả cam ở lô hàng B	1	3	7	10	4

- a) Hãy ước lượng cân nặng trung bình của mỗi quả cam ở lô hàng A và B
b) Nếu so sánh theo số trung bình thì cam ở lô hàng nào nặng hơn?

Bài tập 3: Cân nặng của một số quả mít trong một khu vườn được thống kê ở bảng sau:

Cân nặng (kg)	$[4; 6)$	$[6; 8)$	$[8; 10)$	$[10; 12)$	$[12; 14)$
Số quả mít	6	12	19	9	4

Hãy tính phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm trên. (Kết quả các phép tính làm tròn đến hàng phần trăm.)



Bài tập 4: Thống kê tổng số giờ nắng trong tháng 9 tại một trạm quan trắc đặt ở Cà Mau trong các năm từ 2002 đến 2021 được thống kê như sau:

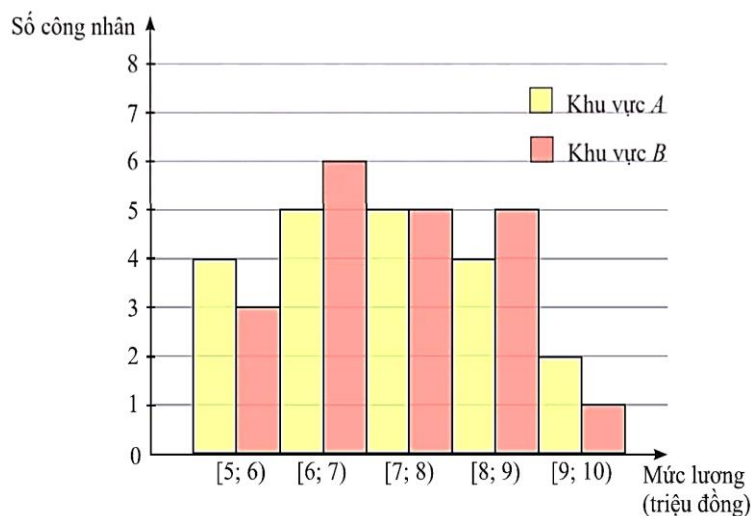
111,6	134,9	130,3	134,2	140,9	109,3	154,4	156,3	116,1	96,7
105,2	80,8	80,8	110	109	139	145	161	126	114

- a) Hãy tính phương sai và độ lệch chuẩn của mẫu số liệu trên.
- b) Hãy lập bảng tần số ghép nhóm với nhóm đầu tiên là $[80;98)$ và độ dài mỗi nhóm bằng 18. Tính phương sai, độ lệch chuẩn của mẫu số liệu ghép nhóm.
- c) Hãy tính sai số tương đối của độ lệch chuẩn của mẫu số liệu ghép nhóm so với độ lệch chuẩn của mẫu số liệu gốc.
- (Kết quả các phép tính làm tròn đến hàng phần nghìn.)

Bài tập 5: Biểu đồ dưới đây mô tả kết quả điều tra về mức lương khởi điểm (đơn vị: triệu đồng) của một số công nhân ở hai khu vực A và B.

- a) Hãy xác định giá trị đại diện cho mỗi nhóm và lập bảng tần số ghép nhóm cho mẫu số liệu đó.
- b) Nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì công nhân ở khu vực nào có mức lương khởi điểm đồng đều hơn?

Mức lương khởi điểm của công nhân ở hai khu vực A và B



Bài tập 6: Giá đóng cửa của một cổ phiếu là giá của cổ phiếu đó cuối một phiên giao dịch. Bảng sau thống kê giá đóng cửa (đơn vị: nghìn đồng) của hai mã cổ phiếu A và B trong 50 ngày giao dịch liên tiếp.

Giá đóng cửa	$[120;122)$	$[122;124)$	$[124;126)$	$[126;128)$	$[128;130)$
Cổ phiếu A	8	9	12	10	11
Cổ phiếu B	16	4	3	6	21

Người ta có thể dùng phương sai và độ lệch chuẩn để so sánh mức độ rủi ro của các loại cổ phiếu có giá trị trung bình gần bằng nhau. Cổ phiếu nào có phương sai, độ lệch chuẩn cao hơn thì được coi là có độ rủi ro lớn hơn.

Theo quan điểm trên, hãy so sánh độ rủi ro của cổ phiếu A và cổ phiếu B.



Bài tập 7: Mai và Ngọc cùng sử dụng vòng đeo tay thông minh để ghi lại số bước chân hai bạn đi mỗi ngày trong một tháng. Kết quả được ghi lại ở bảng sau:

Giá trị đại diện	4	6	8	10	12
Mai	6	7	6	6	5
Ngọc	2	5	10	8	2

- a) Hãy tính số trung bình và độ lệch chuẩn của mẫu số liệu ghép nhóm trên.
b) Nếu so sánh theo độ lệch chuẩn thì bạn nào có số lượng bước chân đi mỗi ngày đều đặn hơn?

Bài tập 8: Bảng dưới đây thống kê cự li ném tạ của một vận động viên.

Cự li (m)	[19;19,5)	[19,5;20)	[20;20,5)	[20,5;21)	[21;21,5)
Tần số	13	45	24	12	6

Hãy tính phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm trên.

BÀI TẬP TRẮC NGHIỆM

PHẦN I. Câu trắc nghiệm nhiều phương án lựa chọn. Mỗi câu hỏi thí sinh chỉ chọn một phương án.

- Câu 1:** Số đặc trưng nào không sử dụng thông tin của nhóm số liệu đầu tiên và nhóm số liệu cuối cùng?
A. Khoảng biến thiên. B. Khoảng tứ phân vị.
C. Phương sai. D. Độ lệch chuẩn.
- Câu 2:** Nếu thay tất cả các tần số trong mẫu số liệu ghép nhóm trên bằng 4 thì số đặc trưng nào sau đây không thay đổi?
A. Khoảng biến thiên. B. Khoảng tứ phân vị.
C. Phương sai. D. Độ lệch chuẩn.
- Câu 3:** Trong các khẳng định sau, khẳng định nào sai?
A. Phương sai luôn luôn là số không âm.
B. Phương sai là bình phương của độ lệch chuẩn.
C. Phương sai càng lớn thì độ phân tán của các giá trị quanh số trung bình càng lớn.
D. Phương sai luôn luôn lớn hơn độ lệch chuẩn.
- Câu 4:** Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một cửa hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

Doanh thu	[5;7)	[7;9)	[9;11)	[11;13)	[13;15)
Số ngày	2	7	7	3	1

Số trung bình của mẫu số liệu trên thuộc khoảng nào trong các khoảng dưới đây?

- A. [7;9). B. [9;11). C. [11;13). D. [13;15).
- Câu 5:** Một siêu thị thống kê số tiền (đơn vị: chục nghìn đồng) mà 44 khách hàng mua hàng ở siêu thị đó trong một ngày. Số liệu được ghi lại trong bảng dưới đây.



Nhóm	Giá trị đại diện	Tần số
[40; 45)	42,5	4
[45; 50)	47,5	14
[50; 55)	52,5	8
[55; 60)	57,5	10
[60; 75)	62,5	6
[65; 70)	67,5	2
		$n = 44$

Phương sai của mẫu số liệu ghép nhóm trên là:

- A. 53,2. B. 46,1. C. 30. D. 11.

Câu 6: Một siêu thị thống kê số tiền (đơn vị: chục nghìn đồng) mà 44 khách hàng mua hàng ở siêu thị đó trong một ngày. Số liệu được ghi lại trong bảng dưới đây.

Nhóm	Giá trị đại diện	Tần số
[40; 45)	42,5	4
[45; 50)	47,5	14
[50; 55)	52,5	8
[55; 60)	57,5	10
[60; 75)	62,5	6
[65; 70)	67,5	2
		$n = 44$

Độ lệch chuẩn của mẫu số liệu ghép nhóm trên (làm tròn kết quả đến hàng đơn vị) là:

- A. 6,8. B. 7,3. C. 3,3. D. 46,1.

Câu 7: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khỏe. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

Quãng đường (km)	[2, 7; 3, 0)	[3, 0; 3, 3)	[3, 3; 3, 6)	[3, 6; 3, 9)	[3, 9; 4, 2)
Số ngày	3	6	5	4	2

Phương sai của mẫu số liệu ghép nhóm là

- A. 3,39. B. 11,62. C. 0,1314. D. 0,36.

Câu 8: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khỏe. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

Quãng đường (km)	[2, 7; 3, 0)	[3, 0; 3, 3)	[3, 3; 3, 6)	[3, 6; 3, 9)	[3, 9; 4, 2)
Số ngày	3	6	5	4	2



Độ lệch chuẩn của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

- A. 3,41.
- B. 11,62.
- C. 0,017.
- D. 0,36.

Câu 9: Bạn Chi rất thích nhảy hiện đại. Thời gian tập nhảy mỗi ngày trong thời gian gần đây của bạn Chi được thống kê lại ở bảng sau:

Thời gian (phút)	[20;25)	[25;30)	[30;35)	[35;40)	[40;45)
Số ngày	6	6	4	1	1

Phương sai của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

- A. 31,77.
- B. 32.
- C. 31.
- D. 31,44.

Câu 10: Dũng là học sinh rất giỏi chơi rubik, bạn có thể giải nhiều loại khối rubik khác nhau. Trong một lần tập luyện giải khối rubik 3×3 , bạn Dũng đã tự thống kê lại thời gian giải rubik trong 25 lần giải liên tiếp ở bảng sau:

Thời gian giải rubik (giây)	[8;10)	[10;12)	[12;14)	[14;16)	[16;18)
Số lần	4	6	8	4	3

Độ lệch chuẩn của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

- A. 5,98.
- B. 6.
- C. 2,44.
- D. 2,5.

Câu 11: Để đánh giá chất lượng của một loại pin điện thoại mới, người ta ghi lại thời gian nghe nhạc liên tục của điện thoại được sạc đầy pin cho đến khi hết pin cho kết quả sau:

Thời gian (giờ)	[5;5,5)	[5,5;6)	[6;6,5)	[6,5;7)	[7;7,5)
Số chiếc điện thoại (tần số)	2	8	15	10	5

Tính độ lệch chuẩn của mẫu số liệu ghép nhóm trên (làm tròn đến 4 chữ số thập phân)

- A. 0,4252 .
- B. 0,5314 .
- C. 0,6214 .
- D. 0,5268 .

Câu 12: Người ta ghi lại tiền lãi (đơn vị: triệu đồng) của một số nhà đầu tư (với số tiền đầu tư như nhau), khi đầu tư vào hai lĩnh vực A, B cho kết quả như sau:

Tiền lãi	[5;10)	[10;15)	[15;20)	[20;25)	[25;30)
Số nhà đầu tư vào lĩnh vực A	2	5	8	6	4
Số nhà đầu tư vào lĩnh vực B	8	4	2	5	6

Tính độ lệch chuẩn cho các mẫu số liệu về tiền lãi của các nhà đầu tư ở lĩnh vực B

- A. 8,039 .
- B. 8,5314 .
- C. 8,6214 .
- D. 8,5268 .

**PHẦN II. Câu trắc nghiệm đúng sai. Trong mỗi ý a), b), c), d) ở mỗi câu, thí sinh chọn đúng hoặc sai.**

Câu 1: Kết quả 40 lần nhảy xa của hai vận động viên Dũng và Huy được lần lượt thống kê trong bảng bên dưới (đơn vị: mét)

Nhóm	Dũng	Huy
[6, 22; 6, 46)	3	2
[6, 46; 6, 70)	7	5
[6, 70; 6, 94)	5	8
[6, 94; 7, 18)	20	19
[7, 18; 7, 42)	5	6
	$n = 40$	$n = 40$

Xét tính đúng sai của các khẳng định sau:

- a) Số trung bình cộng của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Dũng cho bởi Bảng 15 (làm tròn kết quả đến hàng phần trăm) là 6,92 m.
- b) Độ lệch chuẩn của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Dũng cho bởi Bảng 15 (làm tròn kết quả đến hàng phần trăm) là 0,26 m.
- c) Phương sai của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Huy cho bởi Bảng 16 (làm tròn kết quả đến hàng phần trăm) là 0,16.
- d) Kết quả nhảy xa của vận động viên Dũng đồng đều hơn kết quả nhảy xa của vận động viên Huy.

Câu 2: Mẫu số liệu ghép nhóm thống kê mức lương của hai công ty A, B (đơn vị: triệu đồng) được thể hiện như bảng dưới đây

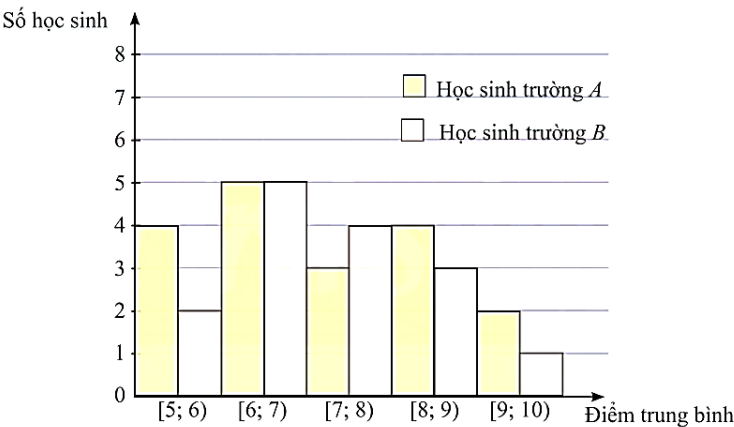
Nhóm	Công ty A	Công ty B
[10;15)	15	25
[15;20)	18	15
[20;25)	10	7
[25;30)	10	5
[30;35)	5	5
[35;40)	2	3
	$n = 60$	$n = 60$

Xét tính đúng sai của các khẳng định sau:

- a) Số trung bình cộng của mẫu số liệu ghép nhóm của công ty A là: $\frac{62}{3}$
- b) Độ lệch chuẩn của mẫu số liệu ghép nhóm của công ty A là: 5
- c) Phương sai của mẫu số liệu ghép nhóm của công ty B (làm tròn kết quả đến hàng phần trăm) là 52,91.
- b) Nhận thấy độ lệch chuẩn của công ty A nhỏ hơn công ty B nên mức lương của công ty A đồng đều hơn.



Câu 3: Biểu đồ sau mô tả kết quả điều tra về điểm trung bình năm học của học sinh hai trường A và B



Người ta lập được bảng tần số ghép nhóm cho mẫu số liệu trên như sau:

Điểm trung bình	[5; 6)	[6; 7)	[7; 8)	[8; 9)	[9; 10)
Giá trị đại diện	5,5	6,5	7,5	8,5	9,5
Học sinh trường A	4	5	3	4	2
Học sinh trường B	2	5	4	3	1

Xét tính đúng sai của các khẳng định sau:

- a) Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm của học sinh trường A là: 6,1
- b) Khoảng tứ phân vị của mẫu số liệu ghép nhóm của học sinh trường B là: 1,73
- c) Nếu so sánh theo khoảng tứ phân vị của mẫu số liệu ghép nhóm thì học sinh trường B có điểm trung bình đồng đều hơn.
- d) Nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì học sinh trường A có điểm trung bình đồng đều hơn.

Câu 4: Bảng dưới đây biểu diễn mẫu số liệu ghép nhóm về nhiệt độ không khí trung bình các tháng năm 2021 tại Hà Nội và Huế (đơn vị: độ C).

Nhóm	Hà Nội	Huế
[16,8;19,8)	2	1
[19,8;22,8)	3	2
[22,8;25,8)	2	3
[25,8;28,8)	1	2
[28,8;31,8)	4	4
	$n = 12$	$n = 12$

Xét tính đúng sai của các khẳng định sau:



- a) Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên số liệu ở Hà Nội là: 8,75
- b) Độ lệch chuẩn của mẫu số liệu ghép nhóm trên số liệu ở Hà Nội (làm tròn kết quả đến hàng phần trăm) là: 3,56
- c) Phương sai của mẫu số liệu ghép nhóm trên số liệu ở Huế là: 12,75
- d) Huế có nhiệt độ không khí trung bình tháng đồng đều hơn vì độ lệch chuẩn nhỏ hơn

Câu 5: Bảng dưới đây thống kê độ ẩm không khí trung bình các tháng năm 2021 tại Đà Lạt và Vũng Tàu (đơn vị: %).

Tháng	1	2	3	4	5	6	7	8	9	10	11	12
Đà Lạt	83	79	79	87	87	87	88	89	90	91	88	86
Vũng Tàu	75	77	78	77	79	79	81	79	81	83	80	77

Người ta lần lượt ghép các số liệu của Đà Lạt, Vũng Tàu thành mẫu số liệu ghép nhóm như sau:

Đà Lạt			Vũng Tàu		
Nhóm	Giá trị đại diện	Tần số	Nhóm	Giá trị đại diện	Tần số
[75; 78,3)	76,65	0	[75; 78,3)	76,65	5
[78,3; 81,6)	79,95	2	[78,3; 81,6)	79,95	6
[81,6; 84,9)	83,25	1	[81,6; 84,9)	83,25	1
[84,9; 88,2)	86,55	6	[84,9; 88,2)	86,55	0
[88,2; 91,5)	89,85	3	[88,2; 91,5)	89,85	0

Xét tính đúng sai của các khẳng định sau:

- a) Xét số liệu ở Đà Lạt ta có khoảng biến thiên là 16,5.
- b) Xét số liệu ở Đà Lạt thì ta có độ lệch chuẩn của mẫu số liệu ghép nhóm (làm tròn kết quả đến hàng phần trăm) là 3,28.
- c) Xét số liệu ở Vũng Tàu thì phương sai của mẫu số liệu ghép nhóm là 4,235.
- d) Đà Lạt có nhiệt độ không khí trung bình tháng đồng đều hơn vì độ lệch chuẩn nhỏ hơn

Câu 6: Bảng sau thống kê lại tổng số giờ nắng trong tháng 6 của các năm từ 2002 đến 2021 tại hai trạm quan trắc đặt ở Nha Trang và Quy Nhơn.

Số giờ nắng	[130;160)	[160;190)	[190;220)	[220;250)	[250;280)	[280;310)
Số năm ở Nha Trang	1	1	1	8	7	2
Số năm ở Quy Nhơn	0	1	2	4	10	3

Xét tính đúng sai của các khẳng định sau:

- a) Xét số liệu ở Nha Trang thì khoảng tứ phân vị của mẫu số liệu ghép nhóm là: 32,64
- b) Nếu so sánh theo khoảng tứ phân vị thì số giờ nắng trong tháng 6 của Quy Nhơn đồng đều hơn



c) Xét số liệu của Quy Nhơn ta có độ lệch chuẩn của mẫu số liệu ghép nhóm (làm tròn kết quả đến hàng phần trăm) là: 30,59

d) Nếu so sánh theo độ lệch chuẩn thì số giờ nắng trong tháng 6 của Nha Trang đồng đều hơn

Câu 7: Thống kê tổng số giờ nắng trong tháng 9 tại một trạm quan trắc đặt ở Cà Mau trong các năm từ 2002 đến 2021 được thống kê như sau:

111,6	134,9	130,3	134,2	140,9	109,3	152,4	156,3	116,1	96,7
105,2	80,8	80,8	110	109	139	145	161	126	114

Người ta lập được bảng tần số ghép nhóm như sau:

Số giờ nắng	[80;98)	[98;116)	[116;134)	[134;152)	[152;170)
Giá trị đại diện	89	107	125	143	161
Số năm	3	6	3	5	3

Xét tính đúng sai của các khẳng định sau:

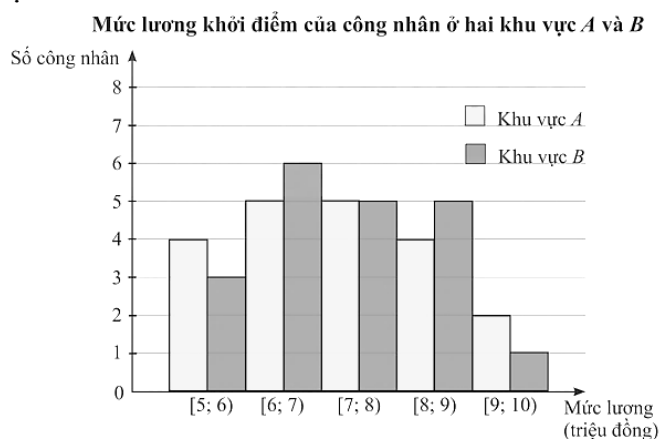
a) Số trung bình của mẫu số liệu ghép nhóm là 124,1.

b) Phương sai của mẫu số liệu ghép nhóm là 566,19.

c) Độ lệch chuẩn của mẫu số liệu ghép nhóm (kết quả các phép tính làm tròn đến hàng phần nghìn) là 23,795.

d) Sai số tương đối của độ lệch chuẩn của mẫu số liệu ghép nhóm so với độ lệch chuẩn của mẫu số liệu gốc (kết quả các phép tính làm tròn đến hàng phần nghìn) là 4,805%.

Câu 8: Biểu đồ dưới đây mô tả kết quả điều tra về mức lương khởi điểm (đơn vị: triệu đồng) của một số công nhân ở hai khu vực A và B.



Người ta lập được bảng tần số ghép nhóm cho mẫu số liệu như sau:

Mức lương	[5;6)	[6;7)	[7;8)	[8;9)	[9;10)
Mức lương đại diện (triệu đồng)	5,5	6,5	7,5	8,5	9,5
Khu vực A	4	5	5	4	2
Khu vực B	3	6	5	5	1



Xét tính đúng sai của các khẳng định sau:

- a) Xét mẫu số liệu của khu vực A ta có số trung bình của mẫu số liệu ghép nhóm là 6,25.
- b) Xét mẫu số liệu của khu vực A ta có độ lệch chuẩn của mẫu số liệu ghép nhóm là $\sqrt{1,5875}$
- c) Xét mẫu số liệu của khu vực B ta có phương sai của mẫu số liệu ghép nhóm là 1,3875.
- d) Nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì mức lương khởi điểm của công nhân khu vực B đồng đều hơn của công nhân khu vực A.

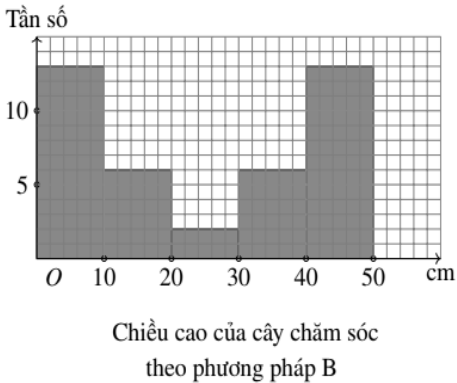
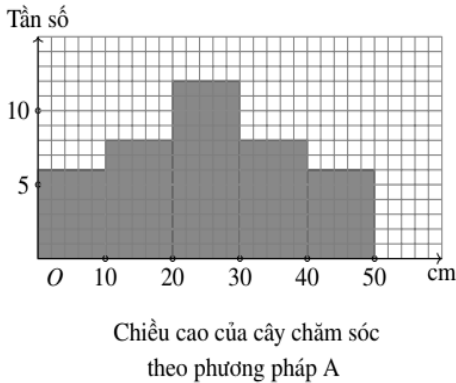
Câu 9: Giá đóng cửa của một cổ phiếu là giá của cổ phiếu đó cuối một phiên giao dịch. Bảng sau thống kê giá đóng cửa (đơn vị: nghìn đồng) của hai mã cổ phiếu A và B trong 50 ngày giao dịch liên tiếp.

Giá đóng cửa	[120;122)	[122;124)	[124;126)	[126;128)	[128;130)
Cổ phiếu A	8	9	12	10	11
Cổ phiếu B	16	4	3	6	21

Xét tính đúng sai của các khẳng định sau:

- a) Xét mẫu số liệu của cổ phiếu A ta có phương sai của mẫu số liệu ghép nhóm là 7,5216
- b) Xét mẫu số liệu của cổ phiếu B ta có số trung bình của mẫu số liệu ghép nhóm là 115,28.
- c) Xét mẫu số liệu của cổ phiếu B ta có độ lệch chuẩn của mẫu số liệu ghép nhóm là $\sqrt{15,4096}$
- d) Người ta có thể dùng phương sai và độ lệch chuẩn để so sánh mức độ rủi ro của các loại cổ phiếu có giá trị trung bình gần bằng nhau. Cổ phiếu nào có phương sai, độ lệch chuẩn cao hơn thì được coi là có độ rủi ro lớn hơn. Theo quan điểm trên, thì cổ phiếu A có độ rủi ro thấp hơn cổ phiếu B.

Câu 10: Một công ty giống cây trồng đã thử nghiệm hai phương pháp chăm sóc khác nhau cho cây hướng dương. Sau hai tuần, người ta thấy cây được chăm sóc theo cả hai phương pháp đều thấp hơn 50 cm.



- a) Khoảng biến thiên của chiều cao các cây được chăm sóc theo mỗi phương pháp A và B bằng nhau.
- b) Trung bình của chiều cao các cây được chăm sóc theo mỗi phương pháp A và B bằng nhau.
- c) Độ lệch chuẩn của chiều cao các cây được chăm sóc theo phương án A là 12,65(cm).
- d) Dựa vào độ lệch chuẩn thì chiều cao của các loại cây được chăm sóc theo phương án B ít bị chênh lệch hơn so với phương án A.

**PHẦN III. Câu trắc nghiệm trả lời ngắn**

Câu 1: Hãy tính độ lệch chuẩn của mẫu số liệu ghép nhóm (kết quả được làm tròn đến hàng phần trăm)

Chiều cao (cm)	[160;164)	[164;168)	[168;172)	[172;176)	[176;180)
Số học sinh	3	5	8	4	1

Câu 2: Để đánh giá chất lượng một loại pin điện thoại mới, người ta ghi lại thời gian nghe nhạc liên tục của điện thoại được sạc đầy pin cho đến khi hết pin cho kết quả như sau:

Thời gian (giờ)	[5;5,5)	[5,5;6)	[6;6,5)	[6,5;7)	[7;7,5)
Số chiếc điện thoại (tần số)	2	8	15	10	5

Tính độ lệch chuẩn của mẫu số liệu ghép nhóm trên (kết quả được làm tròn đến hàng phần trăm)

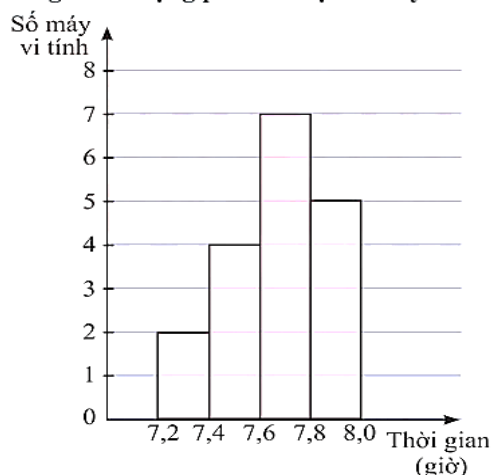
Câu 3: Bảng dưới đây thống kê cự li ném tạ của một vận động viên.

Cự li (m)	[19;19,5)	[19,5;20)	[20;20,5)	[20,5;21)	[21;21,5)
Tần số	13	45	24	12	6

Hãy tính độ lệch chuẩn của mẫu số liệu ghép nhóm trên (kết quả được làm tròn đến hàng phần trăm)

Câu 4: Kết quả khảo sát thời gian sử dụng liên tục (đơn vị: giờ) từ lúc sạc đầy cho đến khi hết của pin một số máy vi tính cùng loại được mô tả bằng biểu đồ bên.

Thời gian sử dụng pin của một số máy vi tính



Hãy xác định độ lệch chuẩn của thời gian sử dụng pin (kết quả được làm tròn đến hàng phần trăm)

Câu 5: Thành tích môn nhảy cao của các vận động viên tại một giải điền kinh dành cho học sinh trung học phổ thông như sau:

Mức xà (cm)	[170;172)	[172;174)	[174;176)	[176;180)
Số vận động viên	3	10	6	1

Hãy xác định độ lệch chuẩn của thời gian sử dụng pin (kết quả được làm tròn đến hàng phần trăm)



Câu 6: Chiều dài của 40 bé trai sơ sinh 12 ngày tuổi chọn ngẫu nhiên ở một bệnh viện được nhà nghiên cứu thống kê trong bảng dưới đây:

Chiều dài (cm)	[44; 46)	[46; 48)	[48; 50)	[50; 52)	[52; 54)	[54; 56)
Số trẻ	3	3	10	15	7	2

Tính độ lệch chuẩn của chiều dài nhóm 40 bé trai sơ sinh (làm tròn kết quả đến hàng phần nghìn).

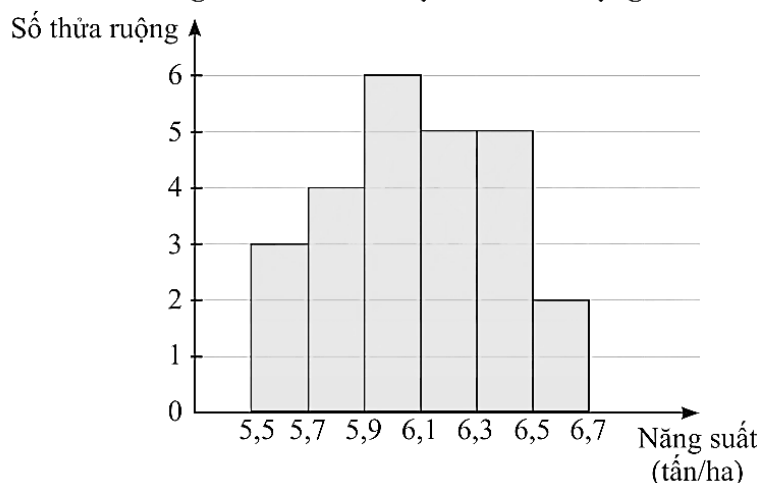
Câu 7: Một bác tài xế thống kê lại độ dài quãng đường (đơn vị: km) bác đã lái xe mỗi ngày trong một tháng ở bảng sau:

Độ dài quãng đường (km)	[50;100)	[100;150)	[150;200)	[200;250)	[250;300)
Số ngày	5	10	9	9	2

Hãy xác định độ lệch chuẩn của mẫu số liệu trên.

Câu 8: Kết quả khảo sát năng suất (đơn vị: tấn/ha) của một số thửa ruộng được minh họa ở biểu đồ sau.

Năng suất lúa của một số thửa ruộng



Người ta lập được bảng tần số ghép nhóm và tần số tương đối ghép nhóm tương ứng của mẫu số liệu trên như sau:

Năng suất	[5,5; 5,7)	[5,7; 5,9)	[5,9; 6,1)	[6,1; 6,3)	[6,3; 6,5)	[6,5; 6,7)
Số thửa ruộng	3	4	6	5	5	2
Giá trị đại diện	5,6	5,8	6,0	6,2	6,4	6,6
Tần số tương đối	3	4	6	5	5	2

Hãy xác định độ lệch chuẩn của mẫu số liệu trên (làm tròn kết quả đến hàng phần nghìn).

Câu 9: Một vận động viên luyện tập chạy cự li 100 m đã ghi lại kết quả luyện tập như sau:

Thời gian (giây)	[10,2; 10,4)	[10,4; 10,6)	[10,6; 10,8)	[10,8; 11)
Số vận động viên	3	7	8	2

Hãy xác định phương sai của mẫu số liệu trên?



Câu 10: Cân nặng của một số quả mít trong một khu vườn được thống kê ở bảng sau:

Cân nặng (kg)	[4; 6)	[6; 8)	[8;10)	[10;12)	[12;14)
Số quả mít	6	12	19	9	4

Hãy tính phương sai của mẫu số liệu ghép nhóm trên. (Kết quả các phép tính làm tròn đến hàng phần trăm)?

-----HẾT-----



**BÀI 02 PHƯƠNG SAI VÀ ĐỘ LỆCH CHUẨN CỦA MSLGN****A LÝ THUYẾT CẦN NHỚ**

Xét mẫu số liệu ghép nhóm cho bởi công thức sau:

Nhóm	$[u_1; u_2)$	$[u_1; u_2)$	...	$[u_k; u_{k+1})$
Giá trị đại diện	c_1	c_2	...	c_k
Tần số	n_1	n_2	...	n_k

Phương sai: Phương sai của mẫu số liệu ghép nhóm, kí hiệu S^2 , được tính bởi công thức

$$S^2 = \frac{1}{n} \left[n_1 (c_1 - \bar{x})^2 + n_2 (c_2 - \bar{x})^2 + \dots + n_k (c_k - \bar{x})^2 \right]$$

Trong đó: $n = n_1 + n_2 + \dots + n_k$ là cỡ mẫu; $\bar{x} = \frac{1}{n} (n_1 c_1 + n_2 c_2 + \dots + n_k c_k)$

Độ lệch chuẩn: Độ lệch chuẩn của mẫu số liệu ghép nhóm, kí hiệu S , là căn bậc hai số học của phương sai, nghĩa là $S = \sqrt{S^2}$

Ý nghĩa:

- Phương sai (độ lệch chuẩn) của mẫu số liệu ghép nhóm là giá trị xấp xỉ cho phương sai (độ lệch chuẩn) của mẫu số liệu gốc. Chúng được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm xung quanh số trung bình của mẫu số liệu. Phương sai và độ lệch chuẩn càng lớn thì dữ liệu càng phân tán.
- Độ lệch chuẩn có cùng đơn vị với đơn vị của mẫu số liệu.

Chú ý:

- a) Phương sai của mẫu số liệu ghép nhóm có thể được tính theo công thức sau:

$$S^2 = \frac{1}{n} (n_1 c_1^2 + n_2 c_2^2 + \dots + n_k c_k^2) - \bar{x}^2$$

- b) Trong thống kê, người ta còn dùng đại lượng sau để đo mức độ phân tán của mẫu số liệu ghép nhóm:

$$\hat{s}^2 = \frac{1}{n-1} \left[n_1 (c_1 - \bar{x})^2 + n_2 (c_2 - \bar{x})^2 + \dots + n_k (c_k - \bar{x})^2 \right]$$

**B //** PHÂN LOẠI VÀ PHƯƠNG PHÁP GIẢI TOÁN**Dạng 1: Tính phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm**

Xét mẫu số liệu ghép nhóm cho bởi công thức sau:

Nhóm	$[u_1; u_2)$	$[u_1; u_2)$	...	$[u_k; u_{k+1})$
Giá trị đại diện	c_1	c_2	...	c_k
Tần số	n_1	n_2	...	n_k

Số trung bình cộng của mẫu số liệu ghép nhóm trên được tính bằng công thức:

$$\bar{x} = \frac{1}{n}(n_1c_1 + n_2c_2 + \dots + n_kc_k)$$

Phương sai: Phương sai của mẫu số liệu ghép nhóm, kí hiệu S^2 , được tính bởi công thức

$$S^2 = \frac{1}{n} \left[n_1(c_1 - \bar{x})^2 + n_2(c_2 - \bar{x})^2 + \dots + n_k(c_k - \bar{x})^2 \right]$$

Trong đó: $n = n_1 + n_2 + \dots + n_k$ là cỡ mẫu; $\bar{x} = \frac{1}{n}(n_1c_1 + n_2c_2 + \dots + n_kc_k)$

Độ lệch chuẩn: Độ lệch chuẩn của mẫu số liệu ghép nhóm, kí hiệu S , là căn bậc hai số học của phương sai, nghĩa là $S = \sqrt{S^2}$

BÀI TẬP TỰ LUẬN

Bài tập 1: Tìm cân nặng trung bình của học sinh lớp 11D cho trong bảng sau

Cân nặng	$[40,5; 45,5)$	$[45,5; 50,5)$	$[50,5; 55,5)$	$[55,5; 60,5)$	$[60,5; 65,5)$	$[65,5; 70,5)$
Số học sinh	10	7	16	4	2	3

Lời giải

Trong mỗi khoảng cân nặng, giá trị đại diện là trung bình cộng của hai giá trị đầu mút nên ta có bảng sau:

Cân nặng (kg)	43	48	53	58	63	68
Số học sinh	10	7	16	4	2	3

Tổng số học sinh là $n = 42$. Cân nặng trung bình củ lớp 11D là:

$$\bar{x} = \frac{10.43 + 7.48 + 16.53 + 4.58 + 2.63 + 3.68}{42} \approx 51,81(\text{kg})$$

Bài tập 2: Kết quả khảo sát cân nặng của 25 quả cam ở mỗi lô hàng A và B được cho bởi bảng sau:

Cân nặng (g)	$[150; 155)$	$[155; 160)$	$[160; 165)$	$[165; 170)$	$[170; 175)$
Số quả cam ở lô hàng A	2	6	12	4	1
Số quả cam ở lô hàng B	1	3	7	10	4

a) Hãy ước lượng cân nặng trung bình của mỗi quả cam ở lô hàng A và B



b) Nếu so sánh theo số trung bình thì cam ở lô hàng nào nặng hơn?

Lời giải

Ta có bảng thống kê số lượng cam theo giá trị đại diện:

Cân nặng (g)	152,5	157,5	162,5	167,5	172,5
Số quả cam ở lô hàng A	2	6	12	4	1
Số quả cam ở lô hàng B	1	3	7	10	4

a) Cân nặng trung bình của một quả cam ở lô hàng A là xấp xỉ bằng:

$$\frac{(2.152,5 + 6.157,5 + 12.162,5 + 4.167,5 + 1.172,5)}{25} = 161,7 \text{ (g)}$$

Cân nặng trung bình của một quả cam ở lô hàng B là xấp xỉ bằng:

$$\frac{(1.152,2 + 3.157,5 + 7.162,5 + 10.167,5 + 4.172,5)}{25} = 165,1 \text{ (g)}$$

b) Nếu so sánh theo số trung bình thì cam ở lô hàng B nặng hơn cam ở lô hàng A.

Bài tập 3: Cân nặng của một số quả mít trong một khu vườn được thống kê ở bảng sau:

Cân nặng (kg)	[4;6)	[6;8)	[8;10)	[10;12)	[12;14)
Số quả mít	6	12	19	9	4

Hãy tính phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm trên. (Kết quả các phép tính làm tròn đến hàng phần trăm.)

Lời giải

Ta có bảng thống kê cân nặng của các quả mít theo giá trị đại diện:

Cân nặng đại diện (kg)	5	7	9	11	13
Số quả mít	6	12	19	9	4

Cỡ mẫu $n = 6 + 12 + 19 + 9 + 4 = 50$.

Số trung bình của mẫu số liệu ghép nhóm là: $\bar{x} = \frac{6.5 + 12.7 + 19.9 + 9.11 + 4.13}{50} = 8,72$.

Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{50} (6 \cdot 5^2 + 12 \cdot 7^2 + 19 \cdot 9^2 + 9 \cdot 11^2 + 4 \cdot 13^2) - 8,72^2 \approx 4,80$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S \approx \sqrt{4,80} \approx 2,19$.

Bài tập 4: Thống kê tổng số giờ nắng trong tháng 9 tại một trạm quan trắc đặt ở Cà Mau trong các năm từ 2002 đến 2021 được thống kê như sau:

111,6	134,9	130,3	134,2	140,9	109,3	154,4	156,3	116,1	96,7
105,2	80,8	80,8	110	109	139	145	161	126	114

a) Hãy tính phương sai và độ lệch chuẩn của mẫu số liệu trên.



- b) Hãy lập bảng tần số ghép nhóm với nhóm đầu tiên là $[80;98)$ và độ dài mỗi nhóm bằng 18. Tính phương sai, độ lệch chuẩn của mẫu số liệu ghép nhóm.
- c) Hãy tính sai số tương đối của độ lệch chuẩn của mẫu số liệu ghép nhóm so với độ lệch chuẩn của mẫu số liệu gốc.
- (Kết quả các phép tính làm tròn đến hàng phần nghìn.)

Lời giải

a) Cỡ mẫu là $n = 20$.

Số trung bình của mẫu số liệu trên là: $\bar{x}_1 = \frac{111,6 + 134,9 + \dots + 114}{20} = 122,755$.

Phương sai của mẫu số liệu trên là: $S_1^2 = \frac{1}{20}(111,6^2 + 134,9^2 + \dots + 114^2) - 122,755^2 \approx 515,453$.

Độ lệch chuẩn của mẫu số liệu trên là: $S_1 \approx \sqrt{515,453} \approx 22,704$.

b) Ta có bảng sau:

Số giờ nắng	$[80;98)$	$[98;116)$	$[116;134)$	$[134;152)$	$[152;170)$
Giá trị đại diện	89	107	125	143	161
Số năm	3	6	3	5	3

Số trung bình của mẫu số liệu ghép nhóm là: $\bar{x}_2 = \frac{3.89 + 6.107 + 3.125 + 5.143 + 3.161}{20} = 124,1$.

Phương sai của mẫu số liệu ghép nhóm là

$$S_2^2 = \frac{1}{20}(3.89^2 + 6.107^2 + 3.125^2 + 5.143^2 + 3.161^2) - 124,1^2 = 566,19.$$

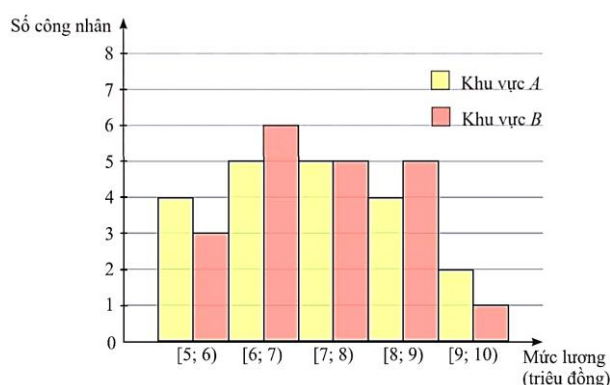
Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S_2 = \sqrt{566,19} \approx 23,795$.

c) Sai số tương đối của độ lệch chuẩn của mẫu số liệu ghép nhóm so với độ lệch chuẩn của mẫu số liệu gốc là: $\frac{|S_2 - S_1|}{S_1} = \frac{|23,795 - 22,704|}{22,704} \cdot 100\% \approx 4,805\%$.

Bài tập 5: Biểu đồ dưới đây mô tả kết quả điều tra về mức lương khởi điểm (đơn vị: triệu đồng) của một số công nhân ở hai khu vực A và B.

- a) Hãy xác định giá trị đại diện cho mỗi nhóm và lập bảng tần số ghép nhóm cho mẫu số liệu đó.
- b) Nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì công nhân ở khu vực nào có mức lương khởi điểm đồng đều hơn?

Mức lương khởi điểm của công nhân ở hai khu vực A và B





Lời giải

a) Ta có bảng sau:

Mức lương	[5;6)	[6;7)	[7;8)	[8;9)	[9;10)
Mức lương đại diện (triệu đồng)	5,5	6,5	7,5	8,5	9,5
Khu vực A	4	5	5	4	2
Khu vực B	3	6	5	5	1

b) Xét mẫu số liệu của khu vực A :

Cỡ mẫu là $n_A = 4 + 5 + 5 + 4 + 2 = 20$.

Số trung bình của mẫu số liệu ghép nhóm là: $\bar{x}_A = \frac{4 \cdot 5,5 + 5 \cdot 6,5 + 5 \cdot 7,5 + 4 \cdot 8,5 + 2 \cdot 9,5}{20} = 7,25$.

Phương sai của mẫu số liệu ghép nhóm là

$$S_A^2 = \frac{1}{20} (4 \cdot 5,5^2 + 5 \cdot 6,5^2 + 5 \cdot 7,5^2 + 4 \cdot 8,5^2 + 2 \cdot 9,5^2) - (7,25)^2 = 1,5875.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S_A = \sqrt{1,5875}$.

Xét mẫu số liệu của khu vực B :

Cỡ mẫu là $n_B = 3 + 6 + 5 + 5 + 1 = 20$.

Số trung bình của mẫu số liệu ghép nhóm là: $\bar{x}_B = \frac{3 \cdot 5,5 + 6 \cdot 6,5 + 5 \cdot 7,5 + 5 \cdot 8,5 + 1 \cdot 9,5}{20} = 7,25$.

Phương sai của mẫu số liệu ghép nhóm là

$$S_B^2 = \frac{1}{20} (3 \cdot 5,5^2 + 6 \cdot 6,5^2 + 5 \cdot 7,5^2 + 5 \cdot 8,5^2 + 1 \cdot 9,5^2) - (7,25)^2 = 1,2875.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S_B = \sqrt{1,2875}$

Do $S_A > S_B$ nên nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì mức lương khởi điểm của công nhân khu vực B đồng đều hơn của công nhân khu vực A.

Bài tập 6: Giá đóng cửa của một cổ phiếu là giá của cổ phiếu đó cuối một phiên giao dịch. Bảng sau thống kê giá đóng cửa (đơn vị: nghìn đồng) của hai mã cổ phiếu A và B trong 50 ngày giao dịch liên tiếp.

Giá đóng cửa	[120;122)	[122;124)	[124;126)	[126;128)	[128;130)
Cổ phiếu A	8	9	12	10	11
Cổ phiếu B	16	4	3	6	21

Người ta có thể dùng phương sai và độ lệch chuẩn để so sánh mức độ rủi ro của các loại cổ phiếu có giá trị trung bình gần bằng nhau. Cổ phiếu nào có phương sai, độ lệch chuẩn cao hơn thì được coi là có độ rủi ro lớn hơn.

Theo quan điểm trên, hãy so sánh độ rủi ro của cổ phiếu A và cổ phiếu B.

Lời giải

Ta có bảng thống kê giá đóng cửa theo giá trị đại diện:



Giá đóng cửa	121	123	125	127	129
Cổ phiếu A	8	9	12	10	11
Cổ phiếu B	16	4	3	6	21

Xét mẫu số liệu của cổ phiếu A :

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_1 = \frac{8.121 + 9.123 + 12.125 + 10.127 + 11.129}{50} = 125,28.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_1^2 = \frac{1}{50} (8.121^2 + 9.123^2 + 12.125^2 + 10.127^2 + 11.129^2) - (125,28)^2 = 7,5216$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_1 = \sqrt{S_1^2} = \sqrt{7,5216}$.

Xét mẫu số liệu của cổ phiếu B :

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_2 = \frac{16.121 + 4.123 + 3.125 + 6.127 + 21.129}{50} = 125,28.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_2^2 = \frac{1}{50} (16.121^2 + 4.123^2 + 3.125^2 + 6.127^2 + 21.129^2) - (125,28)^2 = 12,4096.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_2 = \sqrt{S_2^2} = \sqrt{12,4096}$.

Vậy nếu đánh giá độ rủi ro theo phương sai và độ lệch chuẩn thì cổ phiếu A có độ rủi ro thấp hơn cổ phiếu B.

Bài tập 7: Mai và Ngọc cùng sử dụng vòng đeo tay thông minh để ghi lại số bước chân hai bạn đi mỗi ngày trong một tháng. Kết quả được ghi lại ở bảng sau:

Giá trị đại diện	4	6	8	10	12
Mai	6	7	6	6	5
Ngọc	2	5	10	8	2

a) Hãy tính số trung bình và độ lệch chuẩn của mẫu số liệu ghép nhóm trên.

b) Nếu so sánh theo độ lệch chuẩn thì bạn nào có số lượng bước chân đi mỗi ngày đều đặn hơn?

Lời giải

a) Cỡ mẫu: $n = 30$

Xét mẫu số liệu của Mai:

$$\text{Số trung bình: } \bar{x}_1 = \frac{6.4 + 7.6 + 6.8 + 6.10 + 5.12}{30} = 7,8$$

$$\text{Phương sai: } S_1^2 = \frac{(6.4^2 + 7.6^2 + 6.8^2 + 6.10^2 + 5.12^2)}{30} - 7,8^2 = 7,56$$

$$\text{Độ lệch chuẩn: } \sigma_1 = \sqrt{7,56} \approx 2,75$$

Xét mẫu số liệu của Ngọc:



Số trung bình: $\overline{x}_2 = \frac{2.4 + 5.6 + 13.8 + 8.10 + 2.12}{30} = 8,2$

Phương sai: $S_2^2 = \frac{(2.4^2 + 5.6^2 + 13.8^2 + 8.10^2 + 2.12^2)}{30} - 8,2^2 \approx 3,83$

Độ lệch chuẩn: $\sigma_2 = \sqrt{3,83} \approx 1,96$

b) Nếu so sánh theo độ lệch chuẩn thì bạn Ngọc có số lượng bước chân đi mỗi ngày đều đặn hơn.

Bài tập 8: Bảng dưới đây thống kê cự li ném tạ của một vận động viên.

Cự li (m)	[19;19,5)	[19,5;20)	[20;20,5)	[20,5;21)	[21;21,5)
Tần số	13	45	24	12	6

Hãy tính phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm trên.

Lời giải

Giá trị đại diện	19,25	19,75	20,25	20,75	21,25
Tần số	13	45	24	12	6

Cỡ mẫu: $n = 100$

Số trung bình: $\overline{x} = \frac{13.19,25 + 45.19,75 + 24.20,25 + 12.20,75 + 6.21,25}{100} = 20,015$

Phương sai: $S^2 = \frac{13.19,25^2 + 45.19,75^2 + 24.20,25^2 + 12.20,75^2 + 6.21,25^2}{100} - 20,015^2 \approx 0,28$

Độ lệch chuẩn: $\sigma = \sqrt{0,28} \approx 0,53$



BÀI TẬP TRẮC NGHIỆM

PHẦN I. Câu trắc nghiệm nhiều phương án lựa chọn. Mỗi câu hỏi thí sinh chỉ chọn một phương án.

Câu 1: Số đặc trưng nào không sử dụng thông tin của nhóm số liệu đầu tiên và nhóm số liệu cuối cùng?

- A.** Khoảng biến thiên.
C. Phương sai.
- B.** Khoảng tứ phân vị.
D. Độ lệch chuẩn.

Lời giải

Số đặc trưng không sử dụng thông tin của nhóm số liệu đầu tiên và nhóm số liệu cuối cùng là khoảng tứ phân vị.

Câu 2: Nếu thay tất cả các tần số trong mẫu số liệu ghép nhóm trên bằng 4 thì số đặc trưng nào sau đây không thay đổi?

- A.** Khoảng biến thiên.
B. Khoảng tứ phân vị.
C. Phương sai.
D. Độ lệch chuẩn.

Lời giải

Nếu thay tất cả các tần số trong mẫu số liệu ghép nhóm trên bằng 4 thì số đặc trưng không đổi là khoảng biến thiên.

Câu 3: Trong các khẳng định sau, khẳng định nào sai?

- A.** Phương sai luôn luôn là số không âm.
B. Phương sai là bình phương của độ lệch chuẩn.
C. Phương sai càng lớn thì độ phân tán của các giá trị quanh số trung bình càng lớn.
D. Phương sai luôn luôn lớn hơn độ lệch chuẩn.

Lời giải

Khi $x \in (0;1)$ thì $s^2 < s$ nên khẳng định phương sai luôn lớn hơn độ lệch chuẩn là sai.

Câu 4: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một cửa hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

Doanh thu	[5; 7)	[7; 9)	[9; 11)	[11; 13)	[13; 15)
Số ngày	2	7	7	3	1

Số trung bình của mẫu số liệu trên thuộc khoảng nào trong các khoảng dưới đây?

- A.** [7;9). **B.** [9;11). **C.** [11;13). **D.** [13;15).

Lời giải

Bảng tần số ghép nhóm theo giá trị đại diện là:

Doanh thu	[5; 7)	[7; 9)	[9; 11)	[11; 13)	[13; 15)
Giá trị đại diện	6	8	10	12	14
Số ngày	2	7	7	3	1

Số trung bình: $\bar{x} = \frac{2.6 + 7.8 + 7.10 + 3.12 + 1.14}{20} = 9,4.$



Câu 5: Một siêu thị thống kê số tiền (đơn vị: chục nghìn đồng) mà 44 khách hàng mua hàng ở siêu thị đó trong một ngày. Số liệu được ghi lại trong bảng dưới đây.

Nhóm	Giá trị đại diện	Tần số
[40; 45)	42,5	4
[45; 50)	47,5	14
[50; 55)	52,5	8
[55; 60)	57,5	10
[60; 75)	62,5	6
[65; 70)	67,5	2
		$n = 44$

Phương sai của mẫu số liệu ghép nhóm trên là:

- A. 53,2. **B. 46,1.** C. 30. D. 11.
- Lời giải**

Số trung bình cộng của mẫu số liệu ghép nhóm là:

$$\bar{x} = \frac{4.42,5 + 14.47,5 + 8.52,5 + 10.57,5 + 6.62,5 + 2.67,5}{44} = \frac{585}{11}$$

Phương sai của mẫu số liệu ghép nhóm là:

$$\begin{aligned} s^2 &= \frac{4\left(42,5 - \frac{585}{11}\right)^2 + 14\left(47,5 - \frac{585}{11}\right)^2 + 8\left(52,5 - \frac{585}{11}\right)^2 + 10\left(57,5 - \frac{585}{11}\right)^2}{44} \\ &\quad + \frac{6\left(62,5 - \frac{585}{11}\right)^2 + 2\left(67,5 - \frac{585}{11}\right)^2}{44} \approx 46,12 \end{aligned}$$

Câu 6: Một siêu thị thống kê số tiền (đơn vị: chục nghìn đồng) mà 44 khách hàng mua hàng ở siêu thị đó trong một ngày. Số liệu được ghi lại trong bảng dưới đây.

Nhóm	Giá trị đại diện	Tần số
[40; 45)	42,5	4
[45; 50)	47,5	14
[50; 55)	52,5	8
[55; 60)	57,5	10
[60; 75)	62,5	6
[65; 70)	67,5	2
		$n = 44$

Độ lệch chuẩn của mẫu số liệu ghép nhóm trên (làm tròn kết quả đến hàng đơn vị) là:



A. 6,8.

B. 7,3.

C. 3,3.

D. 46,1.

Lời giải

Số trung bình cộng của mẫu số liệu ghép nhóm là:

$$\bar{x} = \frac{4.42,5 + 14.47,5 + 8.52,5 + 10.57,5 + 6.62,5 + 2.67,5}{44} = \frac{585}{11}$$

Phương sai của mẫu số liệu ghép nhóm là:

$$s^2 = \frac{4\left(42,5 - \frac{585}{11}\right)^2 + 14\left(47,5 - \frac{585}{11}\right)^2 + 8\left(52,5 - \frac{585}{11}\right)^2 + 10\left(57,5 - \frac{585}{11}\right)^2 + 6\left(62,5 - \frac{585}{11}\right)^2 + 2\left(67,5 - \frac{585}{11}\right)^2}{44} \approx 46,12$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s = \sqrt{s^2} = \sqrt{46,12} \approx 6,8$

Câu 7: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khỏe. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

Quãng đường (km)	[2, 7; 3, 0)	[3, 0; 3, 3)	[3, 3; 3, 6)	[3, 6; 3, 9)	[3, 9; 4, 2)
Số ngày	3	6	5	4	2

Phương sai của mẫu số liệu ghép nhóm là

A. 3,39.

B. 11,62.

C. 0,1314.

D. 0,36.

Lời giải

Giá trị đại diện	2,85	3,15	3,45	3,75	4,05
Số ngày	3	6	5	4	2

$$\text{Số trung bình: } \bar{x} = \frac{3.2,85 + 6.3,15 + 5.3,45 + 4.3,75 + 2.4,05}{20} = 3,39$$

$$\text{Phương sai: } S^2 = \frac{3.2,85^2 + 6.3,15^2 + 5.3,45^2 + 4.3,75^2 + 2.4,05^2}{20} - 3,39^2 = 0,1314$$

Câu 8: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khỏe. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

Quãng đường (km)	[2, 7; 3, 0)	[3, 0; 3, 3)	[3, 3; 3, 6)	[3, 6; 3, 9)	[3, 9; 4, 2)
Số ngày	3	6	5	4	2

Độ lệch chuẩn của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

A. 3,41.

B. 11,62.

C. 0,017.

D. 0,36.

Lời giải



Giá trị đại diện	2,85	3,15	3,45	3,75	4,05
Số ngày	3	6	5	4	2

$$\text{Số trung bình: } \bar{x} = \frac{3.2,85 + 6.3,15 + 5.3,45 + 4.3,75 + 2.4,05}{20} = 3,39$$

$$\text{Phương sai: } S^2 = \frac{3.2,85^2 + 6.3,15^2 + 5.3,45^2 + 4.3,75^2 + 2.4,05^2}{20} - 3,39^2 = 0,1314$$

$$\text{Độ lệch chuẩn: } \sigma = \sqrt{0,1314} \approx 0,36$$

Câu 9: Bạn Chi rất thích nhảy hiện đại. Thời gian tập nhảy mỗi ngày trong thời gian gần đây của bạn Chi được thống kê lại ở bảng sau:

Thời gian (phút)	[20;25)	[25;30)	[30;35)	[35;40)	[40;45)
Số ngày	6	6	4	1	1

Phương sai của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

- A. 31,77. B. 32. C. 31. **D. 31,44.**

Lời giải

$$\text{Số trung bình: } \bar{x} = \frac{6.22,5 + 6.27,5 + 4.32,5 + 37,5 + 42,5}{18} \approx 28,33$$

$$\text{Phương sai: } S^2 = \frac{6.22,5^2 + 6.27,5^2 + 4.32,5^2 + 37,5^2 + 42,5^2}{18} - 28,33^2 = 31,25$$

Câu 10: Dũng là học sinh rất giỏi chơi rubik, bạn có thể giải nhiều loại khối rubik khác nhau. Trong một lần tập luyện giải khối rubik 3×3 , bạn Dũng đã tự thống kê lại thời gian giải rubik trong 25 lần giải liên tiếp ở bảng sau:

Thời gian giải rubik (giây)	[8;10)	[10;12)	[12;14)	[14;16)	[16;18)
Số lần	4	6	8	4	3

Độ lệch chuẩn của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

- A. 5,98.** B. 6. C. 2,44. D. 2,5.

Lời giải

Giá trị đại diện	9	11	13	15	17
Số lần	4	6	8	4	3

$$\text{Số trung bình: } \bar{x} = \frac{4.9 + 6.11 + 8.13 + 4.15 + 3.17}{25} = 12,68$$

$$\text{Độ lệch chuẩn: } \sigma = \sqrt{\frac{4.9^2 + 6.11^2 + 8.13^2 + 4.15^2 + 3.17^2}{25} - 12,68^2} \approx 5,98.$$

Câu 11: Để đánh giá chất lượng của một loại pin điện thoại mới, người ta ghi lại thời gian nghe nhạc liên tục của điện thoại được sạc đầy pin cho đến khi hết pin cho kết quả sau:



Thời gian (giờ)	[5;5,5)	[5,5;6)	[6;6,5)	[6,5;7)	[7;7,5)
Số chiếc điện thoại (tần số)	2	8	15	10	5

Tính độ lệch chuẩn của mẫu số liệu ghép nhóm trên (làm tròn đến 4 chữ số thập phân)

- A.** 0,4252. **B.** 0,5314. **C.** 0,6214. **D.** 0,5268.

Lời giải

Thời gian (giờ)	[5;5,5)	[5,5;6)	[6;6,5)	[6,5;7)	[7;7,5)
Giá trị đại diện	5,25	5,75	6,25	6,75	7,25
Số chiếc điện thoại (tần số)	2	8	15	10	5

Số trung bình của mẫu số liệu:

$$\bar{x} = \frac{m_1 \cdot x_1 + \dots + m_k \cdot x_k}{n} = \frac{2 \cdot 5,25 + 8 \cdot 5,75 + 15 \cdot 6,25 + 10 \cdot 6,75 + 5 \cdot 7,25}{40} = 6,35$$

Phương sai của mẫu số liệu ghép nhóm:

$$s^2 = \frac{1}{40} (2 \cdot 5,25^2 + 8 \cdot 5,75^2 + 15 \cdot 6,25^2 + 10 \cdot 6,75^2 + 5 \cdot 7,25^2) - 6,35^2 = 0,2775$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s = \sqrt{s^2} = \sqrt{0,2775} \approx 0,5268$

Câu 12: Người ta ghi lại tiền lãi (đơn vị: triệu đồng) của một số nhà đầu tư (với số tiền đầu tư như nhau), khi đầu tư vào hai lĩnh vực A, B cho kết quả như sau:

Tiền lãi	[5;10)	[10;15)	[15;20)	[20;25)	[25;30)
Số nhà đầu tư vào lĩnh vực A	2	5	8	6	4
Số nhà đầu tư vào lĩnh vực B	8	4	2	5	6

Tính độ lệch chuẩn cho các mẫu số liệu về tiền lãi của các nhà đầu tư ở lĩnh vực B

- A.** 8,039. **B.** 8,5314. **C.** 8,6214. **D.** 8,5268.

Lời giải

Ta có mẫu số liệu ghép nhóm với giá trị đại diện là:

Tiền lãi	[5;10)	[10;15)	[15;20)	[20;25)	[25;30)
Giá trị đại diện	7,5	12,5	17,5	22,5	27,5
Số nhà đầu tư vào lĩnh vực A	2	5	8	6	4
Số nhà đầu tư vào lĩnh vực B	8	4	2	5	6

Phương sai của mẫu số liệu về tiền lãi khi đầu tư vào lĩnh vực B:

$$s_B^2 = \frac{1}{25} (7,5^2 \cdot 8 + 12,5^2 \cdot 4 + 17,5^2 \cdot 2 + 22,5^2 \cdot 5 + 27,5^2 \cdot 6) - 16,9^2 = 64,64$$

Độ lệch chuẩn của mẫu số liệu về tiền lãi khi đầu tư vào lĩnh vực B: $s_B = \sqrt{64,64} \approx 8,039$ (triệu đồng)

**PHẦN II. Câu trắc nghiệm đúng sai. Trong mỗi ý a), b), c), d) ở mỗi câu, thí sinh chọn đúng hoặc sai.**

Câu 1: Kết quả 40 lần nhảy xa của hai vận động viên Dũng và Huy được lần lượt thống kê trong bảng bên dưới (đơn vị: mét)

Nhóm	Dũng	Huy
[6, 22; 6, 46)	3	2
[6, 46; 6, 70)	7	5
[6, 70; 6, 94)	5	8
[6, 94; 7, 18)	20	19
[7, 18; 7, 42)	5	6
	$n = 40$	$n = 40$

Xét tính đúng sai của các khẳng định sau:

- a) Số trung bình cộng của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Dũng cho bởi Bảng 15 (làm tròn kết quả đến hàng phần trăm) là 6,92 m.
- b) Độ lệch chuẩn của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Dũng cho bởi Bảng 15 (làm tròn kết quả đến hàng phần trăm) là 0,26 m.
- c) Phương sai của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Huy cho bởi Bảng 16 (làm tròn kết quả đến hàng phần trăm) là 0,16.
- d) Kết quả nhảy xa của vận động viên Dũng đồng đều hơn kết quả nhảy xa của vận động viên Huy.

Lời giải

a) Đúng: Số trung bình cộng của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Dũng là:

$$\bar{x}_D = \frac{3.6,34 + 7.6,58 + 5.6,82 + 20.7,06 + 5.7,30}{40} = \frac{276,88}{40} \approx 6,92 \text{ (m)}$$

b) Đúng: Phương sai của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Dũng (làm tròn kết quả đến hàng phần trăm) là:

$$s_D^2 = \frac{1}{40} \left[3.(6,34 - 6,92)^2 + 7.(6,58 - 6,92)^2 + 5.(6,82 - 6,92)^2 + 20.(7,06 - 6,92)^2 + 5.(7,30 - 6,92)^2 \right] = \frac{2,9824}{40} \approx 0,07.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm trên là: $s_D \approx \sqrt{0,07} \approx 0,26 \text{ (m)}$

c) Sai: Số trung bình cộng của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Huy là:

$$\bar{x}_H = \frac{2.6,34 + 5.6,58 + 8.6,82 + 19.7,06 + 6.7,30}{40} = \frac{278,08}{40} \approx 6,95 \text{ (m)}$$

Vậy phương sai của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Huy (làm tròn kết quả đến hàng phần trăm) là:



$$s_H^2 = \frac{1}{40} \left[2 \cdot (6,34 - 6,95)^2 + 5 \cdot (6,58 - 6,95)^2 + 8 \cdot (6,82 - 6,95)^2 + 19 \cdot (7,06 - 6,95)^2 + 6 \cdot (7,30 - 6,95)^2 \right] = \frac{2,5288}{40} \approx 0,06$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm trên là: $s_H \approx \sqrt{0,06} \approx 0,24$ (m)

d) Sai: Do $s_H \approx 0,24 < s_D \approx 0,26$ nên kết quả nhảy xa của vận động viên Huy đồng đều hơn kết quả nhảy xa của vận động viên Dũng.

Câu 2: Mẫu số liệu ghép nhóm thống kê mức lương của hai công ty A, B (đơn vị: triệu đồng) được thể hiện như bảng dưới đây

Nhóm	Công ty A	Công ty B
[10;15)	15	25
[15;20)	18	15
[20;25)	10	7
[25;30)	10	5
[30;35)	5	5
[35;40)	2	3
	$n = 60$	$n = 60$

Xét tính đúng sai của các khẳng định sau:

- a) Số trung bình cộng của mẫu số liệu ghép nhóm của công ty A là: $\frac{62}{3}$
- b) Độ lệch chuẩn của mẫu số liệu ghép nhóm của công ty A là: 5
- c) Phương sai của mẫu số liệu ghép nhóm của công ty B (làm tròn kết quả đến hàng phần trăm) là 52,91.
- b) Nhận thấy độ lệch chuẩn của công ty A nhỏ hơn công ty B nên mức lương của công ty A đồng đều hơn.

Lời giải

- a) Đúng: Số trung bình cộng của mẫu số liệu ghép nhóm của công ty A là:

$$\bar{x}_A = \frac{15 \cdot 12,5 + 18 \cdot 17,5 + 10 \cdot 22,5 + 10 \cdot 27,5 + 5 \cdot 32,5 + 2 \cdot 37,5}{60} = \frac{62}{3}$$

- b) Sai: Phương sai của mẫu số liệu ghép nhóm của công ty A là:

$$s_A^2 = \frac{15 \cdot \left(12,5 - \frac{62}{3}\right)^2 + 18 \cdot \left(17,5 - \frac{62}{3}\right)^2 + 10 \cdot \left(22,5 - \frac{62}{3}\right)^2 + 10 \cdot \left(27,5 - \frac{62}{3}\right)^2 + 5 \cdot \left(32,5 - \frac{62}{3}\right)^2 + 2 \cdot \left(37,5 - \frac{62}{3}\right)^2}{60} \approx 49,14$$



Độ lệch chuẩn của mẫu số liệu ghép nhóm của công ty A là: $s_A = \sqrt{s_A^2} = \sqrt{49,13} \approx 7$

c) Sai: Số trung bình cộng của mẫu số liệu ghép nhóm của công ty B là:

$$\bar{x}_B = \frac{25.12,5 + 15.17,5 + 7.22,5 + 5.27,5 + 5.32,5 + 3.37,5}{60} = \frac{229}{12}$$

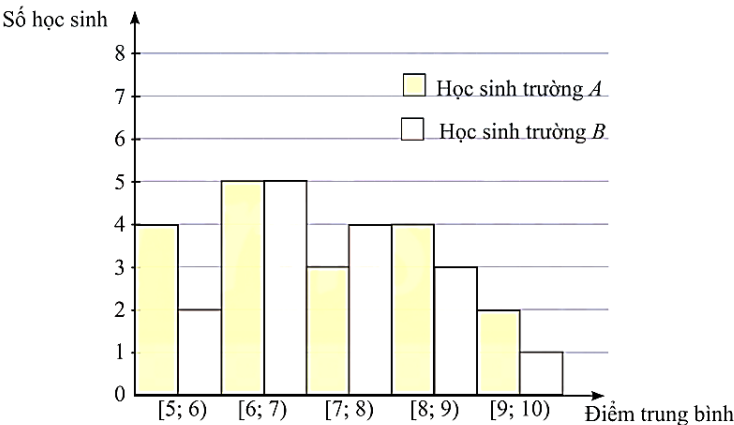
Phương sai của mẫu số liệu ghép nhóm của công ty B là:

$$s_B^2 = \frac{25.\left(12,5 - \frac{229}{12}\right)^2 + 15.\left(17,5 - \frac{229}{12}\right)^2 + 7.\left(22,5 - \frac{229}{12}\right)^2 + 5.\left(27,5 - \frac{229}{12}\right)^2 + 5.\left(32,5 - \frac{229}{12}\right)^2 + 3.\left(37,5 - \frac{229}{12}\right)^2}{60} \approx 57,91$$

d) Đúng: Độ lệch chuẩn của mẫu số liệu ghép nhóm của công ty B là: $s_B = \sqrt{s_B^2} = \sqrt{57,91} \approx 7,61$

Nhận thấy độ lệch chuẩn của công ty A nhỏ hơn công ty B nên mức lương của công ty A đồng đều hơn.

Câu 3: Biểu đồ sau mô tả kết quả điều tra về điểm trung bình năm học của học sinh hai trường A và B



Người ta lập được bảng tần số ghép nhóm cho mẫu số liệu trên như sau:

Điểm trung bình	[5; 6)	[6; 7)	[7; 8)	[8; 9)	[9; 10)
Giá trị đại diện	5,5	6,5	7,5	8,5	9,5
Học sinh trường A	4	5	3	4	2
Học sinh trường B	2	5	4	3	1

Xét tính đúng sai của các khẳng định sau:

- a) Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm của học sinh trường A là: 6,1
- b) Khoảng tứ phân vị của mẫu số liệu ghép nhóm của học sinh trường B là: 1,73
- c) Nếu so sánh theo khoảng tứ phân vị của mẫu số liệu ghép nhóm thì học sinh trường B có điểm trung bình đồng đều hơn.



d) Nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì học sinh trường A có điểm trung bình đồng đều hơn.

Lời giải

a) Đúng: Cỡ mẫu: $n_A = 18$

Gọi $x_1; x_2; \dots; x_{18}$ là mẫu số liệu gốc về điểm trung bình năm học của học sinh hai trường A được xếp theo thứ tự không giảm.

Ta có: $x_1; \dots; x_4 \in [5; 6); x_5; \dots; x_9 \in [6; 7); x_{10}; \dots; x_{12} \in [7; 8); x_{12}; \dots; x_{16} \in [8; 9); x_{17}; x_{18} \in [9; 10)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $x_5 \in [6; 7)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là: $Q_1 = 6 + \frac{\frac{18}{4} - 4}{5}(7 - 6) = 6,1$

Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{14} \in [8; 9)$.

Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là: $Q_3 = 8 + \frac{\frac{3 \cdot 18}{4} - (4 + 5 + 3)}{4}(9 - 8) = 8,375$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 2,275$

b) Đúng: Cỡ mẫu: $n_B = 15$

Gọi $y_1; y_2; \dots; y_{15}$ là mẫu số liệu gốc về điểm trung bình năm học của học sinh hai trường B được xếp theo thứ tự không giảm.

Ta có: $y_1; y_2 \in [5; 6); y_3; \dots; y_7 \in [6; 7); y_8; \dots; y_{11} \in [7; 8); y_{12}; \dots; y_{14} \in [8; 9); y_{15} \in [9; 10)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $y_4 \in [6; 7)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là: $Q'_1 = 6 + \frac{\frac{15}{4} - 2}{5}(7 - 6) = 6,35$

Tứ phân vị thứ ba của mẫu số liệu gốc là $y_{12} \in [8; 9)$.

Do đó tứ phân vị thứ ba của mẫu số liệu ghép nhóm là: $Q'_3 = 8 + \frac{\frac{3 \cdot 15}{4} - (2 + 5 + 4)}{3}(9 - 8) = 8,08$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta'_Q = Q'_3 - Q'_1 = 1,73$

c) Đúng: Vậy nếu so sánh theo khoảng tứ phân vị của mẫu số liệu ghép nhóm thì học sinh trường B có điểm trung bình đồng đều hơn.

d) Sai: Xét số liệu của trường A :



$$\text{Số trung bình: } \bar{x}_A = \frac{4.5,5 + 5.6,5 + 3.7,5 + 4.8,5 + 2.9,5}{18} = 7,22$$

$$\text{Độ lệch chuẩn: } \sigma_A = \sqrt{\frac{4.5,5^2 + 5.6,5^2 + 3.7,5^2 + 4.8,5^2 + 2.9,5^2}{18} - 7,22^2} \approx 1,79$$

Xét số liệu của trường B :

$$\text{Số trung bình: } \bar{x}_B = \frac{2.5,5 + 5.6,5 + 4.7,5 + 3.8,5 + 1.9,5}{15} = 7,23$$

$$\text{Độ lệch chuẩn: } \sigma_B = \sqrt{\frac{2.5,5^2 + 5.6,5^2 + 4.7,5^2 + 3.8,5^2 + 1.9,5^2}{15} - 7,23^2} \approx 1,31$$

Vậy nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì học sinh trường B có điểm trung bình đồng đều hơn.

Câu 4: Bảng dưới đây biểu diễn mẫu số liệu ghép nhóm về nhiệt độ không khí trung bình các tháng năm 2021 tại Hà Nội và Huế (đơn vị: độ C).

Nhóm	Hà Nội	Huế
[16,8;19,8)	2	1
[19,8;22,8)	3	2
[22,8;25,8)	2	3
[25,8;28,8)	1	2
[28,8;31,8)	4	4
	$n = 12$	$n = 12$

Xét tính đúng sai của các khẳng định sau:

- Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên số liệu ở Hà Nội là: 8,75
- Độ lệch chuẩn của mẫu số liệu ghép nhóm trên số liệu ở Hà Nội (làm tròn kết quả đến hàng phần trăm) là: 3,56
- Phương sai của mẫu số liệu ghép nhóm trên số liệu ở Huế là: 12,75
- Huế có nhiệt độ không khí trung bình tháng đồng đều hơn vì độ lệch chuẩn nhỏ hơn

Lời giải

a) Đúng: Xét số liệu ở Hà Nội:

$$\text{Khoảng biến thiên: } R = 31,8 - 16,8 = 15$$

Số phân tử của mẫu là $n = 12$

Tần số tích lũy của các nhóm lần lượt là $cf_1 = 2, cf_2 = 5, cf_3 = 7, cf_4 = 8, cf_5 = 12$



Ta có: $\frac{n}{4} = \frac{12}{4} = 3$ mà $2 < 3 < 5$ suy ra nhóm 2 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 3.

Xét nhóm 2 là nhóm $[19,8;22,8)$ có $s = 19,8, h = 3, n_2 = 3$ và nhóm 1 là nhóm $[16,8;19,8)$ có $cf_1 = 2$. Ta có tứ phân vị thứ nhất là: $Q_1 = s + \left(\frac{3 - cf_1}{n_2} \right) \cdot h = 19,8 + \left(\frac{3 - 2}{3} \right) \cdot 3 = 20,8$

Ta có: $\frac{3n}{4} = \frac{3 \cdot 12}{4} = 9$ mà $8 < 9 < 12$ suy ra nhóm 5 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 9.

Xét nhóm 5 là nhóm $[28,8;31,8)$ có $t = 28,8, l = 3, n_5 = 4$ và nhóm 4 là nhóm $[25,8;28,8)$ có $cf_4 = 8$

Ta có tứ phân vị thứ ba là: $Q_3 = t + \left(\frac{9 - cf_4}{n_5} \right) \cdot l = 28,8 + \left(\frac{9 - 8}{4} \right) \cdot 3 = 29,55$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $Q_3 - Q_1 = 29,55 - 20,8 = 8,75$

b) Sai: Số trung bình cộng của mẫu số liệu ghép nhóm là:

$$\bar{x}_1 = \frac{2 \cdot 18,3 + 3 \cdot 21,3 + 2 \cdot 24,3 + 27,3 + 4 \cdot 30,3}{12} = 24,8$$

Phương sai của mẫu số liệu ghép nhóm là:

$$s_1^2 = \frac{2(18,3 - 24,8)^2 + 3(21,3 - 24,8)^2 + 2(24,3 - 24,8)^2 + (27,3 - 24,8)^2 + 4(30,3 - 24,8)^2}{12} = 20,75$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s_1 = \sqrt{s_1^2} = \sqrt{20,75} \approx 4,56$

c) Sai: Xét số liệu ở Huế:

Khoảng biến thiên: $R = 31,8 - 16,8 = 15$

Số phần tử của mẫu là $n = 12$

Tần số tích lũy của các nhóm lần lượt là $cf_1 = 1, cf_2 = 3, cf_3 = 6, cf_4 = 8, cf_5 = 12$

Ta có: $\frac{n}{4} = \frac{12}{4} = 3$ suy ra nhóm 2 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 3. Xét nhóm 2 là nhóm $[19,8;22,8)$ có $s = 19,8, h = 3, n_2 = 2$ và nhóm 1 là nhóm $[16,8;19,8)$ có $cf_1 = 1$

Ta có tứ phân vị thứ nhất là: $Q_1 = s + \left(\frac{3 - cf_1}{n_2} \right) \cdot h = 19,8 + \left(\frac{3 - 1}{2} \right) \cdot 3 = 22,8$

Ta có: $\frac{3n}{4} = \frac{3 \cdot 12}{4} = 9$ mà $8 < 9 < 12$ suy ra nhóm 5 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 9.



Xét nhóm 5 là nhóm $[28,8;31,8)$ có $t = 28,8, l = 3, n_5 = 4$ và nhóm 4 là nhóm $[25,8;28,8)$ có $cf_4 = 8$

Ta có tứ phân vị thứ ba là: $Q_3 = t + \left(\frac{9 - cf_4}{n_5} \right) \cdot l = 28,8 + \left(\frac{9 - 8}{4} \right) \cdot 3 = 29,55$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $Q_3 - Q_1 = 29,55 - 22,8 = 6,75$

Số trung bình cộng của mẫu số liệu ghép nhóm là:

$$\bar{x}_2 = \frac{18,3 + 2 \cdot 21,3 + 3 \cdot 24,3 + 2 \cdot 27,3 + 4 \cdot 30,3}{12} = 25,8$$

Phương sai của mẫu số liệu ghép nhóm là:

$$s_2^2 = \frac{(18,3 - 25,8)^2 + 3(21,3 - 25,8)^2 + 3(24,3 - 25,8)^2 + 2(27,3 - 25,8)^2 + 4(30,3 - 25,8)^2}{12} = 15,75$$

d) Đúng: Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s_2 = \sqrt{s_2^2} = \sqrt{15,75} \approx 3,97$

Huế có nhiệt độ không khí trung bình tháng đồng đều hơn vì độ lệch chuẩn nhỏ hơn

Câu 5: Bảng dưới đây thống kê độ ẩm không khí trung bình các tháng năm 2021 tại Đà Lạt và Vũng Tàu (đơn vị: %).

Tháng	1	2	3	4	5	6	7	8	9	10	11	12
Đà Lạt	83	79	79	87	87	87	88	89	90	91	88	86
Vũng Tàu	75	77	78	77	79	79	81	79	81	83	80	77

Người ta lần lượt ghép các số liệu của Đà Lạt, Vũng Tàu thành mẫu số liệu ghép nhóm như sau:

Đà Lạt			Vũng Tàu		
Nhóm	Giá trị đại diện	Tần số	Nhóm	Giá trị đại diện	Tần số
$[75; 78,3)$	76,65	0	$[75; 78,3)$	76,65	5
$[78,3; 81,6)$	79,95	2	$[78,3; 81,6)$	79,95	6
$[81,6; 84,9)$	83,25	1	$[81,6; 84,9)$	83,25	1
$[84,9; 88,2)$	86,55	6	$[84,9; 88,2)$	86,55	0
$[88,2; 91,5)$	89,85	3	$[88,2; 91,5)$	89,85	0

Xét tính đúng sai của các khẳng định sau:

- Xét số liệu ở Đà Lạt ta có khoảng biến thiên là 16,5.
- Xét số liệu ở Đà Lạt thì ta có độ lệch chuẩn của mẫu số liệu ghép nhóm (làm tròn kết quả đến hàng phần trăm) là 3,28.
- Xét số liệu ở Vũng Tàu thì phương sai của mẫu số liệu ghép nhóm là 4,235.



d) Đà Lạt có nhiệt độ không khí trung bình tháng đồng đều hơn vì độ lệch chuẩn nhỏ hơn

Lời giải

a) Đúng: Xét số liệu ở Đà Lạt:

Khoảng biến thiên: $R = 91,5 - 75 = 16,5$

b) Đúng: Số phân tử của mẫu là $n = 12$

Tần số tích lũy của các nhóm lần lượt là $cf_1 = 0, cf_2 = 2, cf_3 = 3, cf_4 = 9, cf_5 = 12$

Ta có: $\frac{n}{4} = \frac{12}{4} = 3$ suy ra nhóm 3 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 3.

Xét nhóm 3 là nhóm $[81,6;84,9)$ có $s = 81,6, h = 3,3, n_3 = 1$ và nhóm 2 là nhóm $[78,3;81,6)$ có

$cf_2 = 2$. Ta có tứ phân vị thứ nhất là: $Q_1 = s + \left(\frac{3 - cf_2}{n_3} \right) \cdot h = 81,6 + \left(\frac{3 - 2}{1} \right) \cdot 3,3 = 84,9$

Ta có: $\frac{3n}{4} = \frac{3 \cdot 12}{4} = 9$ suy ra nhóm 4 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 9.

Xét nhóm 4 là nhóm $[84,9;88,2)$ có $t = 84,9, l = 3,3, n_4 = 6$ và nhóm 3 là nhóm $[81,6;84,9)$ có $cf_3 = 3$

Ta có tứ phân vị thứ ba là: $Q_3 = t + \left(\frac{9 - cf_3}{n_4} \right) \cdot l = 84,9 + \left(\frac{9 - 3}{6} \right) \cdot 3,3 = 88,2$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $Q_3 - Q_1 = 88,2 - 84,9 = 3,3$

Số trung bình cộng của mẫu số liệu ghép nhóm là:

$$\bar{x}_1 = \frac{0.76,65 + 2.79,95 + 83,25 + 6.86,55 + 3.89,85}{12} = 86$$

Phương sai của mẫu số liệu ghép nhóm là:

$$s_1^2 = \frac{0(76,65 - 86)^2 + 2(79,95 - 86)^2 + (83,25 - 86)^2 + 6(86,55 - 86)^2 + 3(89,85 - 86)^2}{12} = 10,7825$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s_1 = \sqrt{s_1^2} = \sqrt{10,7825} \approx 3,28$

c) Xét số liệu ở Vũng Tàu:

Khoảng biến thiên: $R = 91,5 - 75 = 16,5$

Số phân tử của mẫu là $n = 12$

Tần số tích lũy của các nhóm lần lượt là $cf_1 = 5, cf_2 = 11, cf_3 = 12, cf_4 = 12, cf_5 = 12$

Ta có: $\frac{n}{4} = \frac{12}{4} = 3$ suy ra nhóm 1 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 3.



Xét nhóm 1 là nhóm $[75; 78, 3)$ có $s = 75, h = 3, 3, n_1 = 5$

Ta có tứ phân vị thứ nhất là: $Q_1 = s + \left(\frac{3 - cf_0}{n_1} \right) \cdot h = 75 + \left(\frac{3 - 0}{5} \right) \cdot 3, 3 = 76, 98$

Ta có: $\frac{3n}{4} = \frac{3 \cdot 12}{4} = 9$ suy ra nhóm 2 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 9.

Xét nhóm 2 là nhóm $[78, 3; 81, 6)$ có $t = 78, 3, l = 3, 3, n_2 = 6$ và nhóm 1 là nhóm $[75; 78, 3)$ có $cf_1 = 5$. Ta có tứ phân vị thứ ba là: $Q_3 = t + \left(\frac{9 - cf_1}{n_2} \right) \cdot l = 75 + \left(\frac{9 - 5}{6} \right) \cdot 3, 3 = 77, 2$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $Q_3 - Q_1 = 77, 2 - 76, 98 = 0, 22$

Số trung bình cộng của mẫu số liệu ghép nhóm là: $\bar{x}_2 = \frac{5 \cdot 76, 65 + 6 \cdot 79, 95 + 83, 25}{12} = 78, 85$

Phương sai của mẫu số liệu ghép nhóm là:

$$s_2^2 = \frac{5(76, 65 - 78, 85)^2 + 6(79, 95 - 78, 85)^2 + (83, 25 - 78, 85)^2}{12} = 4, 235$$

d) Sai: Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s_2 = \sqrt{s_2^2} = \sqrt{4, 235} \approx 2, 06$

Vũng Tàu có nhiệt độ không khí trung bình tháng đồng đều hơn vì độ lệch chuẩn nhỏ hơn.

Câu 6: Bảng sau thống kê lại tổng số giờ nắng trong tháng 6 của các năm từ 2002 đến 2021 tại hai trạm quan trắc đặt ở Nha Trang và Quy Nhơn.

Số giờ nắng	[130; 160)	[160; 190)	[190; 220)	[220; 250)	[250; 280)	[280; 310)
Số năm ở Nha Trang	1	1	1	8	7	2
Số năm ở Quy Nhơn	0	1	2	4	10	3

Xét tính đúng sai của các khẳng định sau:

- Xét số liệu ở Nha Trang thì khoảng tứ phân vị của mẫu số liệu ghép nhóm là: 32, 64
- Nếu so sánh theo khoảng tứ phân vị thì số giờ nắng trong tháng 6 của Quy Nhơn đồng đều hơn
- Xét số liệu của Quy Nhơn ta có độ lệch chuẩn của mẫu số liệu ghép nhóm (làm tròn kết quả đến hàng phần trăm) là: 30, 59
- Nếu so sánh theo độ lệch chuẩn thì số giờ nắng trong tháng 6 của Nha Trang đồng đều hơn

Lời giải

a) Sai: Cỡ mẫu: $n = 20$

Gọi $x_1; x_2; \dots; x_{20}$ là mẫu số liệu gốc về số giờ nắng trong tháng 6 trong 20 năm của Nha Trang được xếp theo thứ tự không giảm.



Ta có: $x_1 \in [130; 160); x_2 \in [160; 190); x_3 \in [190; 220); x_4; \dots; x_{11} \in [220; 250); x_{12}; \dots;$

$x_{18} \in [250; 280); x_{19}; x_{20} \in [280; 310)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_5 + x_6) \in [220; 250)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là:

$$Q_1 = 220 + \frac{\frac{20}{4} - (1+1+1)}{8}(250 - 220) = 227,5$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{15} + x_{16}) \in [250; 280)$.

Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là:

$$Q_3 = 250 + \frac{\frac{3 \cdot 20}{4} - (1+1+1+8)}{7}(280 - 250) = \frac{1870}{7}$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 39,64$

b) Đúng: Gọi $y_1; y_2; \dots; y_{50}$ là mẫu số liệu gốc về số giờ nắng trong tháng 6 trong 20 năm của Quy Nhơn được xếp theo thứ tự không giảm.

Ta có:

$y_1; \dots; y_2 \in [160; 190); y_3 \in [190; 220); y_4; \dots; y_7 \in [220; 250); y_8; \dots; y_{17} \in [250; 280);$
 $y_{18}; \dots; y_{20} \in [280; 310)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(y_5 + y_6) \in [220; 250)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là:

$$Q'_1 = 220 + \frac{\frac{20}{4} - (1+2)}{4}(250 - 220) = 235$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(y_{15} + y_{16}) \in [250; 280)$.

Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là:

$$Q'_3 = 250 + \frac{\frac{3 \cdot 20}{4} - (1+2+4)}{10}(280 - 250) = 274$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta'_Q = Q'_3 - Q'_1 = 39$



Vậy nếu so sánh theo khoảng tứ phân vị thì số giờ nắng trong tháng 6 của Quy Nhơn đồng đều hơn

Xét số liệu của Nha Trang:

$$\text{Số trung bình: } \bar{x}_x = \frac{1.145 + 1.175 + 1.205 + 8.235 + 7.265 + 2.295}{20} = 242,5$$

$$\text{Độ lệch chuẩn: } \sigma_x = \sqrt{\frac{1.145^2 + 1.175^2 + 1.205^2 + 8.235^2 + 7.265^2 + 2.295^2}{20} - 242,5^2} \approx 35,34$$

c) Đúng: Xét số liệu của Quy Nhơn:

$$\text{Số trung bình: } \bar{x}_y = \frac{1.175 + 2.205 + 4.235 + 10.265 + 3.295}{20} = 253$$

$$\text{Độ lệch chuẩn: } \sigma_y = \sqrt{\frac{1.175^2 + 2.205^2 + 4.235^2 + 10.265^2 + 3.295^2}{20} - 253^2} \approx 30,59$$

d) Sai: Vậy nếu so sánh theo độ lệch chuẩn thì số giờ nắng trong tháng 6 của Quy Nhơn đồng đều hơn.

Câu 7: Thống kê tổng số giờ nắng trong tháng 9 tại một trạm quan trắc đặt ở Cà Mau trong các năm từ 2002 đến 2021 được thống kê như sau:

111,6	134,9	130,3	134,2	140,9	109,3	152,4	156,3	116,1	96,7
105,2	80,8	80,8	110	109	139	145	161	126	114

Người ta lập được bảng tần số ghép nhóm như sau:

Số giờ nắng	[80;98)	[98;116)	[116;134)	[134;152)	[152;170)
Giá trị đại diện	89	107	125	143	161
Số năm	3	6	3	5	3

Xét tính đúng sai của các khẳng định sau:

a) Số trung bình của mẫu số liệu ghép nhóm là 124,1.

b) Phương sai của mẫu số liệu ghép nhóm là 566,19.

c) Độ lệch chuẩn của mẫu số liệu ghép nhóm (kết quả các phép tính làm tròn đến hàng phần nghìn) là 23,795.

d) Sai số tương đối của độ lệch chuẩn của mẫu số liệu ghép nhóm so với độ lệch chuẩn của mẫu số liệu gốc (kết quả các phép tính làm tròn đến hàng phần nghìn) là 4,805%.

Lời giải

a) Đúng: Xét theo mẫu số liệu gốc. Cỡ mẫu là $n = 20$.

$$\text{Số trung bình của mẫu số liệu trên là: } \bar{x}_1 = \frac{111,6 + 134,9 + \dots + 114}{20} = 122,755.$$



Phương sai của mẫu số liệu trên là: $S_1^2 = \frac{1}{20}(111,6^2 + 134,9^2 + \dots + 114^2) - 122,755^2 \approx 515,453$.

Độ lệch chuẩn của mẫu số liệu trên là: $S_1 \approx \sqrt{515,453} \approx 22,704$.

Xét theo mẫu số liệu ghép nhóm

Số trung bình của mẫu số liệu ghép nhóm là $\bar{x}_2 = \frac{3.89 + 6.107 + 3.125 + 5.143 + 3.161}{20} = 124,1$.

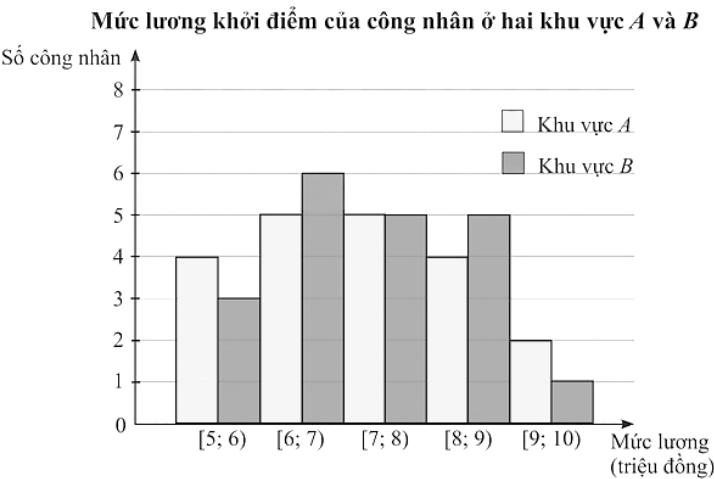
b) Đúng: Phương sai của mẫu số liệu ghép nhóm là

$$S_2^2 = \frac{1}{20}(3.89^2 + 6.107^2 + 3.125^2 + 5.143^2 + 3.161^2) - 124,1^2 = 566,19.$$

c) Đúng: Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S_2 = \sqrt{566,19} \approx 23,795.77$

d) Đúng: Sai số tương đối của độ lệch chuẩn của mẫu số liệu ghép nhóm so với độ lệch chuẩn của mẫu số liệu gốc là: $\frac{|S_2 - S_1|}{S_1} = \frac{|23,795 - 22,704|}{22,704} \cdot 100\% \approx 4,805\%$.

Câu 8: Biểu đồ dưới đây mô tả kết quả điều tra về mức lương khởi điểm (đơn vị: triệu đồng) của một số công nhân ở hai khu vực A và B.



Người ta lập được bảng tần số ghép nhóm cho mẫu số liệu như sau:

Mức lương	[5;6)	[6;7)	[7;8)	[8;9)	[9;10)
Mức lương đại diện (triệu đồng)	5,5	6,5	7,5	8,5	9,5
Khu vực A	4	5	5	4	2
Khu vực B	3	6	5	5	1

Xét tính đúng sai của các khẳng định sau:

a) Xét mẫu số liệu của khu vực A ta có số trung bình của mẫu số liệu ghép nhóm là 6,25.

b) Xét mẫu số liệu của khu vực A ta có độ lệch chuẩn của mẫu số liệu ghép nhóm là $\sqrt{1,5875}$

c) Xét mẫu số liệu của khu vực B ta có phương sai của mẫu số liệu ghép nhóm là 1,3875.



d) Nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì mức lương khởi điểm của công nhân khu vực B đồng đều hơn của công nhân khu vực A .

Lời giải

a) Sai: Xét mẫu số liệu của khu vực A . Cỡ mẫu là $n_A = 4 + 5 + 5 + 4 + 2 = 20$.

Số trung bình của mẫu số liệu ghép nhóm là $\bar{x}_A = \frac{4.5,5 + 5.6,5 + 5.7,5 + 4.8,5 + 2.9,5}{20} = 7,25$.

Phương sai của mẫu số liệu ghép nhóm là

$$S_A^2 = \frac{1}{20} (4.5,5^2 + 5.6,5^2 + 5.7,5^2 + 4.8,5^2 + 2.9,5^2) - (7,25)^2 = 1,5875$$

b) Đúng: Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S_A = \sqrt{1,5875}$

Xét mẫu số liệu của khu vực B . Cỡ mẫu là $n_B = 3 + 6 + 5 + 5 + 1 = 20$.

Số trung bình của mẫu số liệu ghép nhóm là: $\bar{x}_B = \frac{3.5,5 + 6.6,5 + 5.7,5 + 5.8,5 + 1.9,5}{20} = 7,25$.

c) Sai: Phương sai của mẫu số liệu ghép nhóm là

$$S_B^2 = \frac{1}{20} (3 \cdot 5,5^2 + 6 \cdot 6,5^2 + 5 \cdot 7,5^2 + 5 \cdot 8,5^2 + 1 \cdot 9,5^2) - (7,25)^2 = 1,2875$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S_B = \sqrt{1,2875}$

d) Đúng: Do $S_A > S_B$ nên nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì mức lương khởi điểm của công nhân khu vực B đồng đều hơn của công nhân khu vực A .

Câu 9: Giá đóng cửa của một cổ phiếu là giá của cổ phiếu đó cuối một phiên giao dịch. Bảng sau thống kê giá đóng cửa (đơn vị: nghìn đồng) của hai mã cổ phiếu A và B trong 50 ngày giao dịch liên tiếp.

Giá đóng cửa	[120;122)	[122;124)	[124;126)	[126;128)	[128;130)
Cổ phiếu A	8	9	12	10	11
Cổ phiếu B	16	4	3	6	21

Xét tính đúng sai của các khẳng định sau:

a) Xét mẫu số liệu của cổ phiếu A ta có phương sai của mẫu số liệu ghép nhóm là 7,5216

b) Xét mẫu số liệu của cổ phiếu B ta có số trung bình của mẫu số liệu ghép nhóm là 115,28.

c) Xét mẫu số liệu của cổ phiếu B ta có độ lệch chuẩn của mẫu số liệu ghép nhóm là $\sqrt{15,4096}$

d) Người ta có thể dùng phương sai và độ lệch chuẩn để so sánh mức độ rủi ro của các loại cổ phiếu có giá trị trung bình gần bằng nhau. Cổ phiếu nào có phương sai, độ lệch chuẩn cao hơn thì được coi là có độ rủi ro lớn hơn. Theo quan điểm trên, thì cổ phiếu A có độ rủi ro thấp hơn cổ phiếu B .



Lời giải

Ta có bảng thống kê giá đóng cửa theo giá trị đại diện:

Giá đóng cửa	121	123	125	127	129
Cổ phiếu A	8	9	12	10	11
Cổ phiếu B	16	4	3	6	21

a) Đúng: Xét mẫu số liệu của cổ phiếu A :

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_1 = \frac{8.121 + 9.123 + 12.125 + 10.127 + 11.129}{50} = 125,28.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_1^2 = \frac{1}{50} (8.121^2 + 9.123^2 + 12.125^2 + 10.127^2 + 11.129^2) - (125,28)^2 = 7,5216$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_1 = \sqrt{S_1^2} = \sqrt{7,5216}$.

b) Sai: Xét mẫu số liệu của cổ phiếu B :

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_2 = \frac{16.121 + 4.123 + 3.125 + 6.127 + 21.129}{50} = 125,28.$$

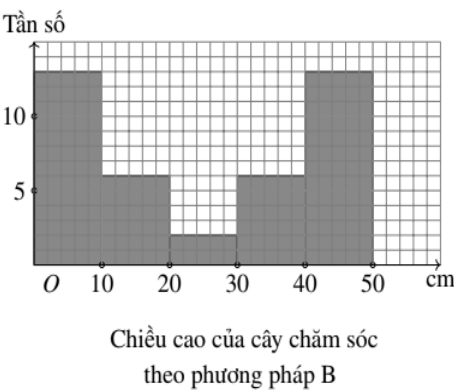
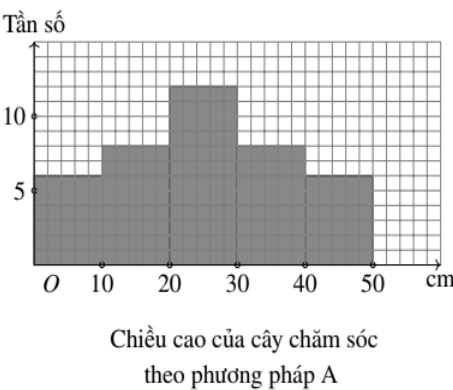
c) Sai: Phương sai của mẫu số liệu ghép nhóm là

$$S_2^2 = \frac{1}{50} (16.121^2 + 4.123^2 + 3.125^2 + 6.127^2 + 21.129^2) - (125,28)^2 = 12,4096.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_2 = \sqrt{S_2^2} = \sqrt{12,4096}$.

d) Vậy nếu đánh giá độ rủi ro theo phương sai và độ lệch chuẩn thì cổ phiếu A có độ rủi ro thấp hơn cổ phiếu B .

Câu 10: Một công ty giống cây trồng đã thử nghiệm hai phương pháp chăm sóc khác nhau cho cây hướng dương. Sau hai tuần, người ta thấy cây được chăm sóc theo cả hai phương pháp đều thấp hơn 50 cm.





- a) Khoảng biến thiên của chiều cao các cây được chăm sóc theo mỗi phương pháp A và B bằng nhau.
- b) Trung bình của chiều cao các cây được chăm sóc theo mỗi phương pháp A và B bằng nhau.
- c) Độ lệch chuẩn của chiều cao các cây được chăm sóc theo phương án A là 12,65 (cm).
- d) Dựa vào độ lệch chuẩn thì chiều cao của các loại cây được chăm sóc theo phương án B ít bị chênh lệch hơn so với phương án A .

Lời giải

a) Đúng: Khoảng biến thiên của chiều cao các cây được chăm sóc theo mỗi phương pháp A và B bằng nhau và cùng bằng 50.

b) Đúng: Ước tính số trung bình và độ lệch chuẩn của chiều cao các cây được chăm sóc theo mỗi phương pháp. Cỡ mẫu của hai mẫu số liệu thống kê là $N = 40$. Ta có bảng tần số ghép nhóm về chiều cao của cây được chăm sóc theo phương pháp A như sau:

Chiều cao (cm)	[0; 10)	[10; 20)	[20; 30)	[30; 40)	[40; 50)
Giá trị đại diện	5	15	25	35	45
Tần số	6	8	12	8	6

Chiều cao trung bình của các cây được chăm sóc theo phương án A là:

$$\bar{x}_A = \frac{5.6 + 15.8 + 25.12 + 35.8 + 45.6}{40} = 25$$

Ta có bảng tần số ghép nhóm về chiều cao của cây được chăm sóc theo phương pháp B như sau:

Chiều cao (cm)	[0; 10)	[10; 20)	[20; 30)	[30; 40)	[40; 50)
Giá trị đại diện	5	15	25	35	45
Tần số	13	6	2	6	13

Chiều cao trung bình của các cây được chăm sóc theo phương án B là:

$$\bar{x}_B = \frac{5.13 + 15.6 + 25.2 + 35.6 + 45.13}{40} = 25 \text{ cm.}$$

c) Độ lệch chuẩn của chiều cao các cây được chăm sóc theo phương án A là:

$$s_A = \sqrt{\frac{5^2.6 + 15^2.8 + 25^2.12 + 35^2.8 + 45^2.6}{40} - 25^2} \approx 12,65$$

d) Độ lệch chuẩn của chiều cao các cây được chăm sóc theo phương án B là:

$$s_B = \sqrt{\frac{5^2.13 + 15^2.6 + 25^2.2 + 35^2.6 + 45^2.13}{40} - 25^2} \approx 17,03 \text{ cm}^2$$

**PHẦN III. Câu trắc nghiệm trả lời ngắn****Câu 1:** Hãy tính độ lệch chuẩn của mẫu số liệu ghép nhóm (kết quả được làm tròn đến hàng phần trăm)

Chiều cao (cm)	[160;164)	[164;168)	[168;172)	[172;176)	[176;180)
Số học sinh	3	5	8	4	1

Lời giảiCỡ mẫu: $n = 21$. Giá trị trung bình của mẫu số liệu mới:

$$\bar{x} = \frac{1}{n}(n_1c_1 + n_2c_2 + \dots + n_kc_k) = \frac{1}{21}(3.162 + 5.166 + 8.170 + 4.174 + 1.178) = \frac{3550}{21}$$

Phương sai của mẫu số liệu mới: $S^2 = \frac{1}{n}[n_1(c_1 - \bar{x})^2 + n_2(c_2 - \bar{x})^2 + \dots + n_k(c_k - \bar{x})^2]$

$$= \frac{1}{21}\left[3\left(162 - \frac{3550}{21}\right)^2 + 5\left(166 - \frac{3550}{21}\right)^2 + \dots + 1\left(178 - \frac{3550}{21}\right)^2\right] = \frac{8000}{441}$$

Độ lệch chuẩn của mẫu số liệu mới: $\sigma = \sqrt{S^2} = \sqrt{\frac{8000}{441}} = \frac{40\sqrt{5}}{21} \approx 4,26$ **Câu 2:** Để đánh giá chất lượng một loại pin điện thoại mới, người ta ghi lại thời gian nghe nhạc liên tục của điện thoại được sạc đầy pin cho đến khi hết pin cho kết quả như sau:

Thời gian (giờ)	[5;5,5)	[5,5;6)	[6;6,5)	[6,5;7)	[7;7,5)
Số chiếc điện thoại (tần số)	2	8	15	10	5

Tính độ lệch chuẩn của mẫu số liệu ghép nhóm trên (kết quả được làm tròn đến hàng phần trăm)

Lời giải

Chọn giá trị đại diện cho các nhóm số liệu, ta có:

Thời gian (giờ)	[5;5,5)	[5,5;6)	[6;6,5)	[6,5;7)	[7;7,5)
Giá trị đại diện	5,25	5,75	6,25	6,75	7,25
Số chiếc điện thoại (tần số)	2	8	15	10	5

Thời gian trung bình nghe nhạc liên tục của điện thoại là:

$$\bar{x} = \frac{1}{40}(5,25.2 + 5,75.8 + 6,25.15 + 6,75.10 + 7,25.5) = 6,35$$

Phương sai của mẫu số liệu là:

$$s^2 = \frac{1}{40}(5,25^2.2 + 5,75^2.8 + 6,25^2.15 + 6,75^2.10 + 7,25^2.5) - 6,35^2 = 0,2775$$

Độ lệch chuẩn của mẫu số liệu là: $\sqrt{0,2775} = \frac{\sqrt{111}}{20} \approx 0,53$



Câu 3: Bảng dưới đây thống kê cự li ném tạ của một vận động viên.

Cự li (m)	[19;19,5)	[19,5;20)	[20;20,5)	[20,5;21)	[21;21,5)
Tần số	13	45	24	12	6

Hãy tính độ lệch chuẩn của mẫu số liệu ghép nhóm trên (kết quả được làm tròn đến hàng phần trăm)

Lời giải

Giá trị đại diện	19,25	19,75	20,25	20,75	21,25
Tần số	13	45	24	12	6

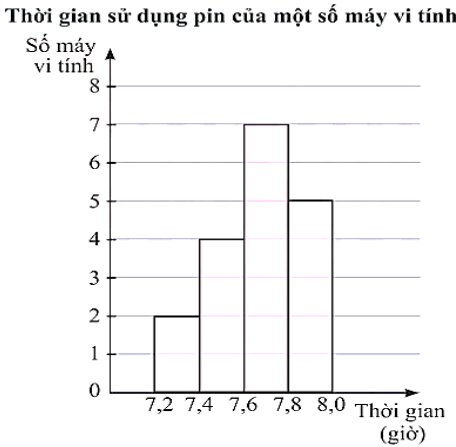
Cỡ mẫu: $n = 100$

Số trung bình: $\bar{x} = \frac{13.19,25 + 45.19,75 + 24.20,25 + 12.20,75 + 6.21,25}{100} = 20,015$

Phương sai: $S^2 = \frac{13.19,25^2 + 45.19,75^2 + 24.20,25^2 + 12.20,75^2 + 6.21,25^2}{100} - 20,015^2 \approx 0,28$

Độ lệch chuẩn: $\sigma = \sqrt{0,28} \approx 0,53$

Câu 4: Kết quả khảo sát thời gian sử dụng liên tục (đơn vị: giờ) từ lúc sạc đầy cho đến khi hết của pin một số máy vi tính cùng loại được mô tả bằng biểu đồ bên.



Hãy xác định độ lệch chuẩn của thời gian sử dụng pin (kết quả được làm tròn đến hàng phần trăm)

Lời giải

Giá trị đại diện	7,3	7,5	7,7	7,9
Số máy	2	4	7	5

Cỡ mẫu: $n = 18$

Số trung bình: $\bar{x} = \frac{2.7,3 + 4.7,5 + 7.7,7 + 5.7,9}{18} \approx 7,67$



Phương sai: $S^2 = \frac{2.7,3^2 + 4.7,5^2 + 7.7,7^2 + 5.7,9^2}{18} - 7,67^2 \approx 0,04$

Độ lệch chuẩn: $\sigma = \sqrt{0,04} \approx 0,19$

Câu 5: Thành tích môn nhảy cao của các vận động viên tại một giải điền kinh dành cho học sinh trung học phổ thông như sau:

Mức xà (cm)	[170;172)	[172;174)	[174;176)	[176;180)
Số vận động viên	3	10	6	1

Hãy xác định độ lệch chuẩn của thời gian sử dụng pin (kết quả được làm tròn đến hàng phần trăm)

Lời giải

Mẫu số liệu với giá trị đại diện

Mức xà (cm)	[170;172)	[172;174)	[174;176)	[176;178)
Giá trị đại diện	171	173	175	177
Số vận động viên	3	10	6	1

Giá trị trung bình: $\bar{x} = \frac{1}{20}(171.3 + 173.10 + 175.6 + 177.1) = 173,5(\text{cm})$

Phương sai của mẫu số liệu: $s^2 = \frac{1}{20}(171^2.3 + 173^2.10 + 175^2.6 + 177^2.1) - 173,5^2 = 2,35$

Độ lệch chuẩn của mẫu số liệu: $s = \sqrt{2,35} = \frac{\sqrt{235}}{10} \approx 1,53(\text{cm})$

Câu 6: Chiều dài của 40 bé trai sơ sinh 12 ngày tuổi chọn ngẫu nhiên ở một bệnh viện được nhà nghiên cứu thống kê trong bảng dưới đây:

Chiều dài (cm)	[44; 46)	[46; 48)	[48; 50)	[50; 52)	[52; 54)	[54; 56)
Số trẻ	3	3	10	15	7	2

Tính độ lệch chuẩn của chiều dài nhóm 40 bé trai sơ sinh (làm tròn kết quả đến hàng phần nghìn).

Lời giải

Bổ sung thêm các giá trị đại diện vào bảng trên ta lập được bảng sau. Chiều dài của 40 bé trai sơ sinh (đơn vị: cm)

Nhóm	c_i	n_i
[44; 46)	45	3
[46; 48)	47	3
[48; 50)	49	10
[50; 52)	51	15
[52; 54)	53	7
[54; 56)	55	2
		$N = 40$



Từ mẫu số liệu đã cho, ta tính được số trung bình là:

$$\bar{x} = \frac{3.45 + 3.47 + 10.49 + 15.51 + 7.53 + 2.55}{40} = \frac{2012}{40} = 50,3.$$

$\bar{x}$ không phải là số nguyên nên để tính phương sai ta tính:

$$\overline{x^2} = \frac{3.45^2 + 3.47^2 + 10.49^2 + 15.51^2 + 7.53^2 + 2.55^2}{40} = 2536.$$

Do đó: $s^2 = \overline{x^2} - (\bar{x})^2 = 2536 - 50,3^2 = 2536 - 2530,09 = 5,91.$

Vậy mẫu số liệu về chiều dài của 40 trẻ sơ sinh có độ lệch chuẩn là $s = \sqrt{5,91} \approx 2,43.$

Câu 7: Một bác tài xế thống kê lại độ dài quãng đường (đơn vị: km) bác đã lái xe mỗi ngày trong một tháng ở bảng sau:

Độ dài quãng đường (km)	[50;100)	[100;150)	[150;200)	[200;250)	[250;300)
Số ngày	5	10	9	9	2

Hãy xác định độ lệch chuẩn của mẫu số liệu trên

Lời giải

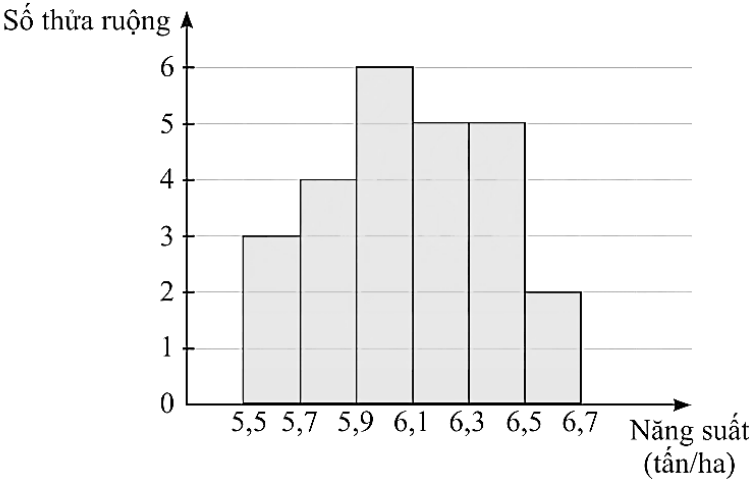
Độ dài quãng đường (km)	75	125	175	225	275
Số ngày	5	10	9	9	2

Số trung bình: $\bar{x} = \frac{5.75 + 10.125 + 9.175 + 4.225 + 2.275}{30} = 155$

Độ lệch chuẩn: $\sigma = \sqrt{\frac{5.75^2 + 10.125^2 + 9.175^2 + 4.225^2 + 2.275^2}{30} - 155^2} = 3100$

Câu 8: Kết quả khảo sát năng suất (đơn vị: tấn/ha) của một số thửa ruộng được minh họa ở biểu đồ sau.

Năng suất lúa của một số thửa ruộng



Người ta lập được bảng tần số ghép nhóm và tần số tương đối ghép nhóm tương ứng của mẫu số liệu trên như sau:



Năng suất	[5, 5; 5, 7)	[5, 7; 5, 9)	[5, 9; 6, 1)	[6, 1; 6, 3)	[6, 3; 6, 5)	[6, 5; 6, 7)
Số thửa ruộng	3	4	6	5	5	2
Giá trị đại diện	5,6	5,8	6,0	6,2	6,4	6,6
Tần số tương đối	3	4	6	5	5	2

Hãy xác định độ lệch chuẩn của mẫu số liệu trên (làm tròn kết quả đến hàng phần nghìn).

Lời giải

$$\text{Số trung bình: } \bar{x} = \frac{3 \cdot 5,6 + 4 \cdot 5,8 + 6 \cdot 6,0 + 5 \cdot 6,2 + 5 \cdot 6,4 + 2 \cdot 6,6}{25} = 6,088$$

$$\text{Độ lệch chuẩn: } \sigma = \sqrt{\frac{5 \cdot 75^2 + 10 \cdot 125^2 + 9 \cdot 175^2 + 4 \cdot 225^2 + 2 \cdot 275^2}{30} - 155^2} \approx 0,29$$

Câu 9: Một vận động viên luyện tập chạy cự li 100 m đã ghi lại kết quả luyện tập như sau:

Thời gian (giây)	[10, 2; 10, 4)	[10, 4; 10, 6)	[10, 6; 10, 8)	[10, 8; 11)
Số vận động viên	3	7	8	2

Hãy xác định phương sai của mẫu số liệu trên?

Lời giải

Chọn giá trị đại diện cho các nhóm số liệu, ta có:

Thời gian (giây)	[10, 2; 10, 4)	[10, 4; 10, 6)	[10, 6; 10, 8)	[10, 8; 11)
Giá trị đại diện	10,3	10,5	10,7	10,9
Số vận động viên	3	7	8	2

Tổng số vận động viên là: $3 + 7 + 8 + 2 = 20$

Thời gian chạy trung bình của các vận động viên là:

$$\bar{x} = \frac{1}{20} (10,3 \cdot 3 + 10,5 \cdot 7 + 10,7 \cdot 8 + 10,9 \cdot 2) = 10,59 \text{ (giây)}$$

Phương sai của mẫu số liệu là:

$$s^2 = \frac{1}{20} (10,3^2 \cdot 3 + 10,5^2 \cdot 7 + 10,7^2 \cdot 8 + 10,9^2 \cdot 2) - 10,59^2 = 0,0299$$

Câu 10: Cân nặng của một số quả mít trong một khu vườn được thống kê ở bảng sau:

Cân nặng (kg)	[4; 6)	[6; 8)	[8; 10)	[10; 12)	[12; 14)
Số quả mít	6	12	19	9	4

Hãy tính phương sai của mẫu số liệu ghép nhóm trên. (Kết quả các phép tính làm tròn đến hàng phần trăm)?



Lời giải

Ta có bảng thống kê cân nặng của các quả mít theo giá trị đại diện:

Cân nặng đại diện (kg)	5	7	9	11	13
Tần số	6	12	19	9	4

Cỡ mẫu $n = 6 + 12 + 19 + 9 + 4 = 50$.

Số trung bình của mẫu số liệu ghép nhóm là: $\bar{x} = \frac{6.5 + 12.7 + 19.9 + 9.11 + 4.13}{50} = 8,72$.

Phương sai của mẫu số liệu ghép nhóm là:

$$S^2 = \frac{1}{50} \left(6.5^2 + 12.7^2 + 19.9^2 + 9.11^2 + 4.13^2 \right) - 8,72^2 \approx 4,80.$$

-----HẾT-----