

**TRẮC NGHIỆM ĐÚNG SAI CÁC SỐ ĐẶC TRƯNG ĐO MỨC ĐỘ PHÂN TÁN
CỦA MẪU SỐ LIỆU GHEP NHÓM**

Câu 1: Phòng bạn Hùng và bạn Vương làm thí nghiệm trồng cây. Mỗi bạn trồng 40 cây cần tây trong cốc, phần gốc của các cây khi bắt đầu trồng đều dài 4 cm. Bảng 1 và bảng 2 lần lượt biểu diễn mẫu số liệu ghép nhóm về số liệu thống kê chiều cao của các cây (đơn vị: centimet) mà bạn Hùng và bạn Vương trồng sau 5 tuần.

Nhóm	Tần số
[20; 25)	2
[25; 30)	16
[30; 35)	20
[35; 40)	2
	$n = 40$

Bảng 1.

Nhóm	Tần số
[20; 25)	5
[25; 30)	9
[30; 35)	25
[35; 40)	1
	$n = 40$

Bảng 2.

- Chiều cao trung bình của mỗi cây do hai bạn Hùng và Vương trồng không bằng nhau.
- Khoảng biến thiên của cả hai mẫu số liệu trên là 20.
- Khoảng tứ phân vị của mẫu số liệu ở bảng 1 là 5,5.
- Chiều cao của các cây mà bạn Vương trồng đồng đều hơn các cây mà bạn Hùng trồng.

Lời giải

a) Chiều cao trung bình của cây do bạn Hùng trồng là: $\overline{x_H} = 30,25 \text{ (cm)}$.

Chiều cao trung bình của cây do bạn Vương trồng là: $\overline{x_V} = 30,25 \text{ (cm)}$.

Suy ra chiều cao trung bình của mỗi cây do hai bạn Hùng và Vương trồng là bằng nhau.

b) Khoảng biến thiên của cả hai mẫu số liệu là $40 - 20 = 20$.

c) Xét mẫu số liệu ở bảng 1.

Tứ phân vị thứ nhất Q_1 của mẫu số liệu đó là:

$$Q_1 = 25 + \left(\frac{10 - 2}{16} \right) \cdot 5 = 27,5 \text{ (cm)}.$$

Tứ phân vị thứ ba Q_3 của mẫu số liệu đó là:

$$Q_3 = 30 + \left(\frac{30 - 18}{20} \right) \cdot 5 = 33 \text{ (cm)}.$$

Suy ra khoảng tứ phân vị của mẫu số liệu ở bảng 1 là $33 - 27,5 = 5,5$.

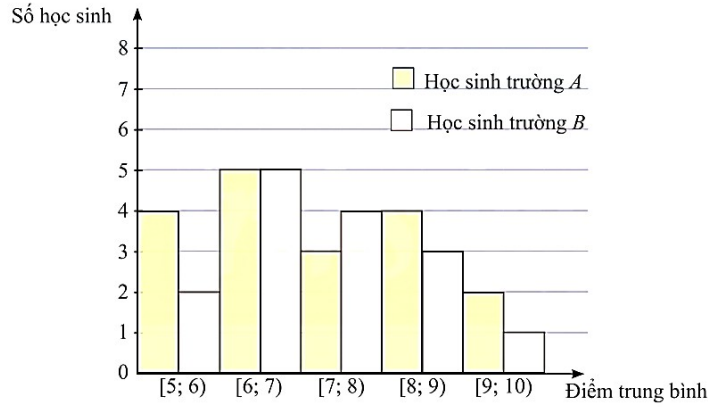
c) Phương sai của mẫu số liệu ở bảng 1 là: $s_H^2 = 11,1875$.

Phương sai của mẫu số liệu ở bảng 2 là: $s_V^2 = 13,6875$.

Suy ra $s_H^2 < s_V^2$. Vậy chiều cao của các cây mà bạn Hùng trồng đồng đều hơn các cây mà bạn Vương trồng.

Câu 2: Biểu đồ sau mô tả kết quả điều tra về điểm trung bình năm học của học sinh hai trường A và B.

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

Các mệnh đề sau đúng hay sai?

- 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: 1,23
- 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	b) Sai	c) Đúng	d) Sai
---------	--------	---------	--------

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: $x_1, \dots, x_4 \in [5; 6); x_5, \dots, x_9 \in [6; 7); x_{10}, \dots, x_{12} \in [7; 8); x_{13}, \dots, x_{16} \in [8; 9); x_{17}, x_{18} \in [9; 10)$

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

$$Q_1 = 6 + \frac{\frac{18}{4} - 4}{5} (7 - 6) = 6,1$$

nhóm là:

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

$$Q_3 = 8 + \frac{\frac{3 \cdot 18}{4} - (4 + 5 + 3)}{4} (9 - 8) = 8,375$$

là:

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 2,275$

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

$$Q_1' = 6 + \frac{\frac{15}{4} - 2}{5} (7 - 6) = 6,35$$

nhóm là:

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

$$Q_3' = 8 + \frac{\frac{3 \cdot 15}{4} - (2 + 5 + 4)}{3} (9 - 8) = 8,08$$

là:

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q' = Q_3' - Q_1' = 1,73$

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

Xét số liệu của trường A:

$$\bar{x}_A = \frac{4,5,5 + 5,6,5 + 3,7,5 + 4,8,5 + 2,9,5}{18} = 7,22$$

Số trung bình:

$$\sigma_A = \sqrt{\frac{4,5,5^2 + 5,6,5^2 + 3,7,5^2 + 4,8,5^2 + 2,9,5^2}{18} - 7,22^2} \approx 1,79$$

Độ lệch chuẩn:

Xét số liệu của trường B:

$$\bar{x}_B = \frac{2,5,5 + 5,6,5 + 4,7,5 + 3,8,5 + 1,9,5}{15} = 7,23$$

Số trung bình:

$$\sigma_B = \sqrt{\frac{2,5,5^2 + 5,6,5^2 + 4,7,5^2 + 3,8,5^2 + 1,9,5^2}{15} - 7,23^2} \approx 1,31$$

Độ lệch chuẩn:

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 3: Bảng 1, Bảng 2 lần lượt 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	Giá trị đại diện	Tần số
[16, 8; 19, 8)	18,3	2
[19, 8; 22, 8)	21,3	3
[22, 8; 25, 8)	24,3	2
[25, 8; 28, 8)	27,3	1
[28, 8; 31, 8)	30,3	4
		$n = 12$

Bảng 1

Nhóm	Giá trị đại diện	Tần số
[16, 8; 19, 8)	18,3	1
[19, 8; 22, 8)	21,3	2
[22, 8; 25, 8)	24,3	3
[25, 8; 28, 8)	27,3	2
[28, 8; 31, 8)	30,3	4
		$n = 12$

Bảng 2

Các mệnh đề sau đúng hay sai?

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

Lời giải

a) Đúng	b) Sai	c) Sai	d) Đúng
---------	--------	--------	---------

- Xét số liệu ở Hà Nội:

+ 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 $cf_1 = 2, cf_2 = 5, cf_3 = 7, cf_4 = 8, cf_5 = 12$

$$\frac{n}{4} = \frac{12}{4} = 3$$

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 $cf_1 = 2$

$$Q_1 = s + \left(\frac{3 - cf_1}{n_2} \right) \cdot h = 19,8 + \left(\frac{3 - 2}{3} \right) \cdot 3 = 20,8$$

Ta có tứ phân vị thứ nhất là:

$$\frac{3n}{4} = \frac{3 \cdot 12}{4} = 9$$

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 $cf_4 = 8$

$$Q_3 = t + \left(\frac{9 - cf_4}{n_5} \right) \cdot l = 28,8 + \left(\frac{9 - 8}{4} \right) \cdot 3 = 29,55$$

Ta có tứ phân vị thứ ba là:

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $Q_3 - Q_1 = 29,55 - 20,8 = 8,75$

$$\bar{x}_1 = \frac{2 \cdot 18,3 + 3 \cdot 21,3 + 2 \cdot 24,3 + 27,3 + 4 \cdot 30,3}{12} = 24,8$$

+ Số trung bình cộng của mẫu số liệu ghép nhóm là:

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: $s_1 = \sqrt{s_1^2} = \sqrt{20,75} \approx 4,56$

- 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 $cf_1 = 1, cf_2 = 3, cf_3 = 6, cf_4 = 8, cf_5 = 12$

$$\frac{n}{4} = \frac{12}{4} = 3$$

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 $cf_1 = 1$

$$Q_1 = s + \left(\frac{3 - cf_1}{n_2} \right) \cdot h = 19,8 + \left(\frac{3 - 1}{2} \right) \cdot 3 = 22,8$$

Ta có tứ phân vị thứ nhất là:

$$\frac{3n}{4} = \frac{3 \cdot 12}{4} = 9$$

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 $cf_4 = 8$

$$Q_3 = t + \left(\frac{9 - cf_4}{n_5} \right) \cdot I = 28,8 + \left(\frac{9 - 8}{4} \right) \cdot 3 = 29,55$$

Ta có tứ phân vị thứ ba là:

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $Q_3 - Q_1 = 29,55 - 22,8 = 6,75$

$$\bar{x}_2 = \frac{18,3 + 2 \cdot 21,3 + 3 \cdot 24,3 + 2 \cdot 27,3 + 4 \cdot 30,3}{12} = 25,8$$

+ Số trung bình cộng của mẫu số liệu ghép nhóm là:

Phương sai của mẫu số liệu ghép nhóm là:

$$s_2^2 = \frac{(18,3 - 25,8)^2 + 3(21,3 - 25,8)^2 + 3(24,3 - 25,8)^2 + 2(27,3 - 25,8)^2 + 4(30,3 - 25,8)^2}{12} = 15,75$$

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 4: Thầy Tuấn thống kê lại điểm trung bình cuối năm của các học sinh lớp 11A và 11B ở bảng sau:

Điểm trung bình \ Lớp	[5; 6)	[6; 7)	[7; 8)	[8; 9)	[9; 10)
11A	1	0	11	22	6
11B	0	6	8	14	12

Các mệnh đề sau đúng hay sai?

a) Khoảng biến thiên của điểm số học sinh lớp 11A là: 5.

b) Nếu so sánh theo khoảng biến thiên thì điểm trung bình của các học sinh lớp 11B ít phân tán hơn điểm trung bình của các học sinh lớp 11A.

c) Xét mẫu số liệu của lớp 11A ta có độ lệch chuẩn của mẫu số liệu ghép nhóm là $\sqrt{0,51}$

d) Nếu so sánh theo độ lệch chuẩn thì học sinh lớp 11A có điểm trung bình ít phân tán hơn học sinh lớp 11B.

Lời giải

a) Đúng	b) Đúng	c) Sai	d) Đúng
---------	---------	--------	---------

-Khoảng biến thiên của điểm số học sinh lớp 11A là: $10 - 5 = 5$.

Khoảng biến thiên của điểm số học sinh lớp 11B là: $10 - 6 = 4$.

Nếu so sánh theo khoảng biến thiên thì điểm trung bình của các học sinh lớp 11B ít phân tán hơn điểm trung bình của các học sinh lớp 11A.

-Ta có bảng thống kê điểm trung bình theo giá trị đại diện:

Giá trị đại diện \ Lớp	5,5	6,5	7,5	8,5	9,5
11A	1	0	11	22	6
11B	0	6	8	14	12

- Xét mẫu số liệu của lớp 11A:

Cỡ mẫu là $n_1 = 1 + 11 + 22 + 6 = 40$.

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_1 = \frac{1 \cdot 5,5 + 11 \cdot 7,5 + 22 \cdot 8,5 + 6 \cdot 9,5}{40} = 8,3.$$

Phương sai của mẫu số liệu ghép nhóm là

$$s_1^2 = \frac{1}{40} (1 \cdot 5,5^2 + 11 \cdot 7,5^2 + 22 \cdot 8,5^2 + 6 \cdot 9,5^2) - 8,3^2 = 0,61.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_1 = \sqrt{0,61}$.

- Xét mẫu số liệu của lớp 11B:

Cỡ mẫu là $n_2 = 6 + 8 + 14 + 12 = 40$.

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_2 = \frac{6 \cdot 6,5 + 8 \cdot 7,5 + 14 \cdot 8,5 + 12 \cdot 9,5}{40} = 8,3.$$

Phương sai của mẫu số liệu ghép nhóm là

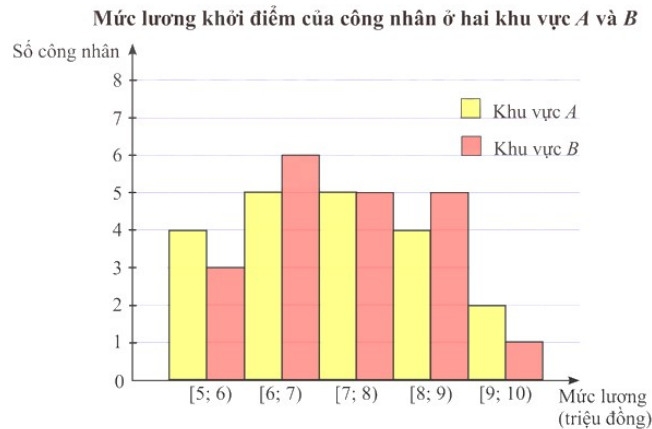
$$S_2^2 = \frac{1}{40} (6 \cdot 6,5^2 + 8 \cdot 7,5^2 + 14 \cdot 8,5^2 + 12 \cdot 9,5^2) - 8,3^2 = 1,06.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_2 = \sqrt{1,06}$.

Do $S_1 < S_2$ nên nếu so sánh theo độ lệch chuẩn thì học sinh lớp 11A có điểm trung bình ít phân tán hơn học sinh lớp 11B.

Chú ý: Trong bài toán trên, kết quả so sánh độ phân tán theo giá trị trung bình và độ lệch chuẩn có sự khác biệt. Điều này là do mẫu số liệu của học sinh lớp 11A có một giá trị ngoại lệ.

Câu 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 .



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

Các mệnh đề sau đúng hay sai?

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	b) Đúng	c) Sai	d) Đúng
--------	---------	--------	---------

- 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

$$S_A^2 = \frac{1}{20} (4.5,5^2 + 5.6,5^2 + 5.7,5^2 + 4.8,5^2 + 2.9,5^2) - (7,25)^2 = 1,5875$$

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

$$\bar{x}_B = \frac{3.5,5 + 6.6,5 + 5.7,5 + 5.8,5 + 1.9,5}{20} = 7,25.$$

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.

Câu 6: Số tiền thưởng cuối năm của nhân viên công ty X được thống kê như sau (đơn vị triệu đồng):

3	5	7	12	15	11	4	5	8	15
16	18	20	22	15	29	28	22	24	26
16	21	25	26	29	6	8	10	17	18

a, Hãy chuyển mẫu số liệu sang dạng ghép nhóm với 6 nhóm có độ dài bằng nhau.

b, Các câu sau là Đúng hay Sai? (làm tròn kết quả đến hàng phần mười)

b1, Tiền thưởng trung bình của các nhân viên là $16,2$.

b2, Khoảng biến thiên của mẫu số liệu ghép nhóm là 26.

b3, Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $14,5$.

Lời giải

a, Cỡ mẫu: $n = 30$

+ Sắp xếp mẫu số liệu theo thứ tự tăng dần ta được:

3	4	5	5	6	7	8	8	10	11
12	15	15	15	16	16	17	18	18	20
21	22	22	24	25	26	26	28	28	29

+ Giá trị lớn nhất của mẫu số liệu là 29 , và giá trị nhỏ nhất của mẫu số liệu là 3 , nên độ dài nhóm là $29 - 3 = 26$

+ Chia thành 6 nhóm có độ dài bằng nhau, ta chọn độ dài mỗi nhóm là $4,5$, ta có mẫu số liệu ghép nhóm như bảng sau:

Nhóm	$[3; 7,5)$	$[7,5; 12)$	$[12; 16,5)$	$[16,5; 21)$	$[21; 25,5)$	$[25,5; 30)$
Tần số	6	4	6	4	5	5

$$\bar{X} = \frac{m_1x_1 + m_2x_2 + \dots + m_6x_6}{n} = 16,2$$

b, + Tiền thưởng trung bình là:

+ Khoảng biến thiên: $30 - 3 = 27$.

+ Khoảng tứ phân vị:

Tứ phân vị thứ nhất Q_1 là x_8 thuộc nhóm $[7,5;12)$ nên nhóm thứ 2 này chứa Q_1 .

Suy ra: $p = 2, a_2 = 7,5; a_3 = 12; m_2 = 4; m_1 = 6$

$$Q_1 = a_2 + \frac{\frac{n}{4} - m_1}{m_2} (a_3 - a_2) = 7,5 + \frac{\frac{30}{4} - 6}{4} (12 - 7,5) = \frac{147}{16} = 9,2$$

Ta có:

Tứ phân vị thứ nhất Q_3 là x_{23} thuộc nhóm $[21;25,5)$ nên nhóm thứ 5 này chứa Q_3 .

Suy ra: $p = 5, a_5 = 21; a_6 = 25,5; m_5 = 5; m_4 = 4$

$$Q_3 = a_5 + \frac{\frac{3n}{4} - (m_1 + m_2 + m_3 + m_4)}{m_5} (a_6 - a_5) = 21 + \frac{\frac{3 \cdot 30}{4} - 20}{5} \cdot 4,5 = \frac{93}{4} = 23,3$$

Ta có:

Vậy khoảng tứ phân vị là: $\Delta_Q = Q_3 - Q_1 = 23,3 - 9,2 = 14,1$

Kết luận: b1, Đúng b2, Sai b3, Sai

Câu 7: Bảng 1 biểu diễn mẫu số liệu ghép nhóm về nhiệt độ ($^{\circ}C$) của tỉnh Nghệ An tháng 5 năm 2024

Nhóm	Giá trị đại diện	Tần số	Tần số tích lũy
$[29;31)$	30	1	1
$[31;33)$	32	4	5
$[33;35)$	34	5	10
$[35;37)$	36	13	26
$[37;39]$	38	7	33
		$n = 30$	

Trong mỗi ý a),b),c),d) chọn đúng hoặc sai (làm tròn đến hàng phần trăm)

- a) Nhóm $[31;33)$ có tần số bằng: 4
b) Một của mẫu số liệu ghép nhóm đã cho là: 13
c) Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên bằng: 2,92
d) Phương sai của mẫu số liệu ghép nhóm trên bằng: 4,57

Lời giải

a) **Đ.** Nhóm $[31;33)$ có tần số bằng: 4.

b) **S.** Ta có nhóm $[35;37)$ có tần số lớn nhất nên Một của mẫu số liệu trên là

$$M_0 = u + \frac{n_i - n_{i-1}}{2n_i - n_{i-1} - n_{i+1}} \cdot g = 35 + \frac{13 - 5}{2 \cdot 13 - 5 - 7} \cdot 2 = 36,14$$

c) **Đ** Ta có số phần tử của mẫu là $n = 30$.

+ Ta có $\frac{n}{4} = 7,5$ nên nhóm $[33;35)$ là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng $7,5$. Nhóm $[33;35)$ có $s = 33; h = 2; n = 5$ và nhóm 2 là nhóm $[31;33)$ có $cf_1 = 5$. Áp dụng công thức ta có tứ phân vị

$$\text{thứ nhất là } Q_1 = 33 + \frac{7,5 - 5}{5} \cdot 2 = 34 \quad (^{\circ}C)$$

+ Ta có $\frac{3n}{4} = 22,5$ nên nhóm $[35; 37)$ là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng $22,5$. Xét nhóm 4 là nhóm $[35; 37)$ có $t = 35; l = 2; n_4 = 13$ và nhóm 3 có tần số tích lũy $cf_4 = 10$. Áp dụng công thức, ta có tứ phân vị thứ 3 là $Q_3 = 35 + \frac{22,5 - 10}{13} \cdot 2 = 36,92$ ($^{\circ}C$)

+ Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm đã cho là $\Delta Q = Q_3 - Q_1 = 36,92 - 34 = 2,92$
 d) **Đ.** Ta có số trung bình cộng của mẫu số liệu ghép nhóm trên là

$$\bar{x} = \frac{1}{30}(1 \cdot 30 + 32 \cdot 4 + 34 \cdot 5 + 36 \cdot 13 + 38 \cdot 7) = 35,4 \quad (^{\circ}C)$$

Vậy phương sai của mẫu số liệu ghép nhóm trên là

$$s^2 = \frac{1}{30}(1(30 - 35,4)^2 + 4(32 - 35,4)^2 + 5(34 - 35,4)^2 + 13(36 - 35,4)^2 + 7(38 - 35,4)^2) = 4,57$$

Câu 8: Cho bảng phân bố tần số ghép lớp cân nặng (đơn vị: kg) của các công nhân trong một công ty như sau:

Các lớp giá trị của X	[50; 52)	[52; 54)	[54; 56)	[56; 58)	[58; 60)	Cộng
Tần số n_i	15	20	45	15	5	100

Các mệnh đề sau đúng hay sai?

- a) Tần suất của nhóm $[52; 54)$ là: 20% .
 b) Số trung vị của mẫu số liệu là: 54,909 .
 c) Khoảng biến thiên của mẫu số liệu trên là: 10 .
 d) Độ lệch chuẩn của mẫu số liệu trên là: 4,35 .

Lời giải

a) Đúng	b) Sai	c) Đúng	d) Sai
----------------	---------------	----------------	---------------

a) Tần số của nhóm $[52; 54)$ là: 20 .

Tần suất của nhóm $[52; 54)$ là: $\frac{20}{100} \cdot 100\% = 20\%$

b) Trung vị của mẫu số liệu là $x_3 \in [54; 56)$. Do đó, trung vị của mẫu số liệu ghép nhóm là:

$$M_e = Q_2 = 54 + \frac{\frac{2 \cdot 100}{4} - (15 + 20)}{45} (56 - 54) = \frac{164}{3} = 54,667$$

c) Khoảng biến thiên của mẫu số liệu là $R = 60 - 50 = 10$.

d) Số trung bình cộng của mẫu số liệu ghép nhóm của công ty là:

$$\bar{x} = \frac{51 \cdot 15 + 53 \cdot 20 + 55 \cdot 45 + 57 \cdot 15 + 59 \cdot 5}{100} = 54,5$$

Phương sai của mẫu số liệu ghép nhóm của công ty là:

$$s^2 = \frac{15 \cdot (51 - 54,5)^2 + 20 \cdot (53 - 54,5)^2 + 45 \cdot (55 - 54,5)^2 + 15 \cdot (57 - 54,5)^2 + 5 \cdot (59 - 54,5)^2}{100} = 4,35$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm của công ty là: $s = \sqrt{s^2} = \sqrt{4,35} \approx 2,085665$.

Câu 9: Cho bảng số liệu dưới đây về thời gian (phút) tập thể dục buổi sáng của hai bạn Bình và Chi trong 30 ngày.

Thời gian	[15; 20)	[20; 25)	[25; 30)	[30; 35)	[35; 40)
Bạn Bình	5	8	10	4	3
Bạn Chi	10	10	5	3	2

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 của Chi là 25 (phút).

b. Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bạn Bình là:

$$Q_1 = \frac{354}{16}$$

c. Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bạn Chi là

$$\frac{314}{9}$$

d. Phương sai của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bạn Bình là

Lời giải

a	b	c	d
Đúng	Sai	Đúng	Đúng

Ta có

Thời gian	[15; 20)	[20; 25)	[25; 30)	[30; 35)	[35; 40)
Giá trị đại diện	17,5	22,5	27,5	32,5	37,5
Bạn Bình	5	8	10	4	3
Bạn Chi	10	10	5	3	2

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

b. Ở bạn Bình. Ta có cỡ mẫu $n = 30$.

Vì $\frac{n}{4} = \frac{30}{4} = 7,5$ và $5 < 7,5 < 5 + 8$ nên tứ phân vị thứ nhất thuộc nhóm $[20; 25)$ và tứ phân vị thứ nhất là

$$Q_1 = 20 + \frac{\frac{30}{4} - 5}{8} \cdot 5 = \frac{345}{16}$$

c. Ở bạn Chi. Ta có cỡ mẫu $n = 30$.

Vì $\frac{n}{4} = \frac{30}{4} = 7,5$ và $7,5 < 10$ nên tứ phân vị thứ nhất thuộc nhóm $[15; 20)$ và tứ phân vị thứ nhất là

$$Q_1 = 15 + \frac{\frac{30}{4} - 0}{10} \cdot 5 = 18,75$$

Vì $\frac{3n}{4} = \frac{3 \cdot 30}{4} = 22,5$ và $10 + 10 < 22,5 < 10 + 10 + 5$ nên tứ phân vị thứ ba thuộc nhóm $[25; 30)$ và tứ phân

$$Q_3 = 25 + \frac{\frac{3 \cdot 30}{4} - (10 + 10)}{5} \cdot 5 = 27,5$$

vị thứ ba là

$$\Delta_Q = Q_3 - Q_1 = 28,75$$

Vậy khoảng tứ phân vị là

$$\bar{x} = \frac{17,5 \cdot 5 + 22,5 \cdot 8 + 27,5 \cdot 10 + 32,5 \cdot 4 + 37,5 \cdot 3}{30} = \frac{157}{6} \approx 26,17$$

Vậy phương sai của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bạn Bình là

$$s_B^2 = \frac{5 \left(17,5 - \frac{157}{6} \right)^2 + 8 \left(22,5 - \frac{157}{6} \right)^2 + 10 \left(27,5 - \frac{157}{6} \right)^2 + 4 \left(32,5 - \frac{157}{6} \right)^2 + 3 \left(37,5 - \frac{157}{6} \right)^2}{30} = \frac{314}{9}$$

Câu 10: Mẫu số liệu dưới đây ghi lại tốc độ của 40 ô tô khi đi qua một trạm đo tốc độ (đơn vị: km/h):

48,5 43 50 55 45 60 53 55,5 44 65
 51 62,5 41 44,5 57 57 68 49 46,5 53,5
 61 49,5 54 62 59 56 47 50 60 61
 49,5 52,5 57 47 60 55 45 47,5 48 61,5

a) Bảng tần số ghép nhóm cho mẫu số liệu trên có sáu nhóm ứng với sáu nửa khoảng là

Nhóm	Tần số
[40; 45)	4
[45; 50)	11
[50; 55)	7
[55; 60)	8
[60; 65)	8
[65; 70)	2
	$n = 40$

b) Mẫu số liệu trên có số trung bình là 54,875

c) Tứ phân vị của mẫu số liệu trên là: $Q_1 = 47,8(km / h); Q_2 = 53,6(km / h); Q_3 = 60(km / h)$.

d) Khoảng biến thiên của mẫu số liệu trên là 25.

Lời giải

a) Bảng tần số ghép nhóm:

Nhóm	Tần số
[40; 45)	4
[45; 50)	11
[50; 55)	7
[55; 60)	8
[60; 65)	8
[65; 70)	2
	$n = 40$

Vậy ý a ĐÚNG

b)S *Số trung bình

$$\bar{x} = \frac{4.42,5 + 11.47,5 + 7.52,5 + 8.57,5 + 8.62,5 + 2.67,5}{40} = 53,875 \text{ (km/h)}$$

Vậy ý b SAI

c)*Số phần tử của mẫu là $n = 40 \Rightarrow \frac{n}{2} = 20$.

Nhóm 3 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 20.

Xét nhóm 3 là nhóm $[50; 55)$ có $r = 50, d = 5, n_3 = 7$ và $cf_2 = 15$

Ta có số trung vị của mẫu số liệu là:

$$M_c = 50 + \frac{20 - 15}{7} \cdot 5 \approx 53,6 \text{ (km/h)}$$

*Số phần tử của mẫu là $n = 40$.

$\frac{n}{4} = 10$
Ta có:

Nhóm 2 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 10.

Xét nhóm 2 là nhóm $[45; 50)$ có $r = 45, d = 5, n_2 = 11$ và $cf_1 = 4$

Ta có tứ phân vị thứ nhất là:

$$Q_1 = 45 + \frac{10 - 4}{11} \cdot 5 \approx 47,8 \text{ (km/h)}$$

$\frac{3n}{4} = 30$
Ta có:

Nhóm 4 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 30.

Xét nhóm 4 là nhóm $[55; 60)$ có $r = 55, d = 5, n_4 = 8$ và $cf_3 = 22$

Ta có tứ phân vị thứ ba là:

$$Q_3 = 55 + \frac{30 - 22}{8} \cdot 5 = 60 \text{ (km/h)}$$

Vậy tứ phân vị của mẫu số liệu trên là: $Q_1 = 47,8 \text{ (km/h)}; Q_2 = 53,6 \text{ (km/h)}; Q_3 = 60 \text{ (km/h)}$.

Vậy ý c ĐÚNG

d)Khoảng biến thiên của mẫu số liệu là $R = a_7 - a_1 = 70 - 40 = 30$

Vậy ý d SAI

Câu 11: Bảng dưới đây cho ta bảng tần số ghép nhóm số liệu thống kê cân nặng của 40 học sinh lớp 12B trong một trường trung học phổ thông (đơn vị: kilôgam).

Nhóm	Số học sinh
$[30; 40)$	2
$[40; 50)$	10
$[50; 60)$	16
$[60; 70)$	8
$[70; 80)$	2
$[80; 90)$	2
	$n = 40$

Các mệnh đề sau đúng hay sai?

a) Số học sinh nặng dưới 50 kg là 12.

- b) Một của mẫu số liệu ghép nhóm trên xấp xỉ bằng $54,29(kg)$.
- c) Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là $\frac{39}{2}$.
- d) Phương sai của mẫu số liệu ghép nhóm là 128.

Lời giải

- a) Số học sinh nặng dưới $50kg$ là $2+10=12$. **Vậy ý a) đúng.**
- b)

Nhóm	Tần số
$[30; 40)$	2
$[40; 50)$	10
$[50; 60)$	16
$[60; 70)$	8
$[70; 80)$	2
$[80; 90)$	2
	$n = 40$

Nhóm chứa một của mẫu số liệu là $[50; 60)$.

Do đó $u_m = 50; n_m = 16; n_{m-1} = 10, n_{m+1} = 8, u_{m+1} - u_m = 60 - 50 = 10$.

Một của mẫu số liệu ghép nhóm xấp xỉ bằng:

$$M_0 = 50 + \frac{16-10}{(16-10)+(16-8)} \cdot 10 = \frac{380}{7} \approx 54,29(kg).$$

Một của mẫu số liệu ghép nhóm trên xấp xỉ bằng $54,29(kg)$. **Vậy ý b) đúng.**

c)

Nhóm	Tần số
$[30; 40)$	2
$[40; 50)$	10
$[50; 60)$	16
$[60; 70)$	8
$[70; 80)$	2
$[80; 90)$	2
	$n = 40$

Cỡ mẫu $n = 40$.

Gọi $x_1, x_2 \in [30; 40); x_3, \dots, x_{12} \in [40; 50); x_{13}, \dots, x_{28} \in [50; 60); x_{29}, \dots, x_{36} \in [60; 70); x_{37}, x_{38} \in [70; 80); x_{39}, x_{40} \in [80; 90)$.

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_{10} + x_{11}) \in [40; 50)$. Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là

$$Q_1 = 40 + \frac{\frac{40}{4} - 2}{10} \cdot (50 - 40) = 48.$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{30} + x_{31}) \in [60; 70)$. Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là

$$Q_3 = 60 + \frac{\frac{3 \cdot 40}{4} - (2 + 10 + 16)}{8} \cdot (70 - 60) = \frac{125}{2}.$$

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

$$\Delta_Q = \frac{125}{2} - 48 = \frac{29}{2}.$$

Vậy ý c) sai.

d)

Ta có bảng cân nặng của các em học sinh theo giá trị đại diện:

Nhóm	Giá trị đại diện	Tần số
$[30; 40)$	35	2
$[40; 50)$	45	10
$[50; 60)$	55	16
$[60; 70)$	65	8
$[70; 80)$	75	2
$[80; 90)$	85	2
		$n = 40$

Cỡ mẫu $n = 2 + 10 + 16 + 8 + 2 + 2 = 40$.

Số trung bình của mẫu số liệu ghép nhóm là

$$\frac{35 \cdot 2 + 45 \cdot 10 + 55 \cdot 16 + 65 \cdot 8 + 75 \cdot 2 + 85 \cdot 2}{40} = \frac{2240}{40} = 56 (kg)$$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{40} (2 \cdot 35^2 + 10 \cdot 45^2 + 16 \cdot 55^2 + 8 \cdot 65^2 + 2 \cdot 75^2 + 2 \cdot 85^2) - 56^2 = 3265 - 3136 = 129.$$

Vậy ý d) sai.

Câu 12: Trong một hội thao, thời gian chạy $200m$ của một nhóm các vận động viên được ghi lại ở bảng sau:

Thời gian (giây)	$[21; 21,5)$	$[21,5; 22)$	$[22; 22,5)$	$[22,5; 23)$	$[23; 23,5)$
Số vận động viên	5	10	30	45	30

Dựa vào bảng số liệu trên, em hãy xét tính đúng, sai của các mệnh đề sau:

- Tần suất của nhóm vận động viên chạy trong khoảng thời gian từ 22 giây đến dưới $22,5$ giây bằng 30%.
- Số trung vị của mẫu số liệu (làm tròn đến chữ số thập phân thứ 2) bằng $22,67$.
- Khoảng biến thiên của mẫu số liệu bằng $R = 2$.
- Độ lệch chuẩn của mẫu số liệu (làm tròn đến chữ số thập phân thứ 2) bằng $0,28$.

Lời giải

a) Cỡ mẫu: $n = 5 + 10 + 30 + 45 + 30 = 120$.

Tần suất của nhóm vận động viên chạy trong khoảng thời gian từ 22 giây đến dưới $22,5$ giây bằng:

$$f_3 = \frac{n_3}{n} = \frac{30}{120} = 25\%$$

Vậy mệnh đề a) sai.

b) Gọi $x_1, x_2, \dots, x_{120}$ là thời gian chạy của 120 vận động viên và dãy này là một dãy không giảm.

Khi đó trung vị là $\frac{x_{60} + x_{61}}{2}$. Do $x_{60}, x_{61} \in [22, 5; 23)$ nên nhóm này chứa trung vị. Ta có:

$$M_e = 22,5 + \frac{\frac{120}{2} - (5 + 10 + 30)}{45} \cdot (23 - 22,5) \approx 22,67$$

Vậy mệnh đề b) đúng.

c) Khoảng biến thiên của mẫu số liệu bằng: $R = 23,5 - 21 = 2,5$.

Vậy mệnh đề c) sai.

d) Giá trị trung bình của mẫu số liệu là:

$$\bar{x} = \frac{5 \cdot 21,25 + 10 \cdot 21,75 + 30 \cdot 22,25 + 45 \cdot 22,75 + 30 \cdot 23,25}{120} \approx 22,60$$

Độ lệch chuẩn của mẫu số liệu (làm tròn đến chữ số thập phân thứ 2) bằng

$$s = \sqrt{\frac{5(-1,35)^2 + 10(-0,85)^2 + 30(-0,35)^2 + 45(0,15)^2 + 30(0,65)^2}{120}} \approx 0,53$$

Vậy mệnh đề d) sai.

Câu 13: Thành tích chạy 50 m của 30 em học sinh lớp 10 trường THPT A được thống kê như bảng sau. (đơn vị: giây)

6,3	6,2	6,5	6,8	6,9	8,2
6,6	6,7	7,0	7,1	7,2	8,3
7,4	7,3	7,2	7,1	7,0	8,4
7,1	7,3	7,5	7,5	7,6	8,7
7,6	7,7	7,8	7,5	7,7	7,8

a. Bảng tần số ghép nhóm của mẫu số liệu trên là:

Lớp thời gian chạy (giây)	$[6,0; 6,5)$	$[6,5; 7,0)$	$[7,0; 7,5)$	$[7,5; 8,0)$	$[8,0; 8,5)$	$[8,5; 9,0)$
Tần số	2	5	10	9	3	1

b. Trung bình mỗi em chạy 50 m hết số thời gian là $7,5$ (giây).

c. Khoảng biến thiên của mẫu số liệu ghép nhóm trên là $R = 3,1$

d. Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là $\Delta_Q = 0,781$.

Lời giải

a. Đúng

Bảng tần số ghép nhóm của mẫu số liệu trên là:

Lớp thời gian chạy (giây)	$[6,0; 6,5)$	$[6,5; 7,0)$	$[7,0; 7,5)$	$[7,5; 8,0)$	$[8,0; 8,5)$	$[8,5; 9,0)$
Tần số	2	5	10	9	3	1

b. Sai

Tổng số học sinh là $n = 30$.

Trung bình mỗi em chạy 50 m hết số thời gian là:

$$\bar{x} = \frac{6,25 \cdot 2 + 6,75 \cdot 5 + 7,25 \cdot 10 + 7,75 \cdot 9 + 8,25 \cdot 3 + 8,75 \cdot 1}{30} = 7,4 \text{ (giây)}.$$

c. Sai

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là $R = 9,0 - 6,0 = 3,0$.

d. Đúng

- Cỡ mẫu là $n = 30$.

Gọi $x_1, \dots, x_{30}$ là thời gian chạy 50 m của 30 học sinh và giả sử dãy này đã được sắp xếp theo thứ tự tăng dần.

Khi đó, trung vị là $\frac{x_{15} + x_{16}}{2}$. Do 2 giá trị x_{15}, x_{16} thuộc nhóm $[7,0; 7,5)$ nên nhóm này chứa trung vị. Do đó:

$$M_e = 7,0 + \frac{\frac{30}{2} - (2 + 5)}{10} \cdot 0,5 = 7,4$$

Tứ phân vị thứ hai Q_2 chính là trung vị M_e .

- Tứ phân vị thứ nhất Q_1

Nhóm chứa Q_1 là nhóm thứ 3: $[7,0; 7,5)$. Khi đó $Q_1 = 7,0 + \frac{\frac{30}{4} - (2 + 5)}{10} \cdot 0,5 = 7,025$.

- Tứ phân vị thứ ba Q_3

Nhóm chứa Q_3 là nhóm thứ 4: $[7,5; 8,0)$.

$$Q_3 = 7,5 + \frac{\frac{3 \cdot 30}{4} - (2 + 5 + 10)}{9} \cdot 0,5 = 7,80(5) \approx 7,806$$

Khi đó $\Delta_Q = Q_3 - Q_1 = 7,806 - 7,025 = 0,781$.

- Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là $\Delta_Q = Q_3 - Q_1 = 7,806 - 7,025 = 0,781$.

Câu 14: Khảo sát thời gian xem điện thoại trong một ngày của một số học sinh khối 12 thu được mẫu số liệu ghép nhóm sau

Thời gian (phút)	[0; 20)	[20; 40)	[40; 60)	[60; 80)	[80; 100)
Số học sinh	4	8	12	10	8

Các khẳng định sau đây đúng hay sai?

- Tổng số học sinh được khảo sát là 42.
- Mốt của mẫu số liệu lớn hơn 54.
- Khoảng tứ phân vị của mẫu số liệu lớn hơn 38.
- Phương sai của mẫu số liệu nhỏ hơn 610.

Lời giải

a) Đúng	b) Sai	c) Đúng	d) Đúng
---------	--------	---------	---------

a) Tổng số học sinh được khảo sát là $n = 4 + 8 + 12 + 10 + 8 = 42$.

b) Nhóm có tần số lớn nhất là $[40; 60)$.

Mốt của mẫu số liệu là

$$M_0 = 40 + \frac{12 - 8}{(12 - 8) + (12 - 10)} \cdot (60 - 40) \approx 53,3.$$

c)

Gọi $x_1, x_2, \dots, x_{42}$ là thời gian xem điện thoại trong ngày của 42 học sinh khối 12 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 $x_1, x_2, \dots, x_{21}$ nên $Q_1 = x_{11}$. Do đó Q_1 thuộc nhóm $[20; 40)$.

Tứ phân vị thứ ba Q_3 là trung vị của dãy $x_{22}, x_{23}, \dots, x_{42}$ nên $Q_3 = x_{32}$. Do đó Q_3 thuộc nhóm $[60; 80)$.

$$Q_1 = 20 + \frac{\frac{42}{4} - 4}{8} \cdot (40 - 20) = 36,25$$

Suy ra

$$Q_3 = 60 + \frac{\frac{3 \cdot 42}{4} - 24}{10} \cdot (80 - 60) = 75$$

Khoảng tứ phân vị của mẫu số liệu là $\Delta Q = Q_3 - Q_1 = 75 - 36,25 = 38,75$.

d) Số trung bình của mẫu số liệu là

$$\bar{x} = \frac{4 \cdot 10 + 8 \cdot 30 + 12 \cdot 50 + 10 \cdot 70 + 8 \cdot 90}{42} \approx 54,76.$$

Phương sai của mẫu số liệu là

$$s^2 = \frac{4 \cdot (10 - 54,76)^2 + 8 \cdot (30 - 54,76)^2 + 12 \cdot (50 - 54,76)^2 + 10 \cdot (70 - 54,76)^2 + 8 \cdot (90 - 54,76)^2}{42} \approx 605,9.$$

Câu 15: Một trang trại phân 1000 quả trứng thành 5 loại, tùy theo khối lượng (đã được làm tròn) của chúng được thống kê bởi bảng dưới đây:

Khối lượng (gam)	$[30; 36)$	$[36; 42)$	$[42; 48)$	$[48; 54)$	$[54; 60)$
Số trứng	45	190	500	250	15

a) Tần suất của khối lượng trứng $[30; 36)$ là 19%.

b) Số trung vị của mẫu số liệu là 43.

c) Khoảng biến thiên của mẫu số liệu 39,18.

d) Độ lệch chuẩn của mẫu số liệu là $\frac{6\sqrt{17}}{5}$.

Lời giải

a) Đúng	b) Sai	c) Sai	d) Đúng
---------	--------	--------	---------

a) Tần suất của khối lượng trứng $[30; 36)$ là $\frac{190}{1000} \cdot 100 = 19\%$.

b) Nhóm chứa trung vị là nhóm $[42; 48)$.

$$M_e = 42 + \frac{\frac{1000}{2} - 235}{500} \cdot (48 - 42) = \frac{2259}{50}.$$

c) Khoảng biến thiên của mẫu số liệu là: $60 - 30 = 30$.

d) Ta có bảng sau:

Khối lượng (gam)	$[30; 36)$	$[36; 42)$	$[42; 48)$	$[48; 54)$	$[54; 60)$
Số trứng	45	190	500	250	15
Giá trị đại diện	33	39	45	51	57

Phương sai là:

$$s^2 = \frac{33^2 \cdot 45 + 39^2 \cdot 190 + 45^2 \cdot 500 + 51^2 \cdot 250 + 57^2 \cdot 15}{1000} - 45^2 = 24,48.$$

Vậy độ lệch chuẩn của mẫu số liệu là

$$s = \sqrt{24,48} = \frac{6\sqrt{17}}{5}.$$

Câu 16: 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

(Nguồn: Tổng cục Thống kê)

Các mệnh đề sau đúng hay sai?

- 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

Lời giải

a) Sai	b) Đúng	c) Đúng	d) 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: $\Delta_Q = Q_3 - Q_1 = 39,64$

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:

$y_1; \dots; y_{18} \in [160; 190); y_{19}; y_{20} \in [190; 220); y_{21}; \dots; y_{27} \in [220; 250); y_{28}; \dots; y_{34} \in [250; 280); y_{35}; \dots; y_{50} \in [280; 310)$

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:

Số trung bình:

$$\bar{x}_X = \frac{1.145 + 1.175 + 1.205 + 8.235 + 7.265 + 2.295}{20} = 242,5$$

Độ lệch chuẩn:

$$\sigma_X = \sqrt{\frac{1.145^2 + 1.175^2 + 1.205^2 + 8.235^2 + 7.265^2 + 2.295^2}{20} - 242,5^2} \approx 35,34$$

Xét số liệu của Quy Nhơn:

Số trung bình:

$$\bar{x}_Y = \frac{1.175 + 2.205 + 4.235 + 10.265 + 3.295}{20} = 253$$

Độ lệch chuẩn:

$$\sigma_Y = \sqrt{\frac{1.175^2 + 2.205^2 + 4.235^2 + 10.265^2 + 3.295^2}{20} - 253^2} \approx 30,59$$

Vậy nếu so sánh theo độ lệch chuẩn thì số giờ nắng trong tháng 6 của Quy Nhơn đồng đều hơn

Câu 17: Điểm kiểm tra giữa học kì I môn Toán của 2 nhóm học sinh được cho trong 2 bảng sau:

⊕ Nhóm học sinh 1

Điểm	Tần số
[4;5)	1
[5;6)	2
[6;7)	4
[7;8)	6
[8;9)	3
[9;10]	2
	$n=18$

Nhóm học sinh 2

Điểm	Tần số
[4;5)	1
[5;6)	1
[6;7)	3
[7;8)	9
[8;9)	3
[9;10]	1
	$n=18$

Chọn Đúng hoặc Sai trong mỗi ý sau:

- a,** Khoảng biến thiên của hai mẫu số liệu đã cho ở hai bảng trên là 6.
b, Khoảng tứ phân vị của mẫu số liệu đã cho ở bảng nhóm học sinh 1 xấp xỉ là 5,8.
c, Điểm kiểm tra trung bình của hai nhóm học sinh trên là xấp xỉ bằng nhau.
d, Điểm kiểm tra của nhóm học sinh 1 đồng đều hơn của nhóm học sinh 2.

Lời giải

a, Đúng

Khoảng biến thiên của hai mẫu số liệu đã cho ở 2 bảng trên là $R = 10 - 4 = 6$ cm.

b, Sai

Xét bảng ghép lớp của nhóm học sinh 1

Số phần tử của mẫu số liệu là $n = 18$

Ta có nhóm 3 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng $\frac{n}{4} = 4,5$

Xét nhóm 3 là nhóm $[6;7)$ có $s = 6, h = 1, n_3 = 4$ và nhóm 2 là nhóm $[5;6)$ có tần số tích lũy $cf_2 = 3$

Áp dụng công thức, ta có tứ phân vị thứ nhất là $Q_1 = 6 + \frac{4,5 - 3}{4} \cdot 1 = \frac{35}{8} = 4,375 \approx 4,4$

Ta có nhóm 5 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng $\frac{3n}{4} = 13,5$

Xét nhóm 5 là nhóm $[8;9)$ có $t = 8, l = 1, n_5 = 3$ và nhóm 4 là nhóm $[7;8)$ có tần số tích lũy $cf_3 = 13$

Áp dụng công thức, ta có tứ phân vị thứ ba là $Q_3 = 8 + \frac{13,5 - 13}{3} \cdot 1 = \frac{49}{6} \approx 8,2$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm đã cho ở bảng nhóm 1 là

$$\Delta Q = Q_3 - Q_1 \approx 8,2 - 4,4 = 3,8$$

c, Đúng

Điểm kiểm tra trung bình của nhóm học sinh 1 là

$$\bar{x}_1 = \frac{1.4,5 + 2.5,5 + 4.6,5 + 6.7,5 + 3.8,5 + 2.9,5}{18} = \frac{131}{18} \approx 7,3$$

Điểm kiểm tra trung bình của nhóm học sinh 2 là

$$\bar{x}_2 = \frac{1.4,5 + 1.5,5 + 3.6,5 + 9.7,5 + 3.8,5 + 1.9,5}{18} = \frac{22}{3} \approx 7,3$$

Vậy điểm kiểm tra trung bình của hai nhóm học sinh trên là xấp xỉ bằng nhau

d, Sai

Phương sai của mẫu số liệu của nhóm học sinh 1 là

$$s_1^2 = \frac{1 \cdot (4,5 - 7,3)^2 + 2 \cdot (5,5 - 7,3)^2 + 4 \cdot (6,5 - 7,3)^2 + 6 \cdot (7,5 - 7,3)^2 + 3 \cdot (8,5 - 7,3)^2 + 2 \cdot (9,5 - 7,3)^2}{18}$$

$$= \frac{389}{225} \approx 1,73$$

Phương sai của mẫu số liệu của nhóm học sinh 2 là

$$s_2^2 = \frac{1 \cdot (4,5 - 7,3)^2 + 1 \cdot (5,5 - 7,3)^2 + 3 \cdot (6,5 - 7,3)^2 + 9 \cdot (7,5 - 7,3)^2 + 3 \cdot (8,5 - 7,3)^2 + 1 \cdot (9,5 - 7,3)^2}{18}$$

$$= \frac{563}{450} \approx 1,25$$

Vì $s_1^2 > s_2^2$ nên điểm kiểm tra của nhóm học sinh 2 đồng đều hơn của nhóm học sinh 1.

Câu 18: 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

Các mệnh đề sau đúng hay sai?

- a) Cỡ mẫu là $n = 20$
b) Khoảng biến thiên của mẫu số liệu trên là $R = 357,9$ mm
c) Tứ phân vị thứ nhất của mẫu số liệu là $Q_1 = 254,9$
d) Khoảng tứ phân vị của mẫu số liệu trên là $\Delta_Q = 162,35$

Lời giải

- a) (ĐÚNG) Cỡ mẫu là $n = 20$.
b) (SAI) Khoảng biến thiên của mẫu số liệu trên là: $R = 522,9 - 147 = 375,9$ (mm).
c) (ĐÚNG) Tứ phân vị thứ nhất là trung vị của mẫu số liệu:
147; 187,1; 200,7; 242,2; 251,4; 258,4 ; 288,5; 298,1; 305 ; 332.

$$Q_1 = \frac{251,4 + 258,4}{2} = 254,9$$

Do đó,

- d) (ĐÚNG) Tứ phân vị thứ ba là trung vị của mẫu số liệu:
341,4; 388,6 ; 400; 413,5; 413,5 ; 421; 432,2; 475; 520; 522,9.

$$Q_3 = \frac{413,5 + 421}{2} = 417,25$$

Do đó,

Khoảng tứ phân vị của mẫu số liệu đã cho là:

$$\Delta_Q = Q_3 - Q_1 = 417,25 - 254,9 = 162,35$$

Câu 19: Bảng dưới đây cho biết chiều cao tổ 1 lớp 12A, 12B

Chiều cao	[145;150)	[150;155)	[155;160)	[160;165)	[165;170)	[170;175)
Số học sinh 12A	3	2	1	4	2	1
Số học sinh 12B	2	3	1	4	0	0

Các mệnh đề sau đúng hay sai

- a) Khoảng biến thiên của mẫu số liệu tổ 1 lớp 12A là 30
b) Khoảng biến thiên của mẫu số liệu tổ 1 lớp 12B là 25
c) Tứ phân vị thứ nhất của mẫu số liệu tổ 1 lớp 12A là 100
d) Tứ phân vị thứ nhất của mẫu số liệu tổ 1 lớp 12B là 151

Lời giải

- a) Khoảng biến thiên của mẫu số liệu tổ 1 lớp 12A là $175 - 145 = 30$

Mệnh đề đúng.

b) Khoảng biến thiên của mẫu số liệu tổ 1 lớp 12B là $165 - 145 = 20$

Mệnh đề Sai.

c) Cỡ mẫu là: 13. Nên tứ phân vị thứ nhất của mẫu số liệu gốc là x_4 nên nhóm chứa tứ phân vị thứ nhất

là nhóm $[150; 155)$ và ta có $Q_1 = 150 + \frac{(\frac{13}{4} - 3)}{2} \cdot 5 \approx 150,6$

Mệnh đề Sai.

d) Cỡ mẫu là: 10. Nên tứ phân vị thứ nhất của mẫu số liệu gốc là x_3 nên nhóm chứa tứ phân vị thứ nhất là nhóm $[150; 155)$ và ta có

$Q_1 = 150 + \frac{(\frac{10}{4} - 2)}{3} \cdot 5 \approx 150,8$

Mệnh đề Sai.

Câu 20: Ta có các bảng thống kê sau:

Nhóm	Giá trị đại diện	Tần số
[6,22 ; 6,46)	6,34	3
[6,46 ; 6,70)	6,58	7
[6,70 ; 6,94)	6,82	5
[6,94 ; 7,18)	7,06	20
[7,18 ; 7,42)	7,30	5
		$n = 40$

Bảng 15

Nhóm	Giá trị đại diện	Tần số
[6,22 ; 6,46)	6,34	2
[6,46 ; 6,70)	6,58	5
[6,70 ; 6,94)	6,82	8
[6,94 ; 7,18)	7,06	19
[7,18 ; 7,42)	7,30	6
		$n = 40$

Bảng 16

Các mệnh đề sau đúng hay sai?

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: $\approx 6,92(m)$.

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	b) Đúng	c) Sai	d) Sai
---------	---------	--------	--------

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 là:

$$\bar{x}_D = \frac{3 \cdot 6,34 + 7 \cdot 6,58 + 5 \cdot 6,82 + 20 \cdot 7,06 + 5 \cdot 7,30}{40} = \frac{276,88}{40} \approx 6,92(m).$$

Vậy phương sai 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:

$$s_D^2 = \frac{1}{40} \left[3 \cdot (6,34 - 6,92)^2 + 7 \cdot (6,58 - 6,92)^2 + 5 \cdot (6,82 - 6,92)^2 + 20 \cdot (7,06 - 6,92)^2 + 5 \cdot (7,30 - 6,92)^2 \right] = \frac{2,9824}{40} \approx 0,07.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm trên là: $s_D \approx \sqrt{0,07} \approx 0,26(m)$.

Số trung bình cộng của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Huy cho bởi Bảng 16 là:

$$\bar{x}_H = \frac{2 \cdot 6,34 + 5 \cdot 6,58 + 8 \cdot 6,82 + 19 \cdot 7,06 + 6 \cdot 7,30}{40} = \frac{278,08}{40} \approx 6,95(m).$$

Vậy 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:

$$\begin{aligned} s_H^2 &= \frac{1}{40} \left[2 \cdot (6,34 - 6,95)^2 + 5 \cdot (6,58 - 6,95)^2 \right. \\ &\quad + 8 \cdot (6,82 - 6,95)^2 + 19 \cdot (7,06 - 6,95)^2 \\ &\quad \left. + 6 \cdot (7,30 - 6,95)^2 \right] = \frac{2,5288}{40} \approx 0,06. \end{aligned}$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm trên là: $s_H \approx \sqrt{0,06} \approx 0,24(m)$

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 21: Bảng 1, Bảng 2 lần lượt biểu diễn mẫu số liệu ghép nhóm thống kê mức lương của hai công ty A, B (đơn vị: triệu đồng)

Nhóm	Giá trị đại diện	Tần số
[10; 15)	12,5	15
[15; 20)	17,5	18
[20; 25)	22,5	10
[25; 30)	27,5	10
[30; 35)	32,5	5
[35; 40)	37,5	2
		$n = 60$

Bảng 1

Nhóm	Giá trị đại diện	Tần số
[10; 15)	12,5	25
[15; 20)	17,5	15
[20; 25)	22,5	7
[25; 30)	27,5	5
[30; 35)	32,5	5
[35; 40)	37,5	3
		$n = 60$

Bảng 2

Các mệnh đề sau đúng hay sai?

- 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	b) Sai	c) Sai	d) Đúng
---------	--------	--------	---------

Số trung bình cộng của mẫu số liệu ghép nhóm của công ty A là:

$$\bar{x}_A = \frac{15 \cdot 12,5 + 18 \cdot 17,5 + 10 \cdot 22,5 + 10 \cdot 27,5 + 5 \cdot 32,5 + 2 \cdot 37,5}{60} = \frac{62}{3}$$

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: $s_A = \sqrt{s_A^2} = \sqrt{49,14} \approx 7$

Số trung bình cộng của mẫu số liệu ghép nhóm của công ty B là:

$$\bar{x}_B = \frac{25 \cdot 12,5 + 15 \cdot 17,5 + 7 \cdot 22,5 + 5 \cdot 27,5 + 5 \cdot 32,5 + 3 \cdot 37,5}{60} = \frac{229}{12}$$

Phương sai của mẫu số liệu ghép nhóm của công ty B là:

$$s_B^2 = \frac{25 \cdot \left(12,5 - \frac{229}{12}\right)^2 + 15 \cdot \left(17,5 - \frac{229}{12}\right)^2 + 7 \cdot \left(22,5 - \frac{229}{12}\right)^2 + 5 \cdot \left(27,5 - \frac{229}{12}\right)^2 + 5 \cdot \left(32,5 - \frac{229}{12}\right)^2 + 3 \cdot \left(37,5 - \frac{229}{12}\right)^2}{60} \approx 57,91$$

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 22: Bảng 24 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

Bảng 24

(Nguồn: Niên giám Thống kê 2021, NXB Thống kê, 2022)

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

Các mệnh đề sau đúng hay sai?

- 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

Lời giải

a) Đúng	b) Đúng	c) Đúng	d) Sai
---------	---------	---------	--------

- Xét số liệu ở Đà Lạt:

+ Khoảng biến thiên: $R = 91,5 - 75 = 16,5$

+ Số phần tử của mẫu là $n = 12$

Tần số tích lũy của các nhóm lần lượt là $cf_1 = 0, cf_2 = 2, cf_3 = 3, cf_4 = 9, cf_5 = 12$

$$\frac{n}{4} = \frac{12}{4} = 3$$

Ta có: 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 $cf_2 = 2$

$$Q_1 = s + \left(\frac{3 - cf_2}{n_3} \right) \cdot h = 81,6 + \left(\frac{3 - 2}{1} \right) \cdot 3,3 = 84,9$$

Ta có tứ phân vị thứ nhất là:

$$\frac{3n}{4} = \frac{3 \cdot 12}{4} = 9$$

Ta có: 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 $cf_3 = 3$

$$Q_3 = t + \left(\frac{9 - cf_3}{n_4} \right) \cdot l = 84,9 + \left(\frac{9 - 3}{6} \right) \cdot 3,3 = 88,2$$

Ta có tứ phân vị thứ ba là:

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:

$$s_1^2 = \frac{0(76,65 - 86)^2 + 2(79,95 - 86)^2 + (83,25 - 86)^2 + 6(86,55 - 86)^2 + 3(89,95 - 86)^2}{12} = 10,7825$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s_1 = \sqrt{s_1^2} = \sqrt{10,7825} \approx 3,28$

- Xét số liệu ở Vũng Tàu:

+ Khoảng biến thiên: $R = 91,5 - 75 = 16,5$

+ Số phần tử của mẫu là $n = 12$

Tần số tích lũy của các nhóm lần lượt là $cf_1 = 5, cf_2 = 11, cf_3 = 12, cf_4 = 12, cf_5 = 12$

$$\frac{n}{4} = \frac{12}{4} = 3$$

Ta có: 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 $s = 75, h = 3,3, n_1 = 5$

$$Q_1 = s + \left(\frac{3 - cf_0}{n_1} \right) \cdot h = 75 + \left(\frac{3 - 0}{5} \right) \cdot 3,3 = 76,98$$

Ta có tứ phân vị thứ nhất là:

Ta có: $\frac{3n}{4} = \frac{3.12}{4} = 9$ 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 $cf_1 = 5$

$$Q_3 = t + \left(\frac{9 - cf_1}{n_2} \right) \cdot l = 75 + \left(\frac{9 - 5}{6} \right) \cdot 3,3 = 77,2$$

Ta có tứ phân vị thứ ba là:

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: $\bar{x}_2 = \frac{5.76,65 + 6.79,95 + 83,25}{12} = 78,85$

Phương sai của mẫu số liệu ghép nhóm là:

$$s_2^2 = \frac{5(76,65 - 78,85)^2 + 6(79,95 - 78,85)^2 + (83,25 - 78,85)^2}{12} = 4,235$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s_2 = \sqrt{s_2^2} = \sqrt{4,235} \approx 2,06$

Vùng Tàu có nhiệt độ không khí trung bình tháng đồng đều hơn vì độ lệch chuẩn nhỏ hơn

Câu 23: Cho mẫu số liệu thống kê cân nặng của 20 củ khoai tây ở một trang trại X dưới đây (Đơn vị: g)

50	70	86	70	80	95	100	60	64	74
53	67	78	82	98	63	56	73	85	63

Các mệnh đề sau đúng hay sai

- Khoảng biến thiên của mẫu số liệu gốc là 50
- Khoảng biến thiên của mẫu số liệu ghép nhóm có độ dài bằng 10 là 50
- Tứ phân vị thứ nhất mẫu số liệu ghép nhóm có độ dài bằng 10 là 64
- Tứ phân vị thứ ba mẫu số liệu ghép nhóm có độ dài bằng 10 là 120

Lời giải

- Khoảng biến thiên của mẫu số liệu gốc là $100 - 50 = 50$

Mệnh đề đúng.

- Ta có bảng số liệu ghép nhóm có độ dài bằng 10 như sau:

Cân nặng	[50;60)	[60;70)	[70;80)	[80;90)	[90;100)
Số củ khoai tây	3	5	5	4	3

Khoảng biến thiên của mẫu số liệu ghép nhóm có độ dài bằng 10 là $100 - 50 = 50$

Mệnh đề đúng.

- Ta có bảng số liệu ghép nhóm có độ dài bằng 10 như sau:

Cân nặng	[50;60)	[60;70)	[70;80)	[80;90)	[90;100)
Số củ khoai tây	3	5	5	4	3

Cỡ mẫu là: 20. Nên tứ phân vị thứ nhất của mẫu số liệu gốc là x_5 nên nhóm chứa tứ phân vị thứ nhất là

nhóm $[60;70)$ và ta có $Q_1 = 60 + \frac{\left(\frac{20}{4} - 3\right)}{5} \cdot 10 = 64$

Mệnh đề đúng.

- Ta có bảng số liệu ghép nhóm có độ dài bằng 10 như sau:

Cân nặng	[50;60)	[60;70)	[70;80)	[80;90)	[90;100)
Số củ khoai tây	3	5	5	4	3

Cỡ mẫu là: 20. Nên tứ phân vị thứ ba của mẫu số liệu gốc là x_{15} nên nhóm chứa tứ phân vị thứ nhất là

$$Q_1 = 80 + \frac{(3 \cdot \frac{20}{4} - 5)}{3} \cdot 10 \approx 113,3$$

nhóm $[80; 90)$ và ta có

Mệnh đề Sai.

Câu 24: Bảng thống kê thời gian (đơn vị: phút) tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình và bạn An

Thời gian (phút)	$[15; 20)$	$[20; 25)$	$[25; 30)$	$[30; 35)$	$[35; 40)$
Số ngày tập của bạn Bình	5	10	10	2	1
Số ngày tập của bạn An	5	5	15	3	0

Các mệnh đề sau đúng hay sai

- Khoảng biến thiên của mẫu số liệu về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn An là 20.
- Khoảng tứ phân vị của mẫu số liệu về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình là 28.
- Khoảng tứ phân vị của mẫu số liệu về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn An là 22.
- Dựa vào khoảng tứ phân vị của hai mẫu số liệu trên thì thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình phân tán hơn bạn An.

Lời giải

- Khoảng biến thiên của mẫu số liệu về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn An là $35 - 15 = 20$

Mệnh đề đúng.

- Cỡ mẫu là: 28. Gọi $x_1, x_2, \dots, x_{28}$ là mẫu số liệu gốc về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình đã được sắp xếp theo thứ tự không giảm. Nên tứ phân vị thứ nhất của

mẫu số liệu gốc là $\frac{1}{2}(x_7 + x_8)$ nên nhóm chứa tứ phân vị thứ nhất là nhóm $[20; 25)$ và ta có

$$Q_1 = 20 + \frac{\left(\frac{1 \cdot 28}{4} - 5\right)}{10} \cdot 5 = 21$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{21} + x_{22})$ nên nhóm chứa tứ phân vị thứ ba là nhóm $[25; 30)$

$$Q_3 = 25 + \frac{\left(\frac{3 \cdot 28}{4} - 15\right)}{10} \cdot 5 = 28$$

và ta có

$$\Delta_Q = Q_3 - Q_1 = 28 - 21 = 7$$

Mệnh đề Sai.

- Cỡ mẫu là: 28. Gọi $x_1, x_2, \dots, x_{28}$ thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn An và giả sử dãy số liệu gốc này đã được sắp xếp theo thứ tự không giảm. Tứ phân vị thứ nhất của

mẫu số liệu gốc là $\frac{1}{2}(x_7 + x_8)$ nên nhóm chứa tứ phân vị thứ nhất là nhóm $[20; 25)$ và ta có

$$Q_1 = 20 + \frac{\left(\frac{1 \cdot 28}{4} - 5\right)}{5} \cdot 5 = 22$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{21} + x_{22})$ nên nhóm chứa tứ phân vị thứ ba là nhóm [25;30)

$$Q_3 = 25 + \frac{\left(\frac{3.28}{4} - 10\right)}{15} \cdot 5 = 26$$

và ta có

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = Q_3 - Q_1 = 26 - 22 = 4$

Mệnh đề Sai.

d) Do $4 < 7$ nên thời gian tập thể dục mỗi buổi sáng trong tháng 2 năm 2023 của bạn Bình phân tán hơn bạn An.

Mệnh đề đúng.

Câu 25: 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

Các mệnh đề sau đúng hay sai

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

b) Khoảng biến thiên của tứ phân vị là 8,95

c) Có một gia đình sử dụng $3m^3$ nước trong một tháng, đây là giá trị ngoại lệ của mẫu số liệu ghép nhóm.

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 thì công ty nên gửi thông báo tiết kiệm nước đến các hộ gia đình có lượng nước tiêu thụ từ $8,95m^3$ nước trở lên.

Lời giải

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

Mệnh đề đúng.

b) Cỡ mẫu là: 160. Gọi $x_1, x_2, \dots, x_{160}$ là mẫu số liệu gốc đã được sắp xếp theo thứ tự không giảm. Nên tứ

phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_{40} + x_{41})$ nên nhóm chứa tứ phân vị thứ nhất là nhóm [6;9)

$$Q_1 = 6 + \frac{\left(\frac{2.160}{4} - 24\right)}{57} \cdot 3 = 8,95$$

và ta có

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{120} + x_{121})$ nên nhóm chứa tứ phân vị thứ ba là nhóm [9;12)

$$Q_3 = 9 + \frac{\left(\frac{3.160}{4} - (24 + 57)\right)}{42} \cdot 3 = 11,79$$

và ta có

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = Q_3 - Q_1 = 11,79 - 8,95 = 2,84$

Mệnh đề Sai.

c) Ta có $3 \in [3;6)$ và $Q_1 - 1,5\Delta_Q = 8,95 - 1,5 \cdot 2,84 = 4,69 > 3$ nên đây là giá trị ngoại lệ

Mệnh đề Đúng.

d) Do $Q_1 = 8,95$ và $Q_3 = 11,79$ nên để tiết kiệm 25% hộ gia đình sử dụng lượng nước nhiều nhất thì phải gửi thông báo cho các gia đình có lượng nước từ $11,79m^3$ trở lên

Mệnh đề Sai.

Câu 26: Khoảng tứ phân vị của mẫu số liệu ghép nhóm là:

- a) Hiệu giữa tứ phân vị thứ ba và tứ phân vị thứ nhất của mẫu số liệu ghép nhóm đó.
 b) Tổng giữa tứ phân vị thứ ba và tứ phân vị thứ nhất của mẫu số liệu ghép nhóm đó.
 c) Hiệu giữa hai tứ phân vị bất kì của mẫu số liệu ghép nhóm đó.
 d) Tổng giữa hai tứ phân vị bất kì của mẫu số liệu ghép nhóm đó.

Lời giải

- a) Mệnh đề đúng.
 b) Mệnh đề sai.
 c) Mệnh đề sai.
 d) Mệnh đề sai.

Câu 27: Bảng tần số ghép nhóm dưới đây thống kê thời gian truy cập internet của các học sinh lớp 12A, 12B như sau:

Thời gian (phút)	[9,5;12,5)	[12,5;15,5)	[15,5;18,5)	[18,5;21,5)	[21,5;24,5)
12A	3	12	15	24	2
12B	13	14	10	10	9

- a) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian truy cập internet của lớp 12A là 15 (phút).
 b) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian truy cập internet của lớp 12B là 12 (phút).
 c) Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian truy cập internet của lớp 12A là $4,75$ (phút).
 d) Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian truy cập internet của lớp 12B là $5,25$ (phút).

Lời giải

- a) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian truy cập internet của học sinh lớp 12A là $R = 24,5 - 9,5 = 15$ (phút).

Mệnh đề đúng.

- b) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian truy cập internet của học sinh lớp 12B là $R = 24,5 - 9,5 = 15$ (phút).

Mệnh đề sai.

c)

Lớp 12A

$$n = 56 \Rightarrow \frac{n}{4} = \frac{56}{4} = 14$$

Ta có:

Tứ phân vị thứ nhất của mẫu số liệu gốc là x_{14}, x_{15}

Do vậy nhóm chứa tứ phân vị thứ nhất là $[12,5;15,5)$.

$$Q_1 = 12,5 + \left[\frac{14-3}{12} \right] \cdot 3 = 15,25$$

Ta có:

$$n = 56 \Rightarrow \frac{3n}{4} = \frac{3 \cdot 56}{4} = 42$$

Ta có:

Tứ phân vị thứ ba của mẫu số liệu gốc là x_{42}, x_{43}

Do vậy nhóm chứa tứ phân vị thứ ba là $[18,5;21,5)$.

$$Q_3 = 18,5 + \left[\frac{42-30}{24} \right] \cdot 3 = 20$$

Ta có:

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 20 - 15,25 = 4,75$.

Mệnh đề đúng.

d)

Lớp 12B

$$n = 56 \Rightarrow \frac{n}{4} = \frac{56}{4} = 14$$

Ta có:

Tứ phân vị thứ nhất của mẫu số liệu gốc là x_{14}, x_{15}

Do vậy nhóm chứa tứ phân vị thứ nhất là $[12,5; 15,5)$.

$$Q_1 = 12,5 + \left[\frac{14 - 13}{14} \right] \cdot 3 = 12,71$$

Ta có:

$$n = 56 \Rightarrow \frac{3n}{4} = \frac{3 \cdot 56}{4} = 42$$

Ta có:

Tứ phân vị thứ ba của mẫu số liệu gốc là x_{42}, x_{43}

Do vậy nhóm chứa tứ phân vị thứ ba là $[18,5; 21,5)$.

$$Q_3 = 18,5 + \left[\frac{42 - 37}{10} \right] \cdot 3 = 20$$

Ta có:

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 20 - 12,71 = 7,29$.

Mệnh đề sai.

Câu 28: 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:

Bảng 1

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

Bảng 2

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

a) Nhiệt độ trung bình cao nhất trong ngày trong tháng Sáu năm 2021 cao hơn nhiệt độ trung bình cao nhất trong ngày trong tháng Sáu năm 2022.

b) Khoảng biến thiên của hai mẫu số liệu trên là 12

c) Khoảng tứ phân vị của mẫu số liệu ở bảng 1 là 3,375

d) Tháng Sáu năm 2021 ở Hà Nội nhiệt độ cao nhất trong ngày phân tán nhiều hơn tháng Sáu năm 2022.

Lời giải

Đáp án: Đ, Đ, S, Đ

Số trung bình ở bảng 1 là $\bar{x}_1 = \frac{0 \cdot 29 + 2 \cdot 31 + 8 \cdot 33 + 5 \cdot 35 + 6 \cdot 37 + 9 \cdot 39}{30} = 35,8$

Số trung bình ở bảng 2 là $\bar{x}_2 = \frac{2 \cdot 29 + 3 \cdot 31 + 4 \cdot 33 + 11 \cdot 35 + 8 \cdot 37 + 2 \cdot 39}{30} = 34,73$

Số trung bình ở bảng 1 là

Khoảng biến thiên của hai mẫu số liệu trên là 12.

Xét bảng 1:

$$Q_1 = 32 + \frac{\frac{30}{4} - 2}{8} \cdot 2 = \frac{267}{8} = 33,375$$

$$Q_2 = M_c = 34 + \frac{\frac{30}{2} - 10}{5} \cdot 2 = 36$$

$$Q_3 = 38 + \frac{\frac{30,3}{4} - 21}{9} \cdot 2 = 38,3$$

Khoảng tứ phân vị của mẫu số liệu ở bảng 1 là $\Delta_Q = Q_3 - Q_1 = 38,3 - 33,375 = 4,925$

Xét bảng 2:

$$Q_1 = 32 + \frac{\frac{30}{4} - 5}{4} \cdot 2 = 33,25$$

$$Q_2 = M_c = 34 + \frac{\frac{30}{2} - 9}{11} \cdot 2 = 35,1$$

$$Q_3 = 36 + \frac{\frac{30,3}{4} - 20}{8} \cdot 2 = 36,625$$

Khoảng tứ phân vị của mẫu số liệu ở bảng 2 là $\Delta_Q = Q_3 - Q_1 = 36,625 - 33,25 = 3,375$

Ta có $4,925 > 3,375$ nên tháng Sáu năm 2021 ở Hà Nội nhiệt độ cao nhất trong ngày phân tán nhiều hơn tháng Sáu năm 2022.

Câu 29: Bảng tần số ghép nhóm dưới đây thống kê số giờ ngủ buổi tối của các học sinh lớp 12A1 và 12A2:

Thời gian	[4;5)	[5;6)	[6;7)	[7;8)	[8;9)
Số học sinh nam	6	10	13	9	7
Số học sinh nữ	4	8	10	11	8

- Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian ngủ của các bạn nam là 5
- Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian ngủ của các bạn nam là 2,09
- Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian ngủ của các bạn nữ trong khoảng (2;3)
- Học sinh nam có thời gian ngủ đồng đều hơn.

Lời giải

- Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian ngủ của các bạn nam là $9 - 4 = 5$

Mệnh đề đúng.

- Tính khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian ngủ của các bạn nam:

Cỡ mẫu $n = 6 + 10 + 13 + 9 + 7 = 45$

Gọi $x_1; x_2; \dots; x_{45}$ là thời gian ngủ của 45 học sinh nam được sắp xếp theo thứ tự tăng dần.

Tứ phân vị thứ nhất của mẫu số liệu là $\frac{x_{11} + x_{12}}{2}$ thuộc nhóm $[5;6)$ nên nhóm chứa tứ phân vị thứ nhất là $[5;6)$. Ta có:

$$Q_1 = 5 + \frac{6-5}{10} \left(\frac{45}{4} - 6 \right) \approx 5,53$$

Tứ phân vị thứ ba của mẫu số liệu là $\frac{x_{34} + x_{35}}{2}$ thuộc nhóm $[7;8)$ nên nhóm chứa tứ phân vị thứ nhất là $[7;8)$. Ta có:

$$Q_3 = 7 + \frac{8-7}{9} \left(\frac{3.45}{4} - 29 \right) \approx 7,53$$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta Q = Q_3 - Q_1 = 7,53 - 5,53 = 2$

Mệnh đề sai.

c) Tính khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian ngủ của các bạn nữ:

Cỡ mẫu $n = 4 + 8 + 10 + 11 + 8 = 41$

Gọi $x_1, x_2, \dots, x_{41}$ là thời gian ngủ của 41 học sinh nữ được sắp xếp theo thứ tự tăng dần.

Tứ phân vị thứ nhất của mẫu số liệu là $\frac{x_{10} + x_{11}}{2}$ thuộc nhóm $[5;6)$ nên nhóm chứa tứ phân vị thứ nhất là $[5;6)$. Ta có:

$$Q_1 = 5 + \frac{6-5}{8} \left(\frac{41}{4} - 4 \right) \approx 5,78$$

Tứ phân vị thứ ba của mẫu số liệu là $\frac{x_{31} + x_{32}}{2}$ thuộc nhóm $[7;8)$ nên nhóm chứa tứ phân vị thứ nhất là $[7;8)$. Ta có:

$$Q_3 = 7 + \frac{8-7}{11} \left(\frac{3.41}{4} - 22 \right) \approx 7,80$$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta Q = Q_3 - Q_1 = 7,80 - 5,78 = 2,02$

Mệnh đề đúng.

d) Vì khoảng tứ phân vị của mẫu số liệu ghép nhóm của học sinh nữ lớn hơn so với học sinh nam. Học sinh nữ có thời gian ngủ đồng đều hơn.

Mệnh đề sai.

Câu 30: Khảo sát tuổi thọ của một loại bóng đèn được hai phân xưởng A và B cùng sản xuất cho ở bảng sau:

Tuổi thọ (tháng)	$[24;27)$	$[27;30)$	$[30;33)$	$[33;36)$	$[36;39)$
Phân xưởng A	4	8	10	6	2
Phân xưởng B	5	7	9	7	2

a) Giá trị đại diện tuổi thọ bóng đèn của nhóm $[27;30)$ là 28 tháng.

b) Khi phương sai của mẫu số liệu giảm 4 lần thì độ lệch chuẩn giảm 16 lần.

c) Tuổi thọ trung bình của bóng đèn mà hai phân xưởng sản xuất là bằng nhau.

d) Khi tuổi thọ trung bình của bóng đèn ở hai phân xưởng không quá chênh lệch, nếu độ lệch chuẩn càng nhỏ thì mẫu số liệu càng phân tán.

Lời giải

a) Giá trị đại diện của nhóm $[27;30)$ là $x_2 = \frac{27+30}{2} = 28,5$.

Mệnh đề sai.

b) Gọi s^2, s là phương sai và độ lệch chuẩn ban đầu.

s'^2, s' là phương sai và độ lệch chuẩn sau khi mẫu số liệu thay đổi.

Theo bài, $s'^2 = \frac{s^2}{4}$

$$s' = \sqrt{s'^2} = \sqrt{\frac{s^2}{4}} = \frac{s}{2}$$

Khi đó, độ lệch chuẩn sau khi mẫu số liệu thay đổi là $\frac{s}{2}$.
 Vậy khi phương sai giảm 4 lần thì độ lệch chuẩn của mẫu số liệu giảm 2 lần.

Mệnh đề sai.

c) Ta có bảng sau:

Giá trị đại diện	25,5	28,5	31,5	34,5	37,5
Phân xưởng A	4	8	10	6	2
Phân xưởng B	5	7	9	7	2

Số trung bình của phân xưởng A là $\bar{x}_A = \frac{25,5 \cdot 4 + 28,5 \cdot 8 + 31,5 \cdot 10 + 34,5 \cdot 6 + 37,5 \cdot 2}{30} = 30,9$

Số trung bình của phân xưởng B là $\bar{x}_B = \frac{25,5 \cdot 5 + 28,5 \cdot 7 + 31,5 \cdot 9 + 34,5 \cdot 7 + 37,5 \cdot 2}{5 + 7 + 9 + 7 + 2} = 30,9$

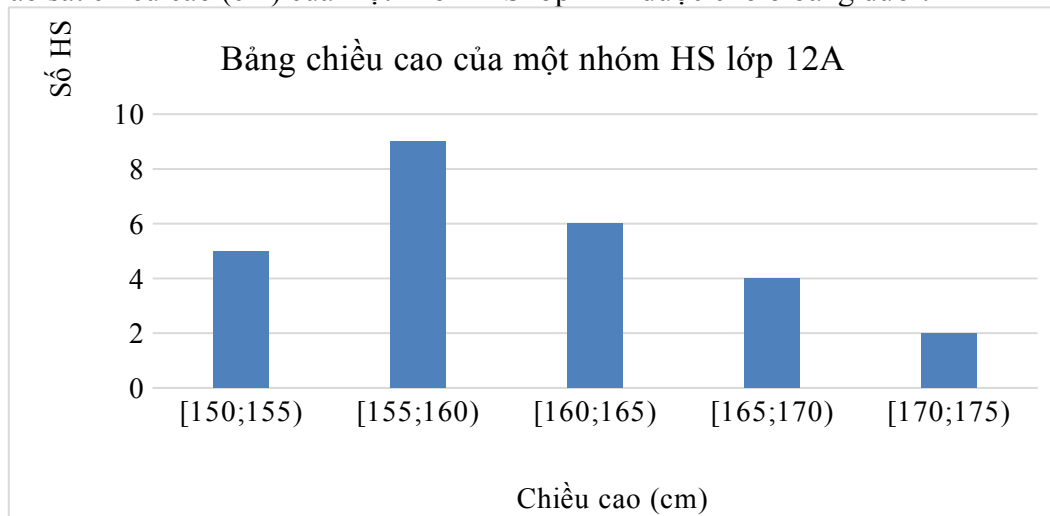
Tuổi thọ trung bình của bóng đèn mà hai phân xưởng sản xuất là bằng nhau.

Mệnh đề đúng.

d) Khi tuổi thọ trung bình của bóng đèn ở hai phân xưởng không quá chênh lệch, nếu độ lệch chuẩn càng lớn thì mẫu số liệu càng phân tán.

Mệnh đề sai.

Câu 31: Khảo sát chiều cao (cm) của một nhóm HS lớp 12A được cho ở bảng dưới:



a) Giá trị đại diện của nhóm chiều cao $[165;170)$ là 167,5 cm.

b) Nếu độ lệch chuẩn của mẫu số liệu tăng 4 lần thì phương sai của mẫu số liệu trên tăng 2 lần.

c) Sau một năm, khảo sát lại chiều cao của nhóm HS trên nhận được bảng số liệu chiều cao ghép nhóm không đổi. Khi đó, chiều cao trung bình của mẫu số liệu gốc tăng lên, nhưng chiều cao trung bình của mẫu số liệu ghép nhóm không thay đổi.

d) Độ lệch chuẩn của mẫu số liệu không phụ thuộc vào chiều cao trung bình của nhóm HS trên là

$$\bar{x} = \frac{2085}{12}$$

Lời giải

a) Giá trị đại diện của nhóm chiều cao $[165;170)$ là $\bar{x}_4 = \frac{165 + 170}{2} = 167,5$.

Mệnh đề đúng.

b) Gọi s^2, s là phương sai và độ lệch chuẩn ban đầu.

s'^2, s' là phương sai và độ lệch chuẩn sau khi mẫu số liệu thay đổi.

Theo bài, $s' = 4s$.

$$s'^2 = (s')^2 = (4s)^2 = 16s^2$$

Khi đó, phương sai sau khi mẫu số liệu thay đổi là

Vậy độ lệch chuẩn của mẫu số liệu tăng 4 lần thì phương sai của mẫu số liệu trên tăng 16 lần.

Mệnh đề sai.

c) Sau một năm chiều cao của từng HS tăng lên nên chiều cao trung bình của mẫu số liệu gốc tăng lên. Sau một năm, bảng số liệu chiều cao ghép nhóm không đổi nên chiều cao trung bình của mẫu số liệu ghép nhóm không đổi.

Mệnh đề đúng.

d) Xử lý số liệu:

Giá trị đại diện	152,5	157,5	162,5	167,5	172,5
Số HS (Tần số)	5	9	6	4	2

Giá trị chiều cao trung bình của mẫu số liệu ghép nhóm trên là:

$$\bar{x} = \frac{152,5 \cdot 5 + 157,5 \cdot 9 + 162,5 \cdot 6 + 167,5 \cdot 4 + 172,5 \cdot 2}{5 + 9 + 6 + 4 + 2} = \frac{2085}{13}$$

Mệnh đề sai.

Câu 32: Cho mẫu số liệu ghép nhóm có tứ phân vị thứ nhất, thứ hai, thứ ba lần lượt là Q_1, Q_2, Q_3 .

a) Khoảng tứ phân vị của mẫu số liệu ghép nhóm đó bằng $Q_3 + Q_1 - Q_2$.

b) Khoảng biến thiên của mẫu số liệu ghép nhóm luôn lớn hơn hoặc bằng khoảng biến thiên của mẫu số liệu gốc.

c) Khoảng biến thiên được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm. Khoảng biến thiên càng nhỏ thì mẫu số liệu càng phân tán.

d) Khoảng tứ phân vị 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.

Lời giải

a) Khoảng tứ phân vị của mẫu số liệu ghép nhóm bằng $Q_3 - Q_1$

Mệnh đề sai.

b)

Mệnh đề đúng.

c) 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.

Mệnh đề sai.

d) Khoảng tứ phân vị 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.

Mệnh đề đúng.

Câu 33: Bác tài xế A và bác tài xế B thống kê lại độ dài quãng đường (đơn vị: km) mà hai 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 bác tài A lái xe	5	10	9	4	2
Số ngày bác tài B lái xe	4	8	12	6	0

a) Khoảng biến thiên về độ dài quãng đường đi mỗi ngày của bác tài A và B ở mẫu số liệu trên bằng nhau.

- b) Tứ phân vị thứ nhất của mẫu số liệu về độ dài quãng đường mỗi ngày của bác tài A lớn hơn bác tài B
- c) Tứ phân vị thứ ba của mẫu số liệu về quãng đường mỗi ngày của bác tài B thuộc nhóm $[150; 200)$.
- d) Theo khoảng biến thiên thì độ dài quãng đường mỗi ngày của bác tài A phân tán hơn độ dài quãng đường mỗi ngày bác tài. **B**

Lời giải

a) Khoảng biến thiên về độ dài quãng đường đi của bác tài A là $300 - 50 = 250$.

Khoảng biến thiên về độ dài quãng đường đi của bác tài B là $250 - 50 = 200$.

Mệnh đề sai.

b) Tứ phân vị thứ nhất của mẫu số liệu về độ dài quãng đường mỗi ngày của bác tài A là

$$Q_1 = 100 + \frac{\frac{30}{4} - 5}{10} \cdot 50 = 112,5$$

Tứ phân vị thứ nhất của mẫu số liệu về độ dài quãng đường mỗi ngày của bác tài B là

$$Q_1 = 100 + \frac{\frac{30}{4} - 4}{8} \cdot 50 = 121,875$$

Mệnh đề sai.

c) Tứ phân vị thứ ba của mẫu số liệu về quãng đường mỗi ngày của bác tài B là x_{23} nên thuộc nhóm $[150; 200)$

Mệnh đề đúng.

d) Vì khoảng biến thiên về độ dài quãng đường đi của bác tài A lớn hơn bác tài B nên mức độ phân tán về độ dài quãng đường bác A nhiều hơn bác. **B**.

Mệnh đề đúng.

Câu 34: 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

a) Cỡ mẫu $n = 20$.

b) Khoảng biến thiên của mẫu số liệu trên là $375,9(mm)$.

c) Nếu chia mẫu số liệu trên thành 4 nhóm với nhóm đầu tiên là $[140; 240)$ thì nhóm thứ tư là $[440; 522,9)$

d) Khoảng biến thiên của mẫu số liệu ghép nhóm bằng mẫu số liệu đã cho.

Lời giải

a) **Đúng.** Cỡ mẫu $n = 20$.

b) **Đúng.** Sắp xếp lại mẫu số liệu trên theo thứ tự không giảm, ta được:

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 trên là: $R = 522,9 - 147 = 375,9(mm)$.

c) **Sai.** Nếu chọn nhóm đầu tiên là $[140; 240)$, thì 3 nhóm còn lại là $[240; 340)$, $[340; 440)$, $[440; 540)$.

d) **Sai.** Khoảng biến thiên của mẫu số liệu ghép nhóm là

$$R' = 540 - 140 = 400(mm)$$

Do đó khoảng biến thiên của mẫu số liệu ghép nhóm lớn hơn mẫu số liệu đã cho.

Câu 35: Một nông trường nghiên cứu sự phát triển của cây keo 3 năm tuổi tại hai khu vườn A và B . Ở khu vườn A kết quả đo chiều cao của các cây keo 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

Ở khu vườn B từ mẫu số liệu về chiều cao của cây keo 3 năm tuổi người ta tính được khoảng tứ phân vị là 0,5. Các mệnh đề sau đúng hay sai?

- a) Cỡ của mẫu số liệu trong bảng trên là 100.
b) Tứ phân vị thứ nhất (làm tròn đến hàng phần trăm) của mẫu số liệu ghép nhóm trong bảng trên là: $Q_1 = 8,87$.
c) Khoảng tứ phân vị (làm tròn đến hàng phần trăm) của mẫu số liệu ghép nhóm trong bảng trên là: $\Delta_Q = 0,30$.
d) Chiều cao của các cây vườn A đồng đều hơn vườn B .

Lời giải

a) Cỡ của mẫu số liệu ghép nhóm: $n = 5 + 12 + 25 + 44 + 14 = 100$, suy ra mệnh đề **đúng**.

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

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_{25} + x_{26})$ nên nhóm chứa tứ phân vị thứ nhất là nhóm $[8, 8; 9, 0)$

Ta có $Q_1 = 8,8 + \frac{\frac{100}{4} - (5+12)}{25} \cdot (9 - 8,8) \approx 8,86$, suy ra mệnh đề **sai**.

c) Tứ vị phân thứ ba của mẫu số liệu là $\frac{1}{2}(x_{75} + x_{76})$ nên nhóm chứa tứ phân vị thứ ba là nhóm $[9, 0; 9, 2)$

Ta có: $Q_3 = 9 + \frac{\frac{3 \cdot 100}{4} - (5+12+25)}{44} \cdot (9,2 - 9) = 9,15$

Khoảng tứ phân vị của mẫu số liệu là $\Delta_Q = Q_3 - Q_1 = 9,15 - 8,86 = 0,29$, suy ra mệnh đề **sai**.

Do $\Delta_Q \approx 0,29 < 0,5$ nên chiều cao của các cây ở khu vườn A đồng đều hơn ở khu vườn B , suy ra mệnh đề **đúng**.

Câu 36: Điểm thi môn Toán tốt nghiệp năm 2024 của học sinh trường THPT A và THPT B ta thu được như sau:

Điểm	[0; 2)	[2; 4)	[4; 6)	[6; 8)	[8; 10)
Số HS trường A	60	125	195	80	40
Số HS trường B	43	149	188	90	30

Các mệnh đề sau đây đúng hay sai?

- a) Điểm trung bình môn Toán của học sinh trường THPT A và THPT B bằng nhau.
b) Phương sai của mẫu số liệu ghép nhóm học sinh trường A bằng 5,01.

- c) Độ lệch chuẩn của mẫu số liệu ghép nhóm học sinh trường B bằng $2,53$.
 d) Điểm thi môn toán của học sinh trường THPT B đồng đều hơn học sinh trường THPT A.

Lời giải

a) Điểm trung bình môn Toán của học sinh trường THPT A và THPT B tương ứng là:

$$\bar{x}_A = \frac{1}{500}(1 \cdot 60 + 3 \cdot 125 + 5 \cdot 195 + 7 \cdot 80 + 9 \cdot 40) = 4,66$$

$$\bar{x}_B = \frac{1}{500}(1 \cdot 43 + 3 \cdot 149 + 5 \cdot 188 + 7 \cdot 90 + 9 \cdot 30) = 4,66$$

Vậy mệnh đề **đúng**.

b) Phương sai của mẫu số liệu ghép nhóm học sinh trường A

$$s_A^2 = \frac{1}{500}[60 \cdot 1^2 + 125 \cdot 3^2 + 195 \cdot 5^2 + 80 \cdot 7^2 + 40 \cdot 9^2] - 4,66^2 = 4,7244$$

Vậy mệnh đề **sai**.

c) Phương sai của mẫu số liệu ghép nhóm học sinh trường B

$$s_B^2 = \frac{1}{500}[43 \cdot 1^2 + 149 \cdot 3^2 + 188 \cdot 5^2 + 90 \cdot 7^2 + 30 \cdot 9^2] - 4,66^2 = 4,1324$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm học sinh trường B bằng $s_B = \sqrt{s_B^2} = \sqrt{4,1324} \approx 2,03$.

Vậy mệnh đề **sai**.

d) $s_B^2 < s_A^2$. Điểm thi môn toán của học sinh trường B đồng đều hơn điểm thi môn toán của học sinh trường A.

Vậy mệnh đề **đúng**.

Câu 37: Lợi nhuận của 50 nhà đầu tư vào hai lĩnh vực A và B được cho như sau:

Lợi nhuận (triệu đồng)	[120;122)	[122;124)	[124;126)	[126;128)	[128;130)
Số nhà đầu tư vào lĩnh vực A	8	9	12	10	11
Số nhà đầu tư vào lĩnh vực B	16	4	3	6	21

a) Lợi nhuận trung bình của các nhà đầu tư vào lĩnh vực A là $125,28$ triệu đồng.

b) Lợi nhuận trung bình của các nhà đầu tư vào lĩnh vực A thấp hơn lợi nhuận trung bình của các nhà đầu tư vào lĩnh vực B.

c) Độ lệch chuẩn của lợi nhuận khi đầu tư vào lĩnh vực A là $2,75$.

d) Lợi nhuận trung bình của các nhà đầu tư vào lĩnh vực B cao hơn lợi nhuận trung bình của các nhà đầu tư vào lĩnh vực A thì độ rủi ro khi đầu tư vào lĩnh vực B thấp hơn khi đầu tư vào lĩnh vực A.

Lời giải

a) Chọn giá trị đại diện cho các nhóm số liệu ta có bảng thống kê sau:

Giá trị đại diện	121	123	125	127	129
Số nhà đầu tư vào lĩnh vực A	8	9	12	10	11
Số nhà đầu tư vào lĩnh vực B	16	4	3	6	21

Lợi nhuận trung bình của các nhà đầu tư vào lĩnh vực A là:

$$\bar{x}_A = \frac{8.121 + 9.123 + 12.125 + 10.127 + 11.129}{50} = 125,28$$

Suy ra mệnh đề **đúng**.

b) Lợi nhuận trung bình của các nhà đầu tư vào lĩnh vực B là:

$$\bar{x}_B = \frac{16.121 + 4.123 + 3.125 + 6.127 + 21.129}{50} = 125,48$$

Lợi nhuận trung bình của các nhà đầu tư vào lĩnh vực A thấp hơn lợi nhuận trung bình của các nhà đầu tư vào lĩnh vực B vì $\bar{x}_A < \bar{x}_B$.

Suy ra mệnh đề **đúng**.

c) Độ lệch chuẩn của lợi nhuận khi đầu tư vào lĩnh vực A là

$$S_A = \sqrt{\frac{8.121^2 + 9.123^2 + 12.125^2 + 10.127^2 + 11.129^2}{50} - 125,28^2} \approx 2,74$$

Suy ra mệnh đề **sai**.

d) Độ lệch chuẩn của lợi nhuận khi đầu tư vào lĩnh vực B là

$$S_B = \sqrt{\frac{16.121^2 + 4.123^2 + 3.125^2 + 6.127^2 + 21.129^2}{50} - 125,48^2} \approx 3,52$$

Đầu tư vào lĩnh vực B rủi ro hơn lĩnh vực A do $S_B > S_A$.

Suy ra mệnh đề **sai**.

Câu 38: 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

Các mệnh đề sau đúng hay sai?

a) Xét mẫu số liệu của cổ phiếu B ta có số trung bình của mẫu số liệu ghép nhóm là 115,28.

b) Xét mẫu số liệu của cổ phiếu A ta có phương sai của mẫu số liệu ghép nhóm là 7,5216

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) Xét mẫu số liệu của cổ phiếu B :

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_2 = \frac{16.121 + 4.123 + 3.125 + 6.127 + 21.129}{50} = 125,28.$$

Suy ra mệnh đề Sai.

b) 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

$$S_1^2 = \frac{1}{50} (8.121^2 + 9.123^2 + 12.125^2 + 10.127^2 + 11.129^2) - (125,28)^2 = 7,5216$$

Suy ra mệnh đề Đúng.

c) Ta có Phương sai của mẫu số liệu ghép nhóm là

$$S_2^2 = \frac{1}{50} (16.121^2 + 4.123^2 + 3.125^2 + 6.127^2 + 21.129^2) - (125,48)^2 = 12,4096.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_2 = \sqrt{S_2^2} = \sqrt{12,4096}$.

Suy ra mệnh đề Sai

d) Ta có độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_1 = \sqrt{S_1^2} = \sqrt{7,5216}$.

$S_2 > S_1$, 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 .

Suy ra mệnh đề Đúng.

Câu 39: 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	[120; 122)	[122; 124)	[124; 126)	[126; 128)	[128; 130)
Giá trị đại diện	121	123	125	127	129
Cổ phiếu A	8	9	12	10	11
Cổ phiếu B	16	4	3	6	21

a) 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 $S_1 = \sqrt{S_1^2} = \sqrt{7,5216}$ suy ra mệnh đề **đúng**.

b) Xét mẫu số liệu của cổ phiếu B :

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_2 = \frac{16.121 + 4.123 + 3.125 + 6.127 + 21.129}{50} = 125,28.$$

Suy ra mệnh đề **sai**.

c) 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 $S_2 = \sqrt{S_2^2} = \sqrt{12,4096}$. Suy ra mệnh đề **sai**.

d) Theo kết quả tính toán ở phần c) ta có:

Phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm của cổ phiếu A lần lượt là:

$$S_1^2 = 7,5216, S_1 = \sqrt{S_1^2} = \sqrt{7,5216}$$

Phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm của cổ phiếu B lần lượt là:

$$S_2^2 = 12,4096, S_2 = \sqrt{S_2^2} = \sqrt{12,4096}$$

Dễ thấy $S_1 < S_2$ 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 . Suy ra mệnh đề **đúng**.

Câu 40: Thời gian học Toán ở nhà mỗi ngày của các bạn tổ 1 và tổ 2 lớp 12A được cho trong bảng sau:

Thời gian (phút)	$[0; 30)$	$[30; 60)$	$[60; 90)$	$[90; 120)$
Số học sinh tổ 1	4	1	3	2
Số học sinh tổ 2	2	5	3	0

a) Tổng số học sinh học Toán ở nhà mỗi ngày dưới 30 phút của 2 tổ là 6.

b) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian học Toán ở nhà mỗi ngày của tổ 1 là 60.

c) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian học Toán ở nhà mỗi ngày của tổ 2 không vượt quá 100.

d) Nếu dựa vào khoảng biến thiên thì thời gian học Toán ở nhà mỗi ngày của các bạn tổ 2 phân tán hơn thời gian học Toán ở nhà mỗi ngày của các bạn tổ 1.

Lời giải

a) **Đúng**. Tổng số học sinh học Toán ở nhà mỗi ngày dưới 30 phút là: $4 + 2 = 6$.

b) **Sai**. Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian học Toán ở nhà mỗi ngày của tổ 1 là $120 - 0 = 120$ (phút)

c) **Đúng**. Ta có khoảng $[90; 120)$ của tổ 2 là 0 học sinh nên khoảng đầu tiên chứa dữ liệu là $[0; 30)$ và khoảng cuối cùng chứa dữ liệu là $[60; 90)$.

Do đó, khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian học Toán ở nhà mỗi ngày của tổ 2 là $90 - 0 = 90$ (phút). $90 < 100$

d) **Sai**. Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian học Toán ở nhà mỗi ngày của tổ 2 nhỏ hơn khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian học Toán ở nhà mỗi ngày của tổ 1.

Câu 41: Bạn Minh thường đi xe buýt từ nhà đến trường để học. Dưới đây là bảng thống kê thời gian của 100 lần bạn Minh đi xe buýt từ nhà đến trường.

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

a) Cỡ mẫu của mẫu số liệu trên là $n = 6$.

b) Tứ phân vị thứ nhất của mẫu số liệu là $Q_1 = \frac{693}{38}$.

c) Tứ phân vị thứ ba của mẫu số liệu là $Q_3 = \frac{68}{38}$.

d) Biết rằng trong 100 lần đi trên, chỉ có đúng một lần bạn Minh đi hết 32 phút. Thời gian của lần đi này là giá trị bất thường của mẫu số liệu ghép nhóm.

Lời giải

a) Ta có : cỡ mẫu $n = 100$, suy ra mệnh đề **sai**.

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

Ta có $x_1, x_2, \dots, x_{22} \in [15; 18)$; $x_{23}, \dots, x_{60} \in [18; 21)$; $x_{61}, \dots, x_{87} \in [21; 24)$; $x_{88}, \dots, x_{95} \in [24; 27)$; $x_{96}, \dots, x_{99} \in [27; 30)$; $x_{100} \in [30; 33)$.

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

liệu ghép nhóm là $Q_1 = 18 + \frac{\frac{100}{4} - 22}{38} \cdot (21 - 18) = \frac{693}{38}$, suy ra mệnh đề **đúng**.

c) 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 $Q_3 = 21 + \frac{3 \cdot \frac{100}{4} - (22 + 38)}{27} \cdot (24 - 21) = \frac{68}{3}$, suy ra mệnh đề **sai**.

d) Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = \frac{68}{3} - \frac{693}{38} = \frac{505}{114} \approx 4,33$.

Trong lần duy nhất bạn Minh đi hết 32 phút, thời gian đi của bạn Minh thuộc nhóm $[30; 33)$.

Vì $Q_3 + 1,5\Delta_Q = \frac{6683}{228} \approx 29,33 < 30$ nên thời gian của bạn Minh đi hết 32 phút là giá trị bất thường của mẫu số liệu ghép nhóm trên, suy ra mệnh đề **đúng**.

Câu 42: Cân nặng của các quả sầu riêng được thu hoạch trong vườn ông An thống kê ở bảng sau:

Cân nặng (kg)	[1; 2)	[2; 3)	[3; 4)	[4; 5)	[5; 6)
Số quả sầu riêng	15	23	28	14	10

a) Mẫu số liệu trên có cỡ mẫu $n = 90$.

b) Cân nặng trung bình của sầu riêng ở vườn ông An xấp xỉ $3,29$.

c) Độ lệch chuẩn của mẫu số liệu ghép nhóm trên xấp xỉ $1,47$.

d) Với cùng số lượng thu hoạch, biết cân nặng trung bình của sầu riêng ở vườn ông Bình xấp xỉ $3,28$. Do đó, cân nặng của sầu riêng ở vườn ông Bình ít phân tán hơn cân nặng của sầu riêng ở vườn ông An. (Các kết quả làm tròn đến hàng phần trăm)

Lời giải

a) Ta có: cỡ mẫu $n = 15 + 23 + 28 + 14 + 10 = 90$. Suy ra mệnh đề **đúng**.

b) Cân nặng trung bình của sầu riêng ở vườn ông An là

$$\bar{x} = \frac{15 \cdot 1,5 + 23 \cdot 2,5 + 28 \cdot 3,5 + 14 \cdot 4,5 + 10 \cdot 5,5}{90} \approx 3,29$$

Suy ra mệnh đề **đúng**.

c) Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{90} (15 \cdot 1,5^2 + 23 \cdot 2,5^2 + 28 \cdot 3,5^2 + 14 \cdot 4,5^2 + 10 \cdot 5,5^2) - 3,29^2 \approx 1,47$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là

$$S \approx \sqrt{1,47} \approx 1,21$$

Suy ra mệnh đề **sai**.
d) Để đo mức độ phân tán của mẫu số liệu ghép nhóm ta dựa vào phương sai và độ lệch chuẩn, không phải dựa vào số trung bình cộng. Do đó, mệnh đề **sai**.

Câu 43: Ông Hà đầu tư số tiền bằng nhau vào hai lĩnh vực kinh doanh A, B. Ông Hà thống kê số tiền thu được mỗi tháng trong vòng 50 tháng theo mỗi lĩnh vực cho kết quả như sau:

Số tiền (triệu đồng)	[5;10)	[10;15)	[15;20)	[20;25)	[25;30)
Số tháng đầu tư vào lĩnh vực A	4	8	30	5	3
Số tháng đầu tư vào lĩnh vực B	19	6	2	7	16

a) Số tiền trung bình thu được khi đầu tư vào lĩnh vực A là: $15,6$ triệu đồng

b) Phương sai của số tiền thu được hàng tháng khi đầu tư vào lĩnh vực B là $76,25$

c) Độ lệch chuẩn của số tiền thu được hàng tháng khi đầu tư vào lĩnh vực A là $20,25$

d) Đầu tư vào lĩnh vực B rủi ro cao hơn đầu tư vào lĩnh vực A

Lời giải

Số tiền trung bình thu được khi đầu tư vào lĩnh vực A và B lần lượt là:

$$\bar{x}_A = \frac{4 \cdot 7,5 + 8 \cdot 12,5 + 30 \cdot 17,5 + 5 \cdot 22,5 + 3 \cdot 27,5}{50} = 17 \text{ triệu đồng}$$

$$\bar{x}_B = \frac{19 \cdot 7,5 + 6 \cdot 12,5 + 2 \cdot 17,5 + 7 \cdot 22,5 + 16 \cdot 27,5}{50} = 17 \text{ triệu đồng}$$

Suy ra ý a) sai

Phương sai của số tiền thu được hàng tháng khi đầu tư vào lĩnh vực B là:

$$S_B^2 = \frac{19(7,5-17)^2 + 6(12,5-17)^2 + 2(17,5-17)^2 + 7(22,5-17)^2 + 16(27,5-17)^2}{50} = 76,25$$

Suy ra ý b) đúng

$$S_B^2 = 76,25 \Rightarrow S_B \approx 8,73 \text{ triệu đồng}$$

Phương sai của số tiền thu được hàng tháng khi đầu tư vào lĩnh vực A là:

$$S_A^2 = \frac{4(7,5-17)^2 + 8(12,5-17)^2 + 30(17,5-17)^2 + 5(22,5-17)^2 + 3(27,5-17)^2}{50} = 20,25$$

$$\Rightarrow S_A = 4,5 \text{ triệu đồng}$$

Suy ra ý c) sai

Ta thấy $S_B > S_A$ nên đầu tư vào lĩnh vực B rủi ro cao hơn

Suy ra ý d) đúng

Câu 44: 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

- a) Số phụ nữ tham gia khảo sát ở khu vực A là 100.
b) Khoảng biến thiên của mẫu số liệu ghép nhóm ở khu vực A là 12.
c) Khoảng biến thiên của mẫu số liệu ghép nhóm ở khu vực A là 15.
d) Nếu so sánh theo khoảng tứ phân vị thì phụ nữ ở khu vực B có độ tuổi kết hôn đồng đều hơn khu vực A.

Lời giải

- a) Ta có $n = 10 + 27 + 31 + 25 + 7 = 100$. Suy ra mệnh đề **đúng**.
b) Khoảng biến thiên của mẫu số liệu ghép nhóm ở khu vực A là $R_1 = 34 - 19 = 15$. Suy ra mệnh đề **sai**.
c) Khoảng biến thiên của mẫu số liệu ghép nhóm ở khu vực A là $R_2 = 31 - 19 = 12$. Suy ra mệnh đề **sai**.
d) Gọi $x_1; x_2; \dots; x_{100}$ là mẫu số liệu về độ tuổi kết hôn của một số phụ nữ ở khu vực A được xếp theo thứ tự không giảm.

Ta có tứ phân vị thứ nhất của mẫu số liệu $x_1; x_2; \dots; x_{100}$ là $\frac{1}{2}(x_{25} + x_{26}) \in [22; 25)$. Do đó, tứ phân vị thứ

$$Q_1 = 22 + \frac{\frac{100}{4} - 10}{27} (25 - 22) = \frac{71}{3}.$$

nhất của mẫu số liệu ghép nhóm là

Ta có tứ phân vị thứ ba của mẫu số liệu $x_1; x_2; \dots; x_{100}$ là $\frac{1}{2}(x_{75} + x_{76}) \in [28; 31)$. Do đó, tứ phân vị thứ

$$Q_3 = 28 + \frac{\frac{3 \cdot 100}{4} - (10 + 27 + 31)}{25} (31 - 28) = \frac{721}{25}.$$

ba của mẫu số liệu ghép nhóm là

Khoảng tứ phân vị của mẫu số liệu ghép nhóm về độ tuổi kết hôn của một số phụ nữ ở khu vực A là:

$$\Delta_Q = Q_3 - Q_1 = \frac{721}{25} - \frac{71}{3} \approx 5,17$$

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

Ta có tứ phân vị thứ nhất của mẫu số liệu $y_1; y_2; \dots; y_{100}$ là $\frac{1}{2}(y_{25} + y_{26}) \in [19; 22)$. Do đó, tứ phân vị thứ

$$Q'_1 = 19 + \frac{\frac{100}{4}}{47} (22 - 19) = \frac{968}{47}.$$

nhất của mẫu số liệu ghép nhóm là

Ta có tứ phân vị thứ ba của mẫu số liệu $y_1; y_2; \dots; y_{100}$ là $\frac{1}{2}(y_{75} + y_{76}) \in [22; 25)$. Do đó, tứ phân vị thứ

$$Q'_3 = 22 + \frac{\frac{3 \cdot 100}{4} - 47}{40} (25 - 22) = \frac{241}{10}.$$

ba của mẫu số liệu ghép nhóm là

Khoảng tứ phân vị của mẫu số liệu ghép nhóm về độ tuổi kết hôn của một số phụ nữ ở khu vực B là:

$$\Delta'_Q = Q'_3 - Q'_1 = \frac{241}{10} - \frac{968}{47} \approx 3,5$$

Vì $\Delta'_Q < \Delta_Q$ nên phụ nữ ở khu vực B có độ tuổi kết hôn đồng đều hơn. Suy ra mệnh đề **đúng**.

Câu 45: Số tiền thưởng cuối năm của nhân viên công ty X được thống kê như sau (đơn vị triệu đồng):

3	5	7	12	15	11	4	5	8	15
16	18	20	22	15	29	28	22	24	26
16	21	25	26	29	6	8	10	17	18

a, Hãy chuyển mẫu số liệu sang dạng ghép nhóm với 6 nhóm có độ dài bằng nhau.

b, Các câu sau là Đúng hay Sai? (làm tròn kết quả đến hàng phần mười)

b1, Tiền thưởng trung bình của các nhân viên là 16,2.

b2, Khoảng biến thiên của mẫu số liệu ghép nhóm là 26.

b3, Khoảng tứ phân vị của mẫu số liệu ghép nhóm là 14,5.

Lời giải

FB tác giả: Huong Giang

a, Cỡ mẫu: $n = 30$

+ Sắp xếp mẫu số liệu theo thứ tự tăng dần ta được:

3	4	5	5	6	7	8	8	10	11
12	15	15	15	16	16	17	18	18	20
21	22	22	24	25	26	26	28	28	29

+ Giá trị lớn nhất của mẫu số liệu là 29, và giá trị nhỏ nhất của mẫu số liệu là 3, nên độ dài nhóm là $29 - 3 = 26$

+ Chia thành 6 nhóm có độ dài bằng nhau, ta chọn độ dài mỗi nhóm là 4,5, ta có mẫu số liệu ghép nhóm như bảng sau:

Nhóm	[3; 7,5)	[7,5; 12)	[12; 16,5)	[16,5; 21)	[21; 25,5)	[25,5; 30)
Tần số	6	4	6	4	5	5

$$\bar{X} = \frac{m_1x_1 + m_2x_2 + \dots + m_6x_6}{n} = 16,2$$

b, + Tiền thưởng trung bình là:

+ Khoảng biến thiên: $30 - 3 = 27$.

+ Khoảng tứ phân vị:

Tứ phân vị thứ nhất Q_1 là x_8 thuộc nhóm $[7,5; 12)$ nên nhóm thứ 2 này chứa Q_1 .

Suy ra: $p = 2, a_2 = 7,5; a_3 = 12; m_2 = 4; m_1 = 6$

Ta có:

$$Q_1 = a_2 + \frac{\frac{n}{4} - m_1}{m_2} (a_3 - a_2) = 7,5 + \frac{\frac{30}{4} - 6}{4} (12 - 7,5) = \frac{147}{16} = 9,2$$

Tứ phân vị thứ nhất Q_3 là x_{23} thuộc nhóm $[21; 25,5)$ nên nhóm thứ 5 này chứa Q_3 .

Suy ra: $p = 5, a_5 = 21; a_6 = 25,5; m_5 = 5; m_4 = 4$

Ta có:

$$Q_3 = a_5 + \frac{\frac{3n}{4} - (m_1 + m_2 + m_3 + m_4)}{m_5} (a_6 - a_5) = 21 + \frac{\frac{3 \cdot 30}{4} - 20}{5} \cdot 4,5 = \frac{93}{4} = 23,3$$

Vậy khoảng tứ phân vị là: $\Delta_Q = Q_3 - Q_1 = 23,3 - 9,2 = 14,1$

Kết luận: b1, Đúng b2, Sai b3, Sai

Câu 46: Bảng 1 biểu diễn mẫu số liệu ghép nhóm về nhiệt độ ($^{\circ}\text{C}$) của tỉnh Nghệ An tháng 5 năm 2024

Nhóm	Giá trị đại diện	Tần số	Tần số tích lũy
[29;31)	30	1	1
[31;33)	32	4	5
[33;35)	34	5	10
[35;37)	36	13	26
[37;39]	38	7	33
		$n = 30$	

Trong mỗi ý a),b),c),d) chọn đúng hoặc sai (làm tròn đến hàng phần trăm)

- a) Nhóm $[31;33)$ có tần số bằng: 4
b) Một của mẫu số liệu ghép nhóm đã cho là: 13
c) Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên bằng: 2,92
d) Phương sai của mẫu số liệu ghép nhóm trên bằng: 4,57

Lời giải

a) **Đ.** Nhóm $[31;33)$ có tần số bằng: 4.

b) **S.** Ta có nhóm $[35;37)$ có tần số lớn nhất nên Một của mẫu số liệu trên là

$$M_0 = u + \frac{n_i - n_{i-1}}{2n_i - n_{i-1} - n_{i+1}} \cdot g = 35 + \frac{13 - 5}{2 \cdot 13 - 5 - 7} \cdot 2 = 36,14$$

c) **Đ** Ta có số phần tử của mẫu là $n = 30$.

+ Ta có $\frac{n}{4} = 7,5$ nên nhóm $[33;35)$ là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng $7,5$. Nhóm $[33;35)$ có $s = 33; h = 2; n = 5$ và nhóm 2 là nhóm $[31;33)$ có $cf_1 = 5$. Áp dụng công thức ta có tứ phân vị thứ nhất là

$$Q_1 = 33 + \frac{7,5 - 5}{5} \cdot 2 = 34 \quad (^{\circ}\text{C})$$

+ Ta có $\frac{3n}{4} = 22,5$ nên nhóm $[35;37)$ là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng $22,5$. Xét nhóm 4 là nhóm $[35;37)$ có $t = 35; l = 2; n_4 = 13$ và nhóm 3 có tần số tích lũy $cf_4 = 10$. Áp dụng công thức, ta có tứ phân vị thứ 3 là

$$Q_3 = 35 + \frac{22,5 - 10}{13} \cdot 2 = 36,92 \quad (^{\circ}\text{C})$$

+ Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm đã cho là $\Delta Q = Q_3 - Q_1 = 36,92 - 34 = 2,92$

d) **Đ.** Ta có số trung bình cộng của mẫu số liệu ghép nhóm trên là

$$\bar{x} = \frac{1}{30} (1 \cdot 30 + 32 \cdot 4 + 34 \cdot 5 + 36 \cdot 13 + 38 \cdot 7) = 35,4 \quad (^{\circ}\text{C})$$

Vậy phương sai của mẫu số liệu ghép nhóm trên là

$$s^2 = \frac{1}{30} (1(30 - 35,4)^2 + 4(32 - 35,4)^2 + 5(34 - 35,4)^2 + 13(36 - 35,4)^2 + 7(38 - 35,4)^2) = 4,57$$

Câu 47: Cho bảng phân bố tần số ghép lớp cân nặng (đơn vị: kg) của các công nhân trong một công ty như sau:

Các lớp giá trị của X	[50; 52)	[52; 54)	[54; 56)	[56; 58)	[58; 60)	Cộng
Tần số n_i	15	20	45	15	5	100

Các mệnh đề sau đúng hay sai?

- a) Tần suất của nhóm [52; 54) là: 20% .
b) Số trung vị của mẫu số liệu là: 54,909 .
c) Khoảng biến thiên của mẫu số liệu trên là: 10 .
d) Độ lệch chuẩn của mẫu số liệu trên là: 4,35 .

Lời giải

a) Đúng	b) Sai	c) Đúng	d) Sai
----------------	---------------	----------------	---------------

a) Tần số của nhóm [52; 54) là: 20 .

Tần suất của nhóm [52; 54) là: $\frac{20}{100} \cdot 100\% = 20\%$.

b) Trung vị của mẫu số liệu là $x_3 \in [54; 56)$. Do đó, trung vị của mẫu số liệu ghép nhóm là:

$$M_e = Q_2 = 54 + \frac{\frac{2 \cdot 100}{4} - (15 + 20)}{45} (56 - 54) = \frac{164}{3} = 54,667$$

c) Khoảng biến thiên của mẫu số liệu là $R = 60 - 50 = 10$.

d) Số trung bình cộng của mẫu số liệu ghép nhóm của công ty là:

$$\bar{x} = \frac{51 \cdot 15 + 53 \cdot 20 + 55 \cdot 45 + 57 \cdot 15 + 59 \cdot 5}{100} = 54,5$$

Phương sai của mẫu số liệu ghép nhóm của công ty là:

$$s^2 = \frac{15 \cdot (51 - 54,5)^2 + 20 \cdot (53 - 54,5)^2 + 45 \cdot (55 - 54,5)^2 + 15 \cdot (57 - 54,5)^2 + 5 \cdot (59 - 54,5)^2}{100} = 4,35$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm của công ty là: $s = \sqrt{s^2} = \sqrt{4,35} \approx 2,085665$

Câu 48: Cho bảng số liệu dưới đây về thời gian (phút) tập thể dục buổi sáng của hai bạn Bình và Chi trong 30 ngày.

Thời gian	[15; 20)	[20; 25)	[25; 30)	[30; 35)	[35; 40)
Bạn Bình	5	8	10	4	3
Bạn Chi	10	10	5	3	2

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 của Chi là 25 (phút).
b. Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bạn Bình là:

$$Q_1 = \frac{354}{16}$$

c. Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bạn Chi là $\frac{314}{9}$

d. Phương sai của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bạn Bình là 9

Lời giải

a	b	c	d
Đúng	Sai	Đúng	Đúng

Ta có

Thời gian	[15; 20)	[20; 25)	[25; 30)	[30; 35)	[35; 40)
-----------	----------	----------	----------	----------	----------

Giá trị đại diện	17,5	22,5	27,5	32,5	37,5
Bạn Bình	5	8	10	4	3
Bạn Chi	10	10	5	3	2

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

b. Ở bạn Bình. Ta có cỡ mẫu $n = 30$.

Vì $\frac{n}{4} = \frac{30}{4} = 7,5$ và $5 < 7,5 < 5 + 8$ nên tứ phân vị thứ nhất thuộc nhóm $[20; 25)$ và tứ phân vị thứ nhất là

$$Q_1 = 20 + \frac{\frac{30}{4} - 5}{8} \cdot 5 = \frac{345}{16}$$

c. Ở bạn Chi. Ta có cỡ mẫu $n = 30$.

Vì $\frac{n}{4} = \frac{30}{4} = 7,5$ và $7,5 < 10$ nên tứ phân vị thứ nhất thuộc nhóm $[15; 20)$ và tứ phân vị thứ nhất là

$$Q_1 = 15 + \frac{\frac{30}{4} - 0}{10} \cdot 5 = 18,75$$

Vì $\frac{3n}{4} = \frac{3 \cdot 30}{4} = 22,5$ và $10 + 10 < 22,5 < 10 + 10 + 5$ nên tứ phân vị thứ ba thuộc nhóm $[25; 30)$ và tứ phân

$$Q_3 = 25 + \frac{\frac{3 \cdot 30}{4} - (10 + 10)}{5} \cdot 5 = 27,5$$

vị thứ ba là

Vậy khoảng tứ phân vị là $\Delta_Q = Q_3 - Q_1 = 28,75$

d. Thời gian trung bình bạn Bình tập thể dục buổi sáng là

$$\bar{x} = \frac{17,5 \cdot 5 + 22,5 \cdot 8 + 27,5 \cdot 10 + 32,5 \cdot 4 + 37,5 \cdot 3}{30} = \frac{157}{6} \approx 26,17$$

Vậy phương sai của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bạn Bình là

$$s_B^2 = \frac{5 \left(17,5 - \frac{157}{6} \right)^2 + 8 \left(22,5 - \frac{157}{6} \right)^2 + 10 \left(27,5 - \frac{157}{6} \right)^2 + 4 \left(32,5 - \frac{157}{6} \right)^2 + 3 \left(37,5 - \frac{157}{6} \right)^2}{30} = \frac{314}{9}$$

Các mệnh đề sau đúng hay sai?

a) Dựa vào khoảng biến thiên của mẫu số liệu ghép nhóm thì chiều cao của học sinh lớp 12A phân tán hơn lớp 12B.

b) Dựa vào khoảng tứ phân vị của mẫu số liệu ghép nhóm thì học sinh lớp 12A có chiều cao phân tán hơn học sinh lớp 12B.

c) Dựa vào phương sai của mẫu số liệu ghép nhóm thì chiều cao của học sinh lớp 12A ít phân tán hơn học sinh lớp 12B.

d) Học sinh lớp 12B có chiều cao đồng đều hơn học sinh lớp 12A vì có độ lệch chuẩn nhỏ hơn.

Lời giải

a) Đúng	b) Sai	c) Sai	d) Đúng
---------	--------	--------	---------

a) Với số liệu của học sinh lớp 12A, có khoảng biến thiên: $R_A = 175 - 145 = 30$

Với số liệu của học sinh lớp 12B, có khoảng biến thiên: $R_B = 175 - 150 = 25$

Vậy ta có $R_A > R_B$ nên chiều cao của học sinh lớp 12A phân tán hơn lớp 12B.

b)

Với mẫu số liệu ghép nhóm của học sinh lớp 12A:

Cỡ mẫu là $n = 2 + 1 + 15 + 11 + 9 + 3 = 41$. Gọi $x_1, x_2, x_3, \dots, x_{41}$ là mẫu số liệu gốc gồm chiều cao của học sinh lớp 12A được sắp xếp theo thứ tự không giảm.

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_{10} + x_{11}) \in [155; 160)$ nên nhóm chứa tứ phân vị thứ nhất là

nhóm $[155; 160)$ và $Q_1 = 155 + \frac{\frac{41}{4} - 3}{15} \cdot 5 \approx 157,42$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{31} + x_{32}) \in [165; 170)$ nên nhóm chứa tứ phân vị thứ ba là

nhóm $[165; 170)$ và $Q_3 = 165 + \frac{\frac{3 \cdot 41}{4} - 29}{9} \cdot 5 \approx 165,97$

Suy ra $\Delta Q_A = Q_3 - Q_1 \approx 8,55$

Với mẫu số liệu ghép nhóm của học sinh lớp 12B:

Cỡ mẫu là $n = 1 + 16 + 11 + 10 + 4 = 42$. Gọi $x_1, x_2, x_3, \dots, x_{42}$ là mẫu số liệu gốc gồm chiều cao của học sinh lớp 12B được sắp xếp theo thứ tự không giảm.

Tứ phân vị thứ nhất của mẫu số liệu gốc là $x_{11} \in [155; 160)$ nên nhóm chứa tứ phân vị thứ nhất là nhóm

$[155; 160)$ và $Q_1 = 155 + \frac{\frac{42}{4} - 1}{16} \cdot 5 \approx 157,97$

Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{32} \in [165; 170)$ nên nhóm chứa tứ phân vị thứ ba là nhóm

$[165; 170)$ và $Q_3 = 165 + \frac{\frac{3 \cdot 42}{4} - 28}{10} \cdot 5 = 166,75$

Suy ra $\Delta Q_B = Q_3 - Q_1 \approx 8,78$

Do $\Delta Q_A < \Delta Q_B$ nên học sinh lớp 12A có chiều cao phân tán ít hơn học sinh lớp 12B.

c) Chọn giá trị đại diện cho các nhóm số liệu ta được bảng:

Chiều cao (cm)	$[145; 150)$	$[150; 155)$	$[155; 160)$	$[160; 165)$	$[165; 170)$	$[170; 175)$
Giá trị đại diện	147,5	152,5	157,5	162,5	167,5	172,5
Số học sinh của lớp 12A	2	1	15	11	9	3
Số học sinh của lớp 12B	0	1	16	11	10	4

Chiều cao trung bình của học sinh lớp 12A là:

$$\bar{x}_A = \frac{1}{41} [2 \cdot 147,5 + 1 \cdot 152,5 + 15 \cdot 157,5 + 11 \cdot 162,5 + 9 \cdot 167,5 + 3 \cdot 172,5] \approx 161,52$$

Chiều cao trung bình của học sinh lớp 12B là:

$$\bar{x}_B = \frac{1}{42} [0 \cdot 147,5 + 1 \cdot 152,5 + 16 \cdot 157,5 + 11 \cdot 162,5 + 10 \cdot 167,5 + 4 \cdot 172,5] = 162,5$$

Phương sai của mẫu số liệu lớp 12A là

$$S_A^2 \approx \frac{1}{41} [2 \cdot 147,5^2 + 1 \cdot 152,5^2 + 15 \cdot 157,5^2 + 11 \cdot 162,5^2 + 9 \cdot 167,5^2 + 3 \cdot 172,5^2] - (161,52)^2 \approx 35,83$$

Phương

sai của mẫu số liệu lớp 12B là

$$S_B^2 = \frac{1}{42} [0 \cdot 147,5^2 + 1 \cdot 152,5^2 + 16 \cdot 157,5^2 + 11 \cdot 162,5^2 + 10 \cdot 167,5^2 + 4 \cdot 172,5^2] - (162,50)^2 \approx 27,38$$

Vậy

dựa vào phương sai của mẫu số liệu ghép nhóm thì chiều cao của học sinh lớp 12A phân tán hơn học sinh lớp 12B.

d) Độ lệch chuẩn của mẫu số liệu lớp 12A là $S_A \approx \sqrt{35,83} \approx 5,99$

Độ lệch chuẩn của mẫu số liệu lớp 12B là $S_B \approx \sqrt{27,38} \approx 5,23$

Vậy học sinh lớp 12B có chiều cao đồng đều hơn học sinh lớp 12A vì có độ lệch chuẩn nhỏ hơn.

Câu 50: Mẫu số liệu dưới đây ghi lại tốc độ của 40 ô tô khi đi qua một trạm đo tốc độ (đơn vị: km/h):

48,5	43	50	55	45	60	53	55,5	44	65
51	62,5	41	44,5	57	57	68	49	46,5	53,5
61	49,5	54	62	59	56	47	50	60	61
49,5	52,5	57	47	60	55	45	47,5	48	61,5

a) Bảng tần số ghép nhóm cho mẫu số liệu trên có sáu nhóm ứng với sáu nửa khoảng là

Nhóm	Tần số
[40; 45)	4
[45; 50)	11
[50; 55)	7
[55; 60)	8
[60; 65)	8
[65; 70)	2
	$n = 40$

b) Mẫu số liệu trên có số trung bình là 54,875

c) Tứ phân vị của mẫu số liệu trên là: $Q_1 = 47,8(km/h)$; $Q_2 = 53,6(km/h)$; $Q_3 = 60(km/h)$.

d) Khoảng biến thiên của mẫu số liệu trên là 25.

Lời giải

a) Bảng tần số ghép nhóm:

Nhóm	Tần số
[40; 45)	4
[45; 50)	11
[50; 55)	7
[55; 60)	8
[60; 65)	8
[65; 70)	2
	$n = 40$

Vậy ý a **ĐÚNG**

b) **S** *Số trung bình

$$\bar{x} = \frac{4.42,5 + 11.47,5 + 7.52,5 + 8.57,5 + 8.62,5 + 2.67,5}{40} = 53,875 \text{ (km/h)}$$

Vậy ý b **SAI**

c) *Số phần tử của mẫu là $n = 40 \Rightarrow \frac{n}{2} = 20$.

Nhóm 3 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 20.

Xét nhóm 3 là nhóm [50; 55) có $r = 50, d = 5, n_3 = 7$ và $cf_2 = 15$

Ta có số trung vị của mẫu số liệu là:

$$M_e = 50 + \frac{20 - 15}{7} \cdot 5 \approx 53,6 \text{ (km/h)}$$

*Số phần tử của mẫu là $n = 40$.

$\frac{n}{4} = 10$
Ta có:

Nhóm 2 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 10.

Xét nhóm 2 là nhóm [45; 50) có $r = 45, d = 5, n_2 = 11$ và $cf_1 = 4$

Ta có tứ phân vị thứ nhất là:

$$Q_1 = 45 + \frac{10 - 4}{11} \cdot 5 \approx 47,8 \text{ (km/h)}$$

$\frac{3n}{4} = 30$
Ta có:

Nhóm 4 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 30.

Xét nhóm 4 là nhóm [55; 60) có $r = 55, d = 5, n_4 = 8$ và $cf_3 = 22$

Ta có tứ phân vị thứ ba là:

$$Q_3 = 55 + \frac{30 - 22}{8} \cdot 5 = 60 \text{ (km/h)}$$

Vậy tứ phân vị của mẫu số liệu trên là: $Q_1 = 47,8(\text{km} / \text{h}); Q_2 = 53,6(\text{km} / \text{h}); Q_3 = 60(\text{km} / \text{h})$.

Vậy ý c ĐÚNG

d) Khoảng biến thiên của mẫu số liệu là $R = a_7 - a_1 = 70 - 40 = 30$

Vậy ý d SAI

Câu 51: Bảng dưới đây cho ta bảng tần số ghép nhóm số liệu thống kê cân nặng của 40 học sinh lớp 12B trong một trường trung học phổ thông (đơn vị: kilôgam).

Nhóm	Số học sinh
$[30; 40)$	2
$[40; 50)$	10
$[50; 60)$	16
$[60; 70)$	8
$[70; 80)$	2
$[80; 90)$	2
	$n = 40$

Các mệnh đề sau đúng hay sai?

a) Số học sinh nặng dưới 50kg là 12.

b) Một của mẫu số liệu ghép nhóm trên xấp xỉ bằng $54,29(\text{kg})$.

c) Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là $\frac{39}{2}$.

d) Phương sai của mẫu số liệu ghép nhóm là 128.

Lời giải

a) Số học sinh nặng dưới 50kg là $2 + 10 = 12$. **Vậy ý a) đúng.**

b)

Nhóm	Tần số
$[30; 40)$	2
$[40; 50)$	10
$[50; 60)$	16
$[60; 70)$	8
$[70; 80)$	2
$[80; 90)$	2
	$n = 40$

Nhóm chứa một của mẫu số liệu là $[50; 60)$.

Do đó $u_m = 50; n_m = 16; n_{m-1} = 10, n_{m+1} = 8, u_{m+1} - u_m = 60 - 50 = 10$.

Một của mẫu số liệu ghép nhóm xấp xỉ bằng:

$$M_0 = 50 + \frac{16 - 10}{(16 - 10) + (16 - 8)} \cdot 10 = \frac{380}{7} \approx 54,29(\text{kg}).$$

Một của mẫu số liệu ghép nhóm trên xấp xỉ bằng $54,29(kg)$. Vậy ý b) đúng.

c)

Nhóm	Tần số
$[30; 40)$	2
$[40; 50)$	10
$[50; 60)$	16
$[60; 70)$	8
$[70; 80)$	2
$[80; 90)$	2
	$n = 40$

Cỡ mẫu $n = 40$.

Gọi $x_1, x_2 \in [30; 40)$; $x_3, \dots, x_{12} \in [40; 50)$; $x_{13}, \dots, x_{28} \in [50; 60)$; $x_{29}, \dots, x_{36} \in [60; 70)$; $x_{37}, x_{38} \in [70; 80)$; $x_{39}, x_{40} \in [80; 90)$.

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_{10} + x_{11}) \in [40; 50)$. Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là

$$Q_1 = 40 + \frac{\frac{40}{4} - 2}{10} \cdot (50 - 40) = 48.$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{30} + x_{31}) \in [60; 70)$. Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là

$$Q_3 = 60 + \frac{\frac{3 \cdot 40}{4} - (2 + 10 + 16)}{8} \cdot (70 - 60) = \frac{125}{2}.$$

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

$$\Delta_Q = \frac{125}{2} - 48 = \frac{29}{2}.$$

Vậy ý c) sai.

d)

Ta có bảng cân nặng của các em học sinh theo giá trị đại diện:

Nhóm	Giá trị đại diện	Tần số
$[30; 40)$	35	2
$[40; 50)$	45	10
$[50; 60)$	55	16
$[60; 70)$	65	8
$[70; 80)$	75	2
$[80; 90)$	85	2
		$n = 40$

Cỡ mẫu $n = 2 + 10 + 16 + 8 + 2 + 2 = 40$.

Số trung bình của mẫu số liệu ghép nhóm là

$$\frac{35.2 + 45.10 + 55.16 + 65.8 + 75.2 + 85.2}{40} = \frac{2240}{40} = 56 (kg)$$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{40} (2.35^2 + 10.45^2 + 16.55^2 + 8.65^2 + 2.75^2 + 2.85^2) - 56^2 = 3265 - 3136 = 129.$$

Vậy ý d) sai.

Câu 52: Trong một hội thao, thời gian chạy $200m$ của một nhóm các vận động viên được ghi lại ở bảng sau:

Thời gian (giây)	[21; 21,5)	[21,5; 22)	[22; 22,5)	[22,5; 23)	[23; 23,5)
Số vận động viên	5	10	30	45	30

Dựa vào bảng số liệu trên, em hãy xét tính đúng, sai của các mệnh đề sau:

- Tần suất của nhóm vận động viên chạy trong khoảng thời gian từ 22 giây đến dưới $22,5$ giây bằng 30%.
- Số trung vị của mẫu số liệu (làm tròn đến chữ số thập phân thứ 2) bằng $22,67$.
- Khoảng biến thiên của mẫu số liệu bằng $R = 2$.
- Độ lệch chuẩn của mẫu số liệu (làm tròn đến chữ số thập phân thứ 2) bằng $0,28$.

Lời giải

a) Cỡ mẫu: $n = 5 + 10 + 30 + 45 + 30 = 120$.

Tần suất của nhóm vận động viên chạy trong khoảng thời gian từ 22 giây đến dưới $22,5$ giây bằng:

$$f_3 = \frac{n_3}{n} = \frac{30}{120} = 25\%$$

Vậy mệnh đề a) sai.

b) Gọi $x_1, x_2, \dots, x_{120}$ là thời gian chạy của 120 vận động viên và dãy này là một dãy không giảm.

Khi đó trung vị là $\frac{x_{60} + x_{61}}{2}$. Do $x_{60}, x_{61} \in [22, 5; 23)$ nên nhóm này chứa trung vị. Ta có:

$$M_e = 22,5 + \frac{\frac{120}{2} - (5 + 10 + 30)}{45} \cdot (23 - 22,5) \approx 22,67$$

Vậy mệnh đề b) đúng.

c) Khoảng biến thiên của mẫu số liệu bằng: $R = 23,5 - 21 = 2,5$.

Vậy mệnh đề c) sai.

d) Giá trị trung bình của mẫu số liệu là:

$$\bar{x} = \frac{5.21,25 + 10.21,75 + 30.22,25 + 45.22,75 + 30.23,25}{120} \approx 22,60$$

Độ lệch chuẩn của mẫu số liệu (làm tròn đến chữ số thập phân thứ 2) bằng

$$s = \sqrt{\frac{5(-1,35)^2 + 10(-0,85)^2 + 30(-0,35)^2 + 45(0,15)^2 + 30(0,65)^2}{120}} \approx 0,53$$

Vậy mệnh đề d) sai.