



VÕ CÔNG TRƯỜNG

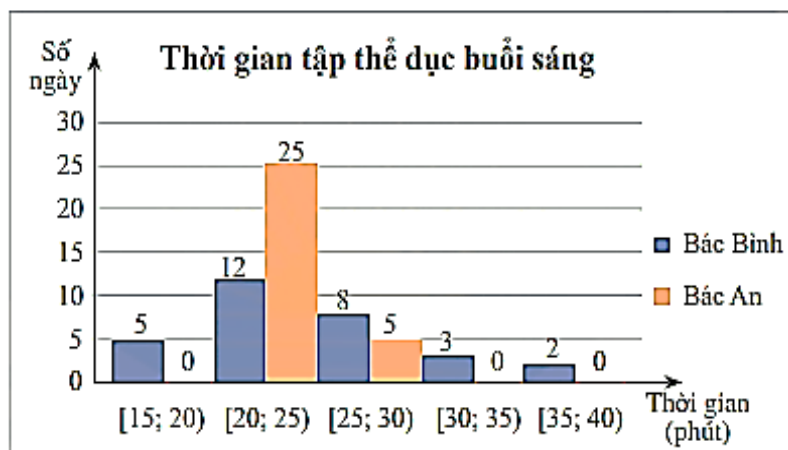
0983 900 570

TOÁN 12

THỐNG KÊ

- ✦ Hệ thống kiến thức chương THỐNG KÊ 10, 11, 12
- ✦ Phương pháp giải các dạng toán thường gặp
- ✦ Ví dụ và lời giải các bài toán cơ bản

Ai là người có thời gian tập thể dục đều hơn?



**MỤC LỤC**

CHƯƠNG ➤ CÁC SỐ ĐẶC TRƯNG ĐO MỨC ĐỘ PHÂN TÁN CỦA MẪU SỐ LIỆU GHÉP NHÓM	2
Bài 1. KHOẢNG BIẾN THIÊN KHOẢNG TỬ PHÂN VỊ CỦA MẪU SỐ LIỆU GHÉP NHÓM	2
Bài 2. PHƯƠNG SAI VÀ ĐỘ LỆCH CHUẨN CỦA MẪU SỐ LIỆU GHÉP NHÓM	6
Phụ lục ➤ THỐNG KÊ LỚP 10 VÀ 11	8
➤ CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CỦA MẪU SỐ LIỆU	8
➤ CÁC SỐ ĐẶC TRƯNG ĐO MỨC ĐỘ PHÂN TÁN CỦA MẪU SỐ LIỆU	9
➤ CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CHO MẪU SỐ LIỆU GHÉP NHÓM	10
✍ BÀI TẬP THAM KHẢO	13
✓ TRẮC NGHIỆM NHIỀU PHƯƠNG ÁN LỰA CHỌN	13
✓ TRẮC NGHIỆM ĐÚNG SAI	37
✓ TRẮC NGHIỆM TRẢ LỜI NGẮN	53
✍ ĐỀ THAM KHẢO KIỂM TRA THƯỜNG XUYÊN	66

CHƯƠNG 4 CÁC SỐ ĐẶC TRƯNG ĐO MỨC ĐỘ PHÂN TÁN CỦA MẪU SỐ LIỆU GHEP NHÓM
Bài 1. KHOẢNG BIẾN THIÊN KHOẢNG TƯ PHÂN VỊ CỦA MẪU SỐ LIỆU GHEP NHÓM
1. Khoảng biến thiên

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

Chú ý:

Xét mẫu số liệu ghép nhóm được cho ở bảng sau:

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

Nếu n_1 và n_{k+1} cùng khác 0 thì $R = u_{k+1} - u_1$.

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

Khoảng biến thiên của mẫu số liệu ghép nhóm là giá trị xấp xỉ khoảng biến thiên của mẫu số liệu gốc và có thể dùng để đo mức độ phân tán của mẫu số liệu.

Khoảng biến thiên $R = u_{k+1} - u_1$ chưa phản ánh được đầy đủ mức độ phân tán của phần lớn các số liệu. Hơn nữa, giá trị của R thường tăng vọt khi xuất hiện giá trị ngoại lệ trong mẫu số liệu. Do đó, để phản ánh mức độ phân tán của số liệu, người ta còn dùng các số đặc trưng khác.

Ví dụ 1: Cô Hà thống kê lại đường kính thân gỗ của một số cây xoan đào 6 năm tuổi được trồng ở một lâm trường ở bảng sau.

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

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

Lời giải

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

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

□ Khoảng biến thiên của mẫu số liệu ghép nhóm là giá trị xấp xỉ khoảng biến thiên của mẫu số liệu gốc và có thể dùng để đo mức độ phân tán của mẫu số liệu.

□ Khoảng biến thiên $R = u_{k+1} - u_1$ chưa phản ánh được đầy đủ mức độ phân tán của phần lớn các số liệu. Hơn nữa, giá trị của R thường tăng vọt khi xuất hiện giá trị ngoại lệ trong mẫu số liệu. Do đó, để phản ánh mức độ phân tán của số liệu, người ta còn dùng các số đặc trưng khác.

2. Khoảng tứ phân vị

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

Tứ phân vị thứ i , kí hiệu là Q_i , với $i = 1, 2, 3$ của mẫu số liệu ghép nhóm (Bảng 1) được xác định như sau:

$$Q_i = u_m + \frac{\frac{i \cdot n}{4} - C}{n_m} (u_{m+1} - u_m)$$

Trong đó:

□ $n = n_1 + n_2 + \dots + n_k$ là cỡ mẫu;

□ $[u_m; u_{m+1})$ là nhóm chứa tứ phân vị thứ i ;

□ n_m là tần số của nhóm chứa tứ phân vị thứ i ;

□ $C = n_1 + n_2 + \dots + n_{m-1}$ (Tổng tần số các nhóm trước nhóm chứa tứ phân vị **gọi là tần số tích lũy**).

Chú ý

Khoảng tứ phân vị của mẫu số liệu ghép nhóm cũng được xác định dựa trên tứ phân vị thứ nhất và tứ phân vị thứ ba như đối với mẫu số liệu không ghép nhóm.

Khoảng tứ phân vị của mẫu số liệu ghép nhóm cho ở Bảng 1, kí hiệu Δ_Q , là hiệu giữa tứ phân vị thứ ba và tứ phân vị thứ nhất: $\Delta_Q = Q_3 - Q_1$.

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

□ Khoảng tứ phân vị của mẫu số liệu ghép nhóm là giá trị xấp xỉ cho khoảng tứ phân vị của mẫu số liệu gốc và có thể dùng để đo mức độ phân tán của nửa giữa của mẫu số liệu (tập hợp gồm 50% số liệu nằm chính giữa mẫu số liệu).

- Khoảng tứ phân vị của mẫu số liệu ghép nhóm càng nhỏ thì dữ liệu càng tập trung xung quanh trung vị.
- Khoảng tứ phân vị được dùng để xác định giá trị ngoại lệ trong mẫu số liệu. Giá trị x trong mẫu số liệu là giá trị ngoại lệ nếu $x > Q_3 + 1,5\Delta_Q$ hoặc $x < Q_1 - 1,5\Delta_Q$.
- Khoảng tứ phân vị của mẫu số liệu ghép nhóm không bị ảnh hưởng nhiều bởi các giá trị ngoại lệ trong mẫu số liệu.

QUY TẮC TÌM CÁC TỨ PHÂN VỊ VÀ KHOẢNG TỨ PHÂN VỊ CỦA MẪU SỐ LIỆU GHEP NHÓM

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

Bước 1: Xác định cỡ mẫu $n = n_1 + n_2 + \dots + n_k$

Bước 2: Xác định các nhóm chứa các tứ phân vị

Sắp xếp lại mẫu số liệu theo thứ tự không giảm, ta được $x_1 \leq x_2 \leq \dots \leq x_n$.

- Trung vị của mẫu, kí hiệu là M_e , là giá trị ở chính giữa dãy $x_1, x_2, \dots, x_n$. Cụ thể:

+ Nếu n lẻ thì trung vị mẫu là $M_e = x_{\frac{n+1}{2}}$ (Số liệu đứng chính giữa)

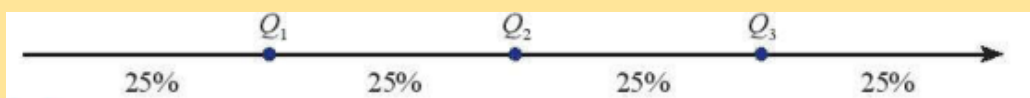
+ Nếu n chẵn thì trung vị mẫu là $M_e = \frac{1}{2} \left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1} \right)$ (Trung bình cộng của 2 số liệu ở giữa)

- Tứ phân vị của một mẫu ngẫu nhiên gồm 3 giá trị, đó là tứ phân vị thứ nhất, thứ hai và thứ ba (lần lượt kí hiệu là Q_1, Q_2, Q_3). Ba giá trị này chia tập hợp dữ liệu đã sắp xếp thành bốn phần đều nhau. Cụ thể:

+ Giá trị tứ phân vị thứ hai, Q_2 , chính là trung vị của mẫu.

+ Giá trị tứ phân vị thứ nhất, Q_1 , là trung vị của nửa số liệu đã sắp xếp bên trái Q_2 (không bao gồm Q_2 nếu n lẻ).

+ Giá trị tứ phân vị thứ ba, Q_3 , là trung vị của nửa số liệu đã sắp xếp bên phải Q_2 (không bao gồm Q_2 nếu n lẻ).



Bước 3: Xác định các đại lượng, yếu tố và thay vào công thức tính tứ phân vị

□ $[u_m; u_{m+1})$ là nhóm chứa tứ phân vị thứ i ;

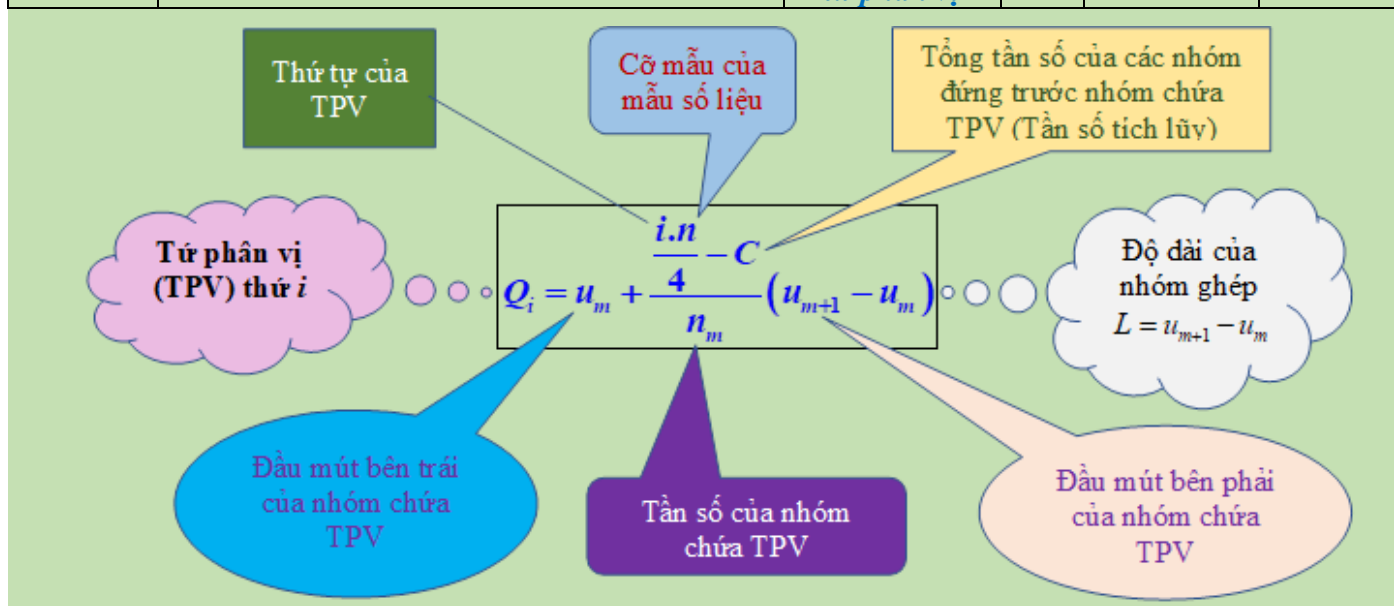
□ n_m là tần số của nhóm chứa tứ phân vị thứ i ;

□ $C = n_1 + n_2 + \dots + n_{m-1}$ (Tổng tần số các nhóm trước nhóm chứa tứ phân vị **gọi là tần số tích lũy**).

$$Q_i = u_m + \frac{\frac{i \cdot n}{4} - C}{n_m} (u_{m+1} - u_m)$$

CÁCH GHI NHỚ VÀ ÁP DỤNG CÔNG THỨC TÍNH TỬ PHÂN VỊ

Nhóm	$[u_1; u_2)$	$[u_2; u_3)$	...	$[u_{m-1}; u_m)$	$[u_m; u_{m+1})$	...	$[u_k; u_{k+1})$	Cỡ mẫu
Tần số	n_1	n_2	...	n_{m-1}	n_m	...	n_k	n
	$C = n_1 + n_2 + \dots + n_{m-1}$				Nhóm chứa tử phân vị			


QUY TẮC TÌM GIÁ TRỊ NGOẠI LỆ

Bước 1: Xác định các tứ phân vị và tính giá trị khoảng tứ phân vị: $\Delta_Q = Q_3 - Q_1$.

Bước 2: Tính $Q_1 - 1,5\Delta_Q$ và $Q_3 + 1,5\Delta_Q$.

Bước 3: Kết luận: các giá trị x (trong mẫu số liệu) nằm ngoài đoạn $\left[Q_1 - \frac{3}{2}\Delta_Q; Q_3 + \frac{3}{2}\Delta_Q \right]$ là **giá trị ngoại lệ**

Vùng chứa giá trị ngoại lệ

x_1	Giá trị ngoại lệ	$Q_1 - 1,5\Delta_Q$		$Q_3 + 1,5\Delta_Q$	Giá trị ngoại lệ	x_n
-------	-------------------------	---------------------	--	---------------------	-------------------------	-------

Ví dụ 2: Hằng ngày ông Thắng đều đi xe buýt từ nhà đến cơ quan. Dưới đây là bảng thống kê thời gian của 100 lần ông Thắng đi xe buýt từ nhà đến cơ quan.

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

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

b) Biết rằng trong 100 lần đi trên, chỉ có đúng một lần ông Thắng đi hết hơn 29 phút. Thời gian của lần đi đó có phải là giá trị ngoại lệ không?

Lời giải

a) Cỡ mẫu $n = 100$.

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

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

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

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

$$Q_1 = 18 + \frac{\frac{100}{4} - 22}{38} \cdot (21 - 18) = \frac{693}{38}.$$

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

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

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

$$\Delta_Q = \frac{68}{3} - \frac{693}{38} = \frac{505}{114}.$$

b) Trong lần duy nhất ông Thắng đi hết hơn 29 phút, thời gian đi của ông thuộc nhóm $[30; 33)$. Vì $Q_3 + 1,5\Delta_Q = \frac{6683}{228} < 30$ nên thời gian của lần ông Thắng đi hết hơn 29 phút là giá trị ngoại lệ của mẫu số liệu ghép nhóm.

Bài 2. PHƯƠNG SAI VÀ ĐỘ LỆCH CHUẨN CỦA MẪU SỐ LIỆU GHEP NHÓM

1. Phương sai và độ lệch chuẩn

Ở lớp 11, ta đã biết giá trị đại diện của nhóm $[a; b)$ là $c = \frac{1}{2}(a + b)$.

Tương tự như cách tính số trung bình, phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm có thể được xác định thông qua giá trị đại diện và tần số của mỗi nhóm.

Xét mẫu số liệu ghép nhóm cho bởi bảng sau:

Nhóm	$[u_1; u_2)$	$[u_2; u_3)$	...	$[u_k; u_{k+1})$
Giá trị đại diện	c_1	c_2	...	c_k
Tần số	n_1	n_2	...	n_k

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

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

Trong đó:

$n = n_1 + n_2 + \dots + n_k$ là cỡ mẫu

$\bar{x} = \frac{1}{n} (n_1 c_1 + n_2 c_2 + \dots + n_k c_k)$ là số trung bình

Độ lệch chuẩn của mẫu số liệu ghép nhóm, kí hiệu S , là căn bậc hai số học của phương sai.

Chú ý

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

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

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

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

2. Ý nghĩa của phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm

(1) Phương sai của mẫu số liệu ghép nhóm là giá trị xấp xỉ cho phương sai của mẫu số liệu gốc. Độ lệch chuẩn của mẫu số liệu ghép nhóm cũng là giá trị xấp xỉ cho độ lệch chuẩn của mẫu số liệu gốc. Chúng được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm. Phương sai và độ lệch chuẩn càng lớn thì dữ liệu càng phân tán.

(2) Độ lệch chuẩn có cùng đơn vị với đơn vị của mẫu số liệu.

Ví dụ 3: Thống kê tổng số giờ nắng trong tháng 9 tại một trạm quan trắc đặt ở Cà Mau trong các năm từ 2002 đến 2021 được thống kê như sau:

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

a) Hãy tính phương sai và độ lệch chuẩn của mẫu số liệu trên.

b) Hãy lập bảng tần số ghép nhóm với nhóm đầu tiên là $[80; 98)$ và độ dài mỗi nhóm bằng 18. Tính phương sai, độ lệch chuẩn của mẫu số liệu ghép nhóm.

c) Hãy tính sai số tương đối của độ lệch chuẩn của mẫu số liệu ghép nhóm so với độ lệch chuẩn của mẫu số liệu gốc.

(Kết quả các phép tính làm tròn đến hàng phần nghìn.)

✎ Lời giải

a) Cỡ mẫu là $n = 20$.

Số trung bình của mẫu số liệu trên là

$$\bar{x}_1 = \frac{111,6 + 134,9 + \dots + 114}{20} = 122,755.$$

Phương sai của mẫu số liệu trên là

$$S_1^2 = \frac{1}{20} (111,6^2 + 134,9^2 + \dots + 114^2) - 122,755^2 \approx 515,453.$$

Độ lệch chuẩn của mẫu số liệu trên là

$$S_1 \approx \sqrt{515,453} \approx 22,704.$$

b) Ta có bảng ghép nhóm sau:

Số giờ nắng	[80; 98)	[98; 116)	[116; 134)	[134; 152)	[152; 170)
Giá trị đại diện	89	107	125	143	161
Số năm	3	6	3	5	3

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_2 = \frac{3.89 + 6.107 + 3.125 + 5.143 + 3.161}{20} = 124,1.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_2^2 = \frac{1}{20} (3.89^2 + 6.107^2 + 3.125^2 + 5.143^2 + 3.161^2) - 124,1^2 = 566,19.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là

$$S_2 = \sqrt{566,19} \approx 23,795.$$

c) Sai số tương đối của độ lệch chuẩn của mẫu số liệu ghép nhóm so với độ lệch chuẩn của mẫu số liệu gốc là

$$\frac{|S_2 - S_1|}{S_1} = \frac{|23,795 - 22,704|}{22,704} \cdot 100\% \approx 4,805\%.$$

Phụ lục ❷ THỐNG KÊ LỚP 10 VÀ 11**❷ CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CỦA MẪU SỐ LIỆU****1. Số trung bình**

- Giả sử ta có một mẫu số liệu là $x_1, x_2, \dots, x_n$.

Số trung bình (hay số trung bình cộng) của mẫu số liệu này, kí hiệu là $\bar{x}$, được tính bởi công thức

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}.$$

- Giả sử mẫu số liệu được cho dưới dạng bảng tần số

Giá trị	x_1	x_2	$\dots$	x_k
Tần số	n_1	n_2	$\dots$	n_k

Khi đó, công thức tính số trung bình trở thành $\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_kx_k}{n}$, trong đó $n = n_1 + n_2 + \dots + n_k$.

Ta gọi n là cỡ mẫu.

Chú ý

Kí hiệu $f_k = \frac{n_k}{n}$ là tần số tương đối (hay còn gọi là tần suất) của x_k trong mẫu số liệu thì số trung bình còn có thể biểu diễn là $\bar{x} = f_1x_1 + f_2x_2 + \dots + f_kx_k$.

Ý nghĩa của số trung bình

Số trung bình của mẫu số liệu được dùng làm đại diện cho các số liệu của mẫu. Nó là một số đo xu thế trung tâm của mẫu đó.

2. Trung vị và tứ phân vị

Sắp xếp lại mẫu số liệu theo thứ tự không giảm, ta được $x_1 \leq x_2 \leq \dots \leq x_n$.

Trung vị của mẫu, kí hiệu là M_e , là giá trị ở chính giữa dãy $x_1, x_2, \dots, x_n$. Cụ thể:

+ Nếu $n = 2k + 1$, $k \in \mathbb{N}$ thì trung vị mẫu là $M_e = x_{k+1}$.

+ Nếu $n = 2k$, $k \in \mathbb{N}$ thì trung vị mẫu là $M_e = \frac{1}{2}(x_k + x_{k+1})$.

Ý nghĩa của trung vị

Trung vị được dùng để đo xu thế trung tâm của mẫu số liệu. Trung vị là giá trị nằm ở chính giữa của mẫu số liệu theo nghĩa: luôn có ít nhất 50% số liệu trong mẫu lớn hơn hoặc bằng trung vị và ít nhất 50% số liệu trong mẫu nhỏ hơn hoặc bằng trung vị. Khi trong mẫu xuất hiện thêm một giá trị rất lớn hoặc rất nhỏ thì số trung bình sẽ bị thay đổi đáng kể nhưng trung vị thì ít thay đổi.

- Tứ phân vị của một mẫu ngẫu nhiên gồm 3 giá trị, đó là tứ phân vị thứ nhất, thứ hai và thứ ba (lần lượt kí hiệu là Q_1, Q_2, Q_3). Ba giá trị này chia tập hợp dữ liệu đã sắp xếp thành bốn phần đều nhau. Cụ thể:

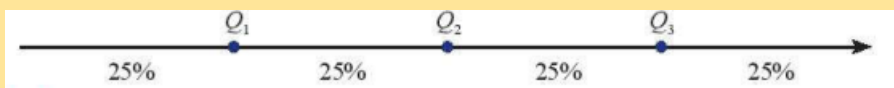
+ Giá trị tứ phân vị thứ hai, Q_2 , chính là trung vị của mẫu.

+ Giá trị tứ phân vị thứ nhất, Q_1 , là trung vị của nửa số liệu đã sắp xếp bên trái Q_2 (không bao gồm Q_2 nếu n lẻ).

+ Giá trị tứ phân vị thứ ba, Q_3 , là trung vị của nửa số liệu đã sắp xếp bên phải Q_2 (không bao gồm Q_2 nếu n lẻ).

Ý nghĩa của tứ phân vị

Các điểm tứ phân vị Q_1, Q_2, Q_3 chia mẫu số liệu đã sắp xếp theo thứ tự từ nhỏ đến lớn thành bốn phần, mỗi phần chứa khoảng 25% tổng số số liệu đã thu thập được. Tứ phân vị thứ nhất Q_1 còn được gọi là tứ phân vị dưới và đại diện cho nửa mẫu số liệu phía dưới. Tứ phân vị thứ ba Q_3 còn được gọi là tứ phân vị trên và đại diện cho nửa mẫu số liệu phía trên.



3. Một

- Cho một mẫu số liệu dưới dạng bảng tần số. Giá trị có tần số lớn nhất được gọi là một của mẫu số liệu, kí hiệu là M_o .

Ý nghĩa của một

Một đặc trưng cho giá trị xuất hiện nhiều nhất trong mẫu. Nếu có tần số xuất hiện bằng nhau thì mẫu số liệu đó không có một.

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

1. Khoảng biến thiên và khoảng tứ phân vị

Sắp xếp mẫu số liệu theo thứ tự không giảm, ta được: $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_n$

Khoảng biến thiên của một mẫu số liệu, kí hiệu là R , là hiệu giữa giá trị lớn nhất và giá trị nhỏ nhất của mẫu số liệu đó, tức là: $R = x_n - x_1$

Khoảng tứ phân vị, kí hiệu là Δ_Q , là hiệu giữa Q_3 và Q_1 , tức là: $\Delta_Q = Q_3 - Q_1$

Ví dụ 4: Hãy tính khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu: 10; 20; 3; 1; 3; 4; 7; 4; 9

Lời giải

Xét mẫu số liệu đã sắp xếp là: 1; 3; 3; 4; 4; 7; 9; 10; 20.

□ Khoảng biến thiên của mẫu số liệu là: $R = 20 - 1 = 19$.

□ Cỡ mẫu là $n = 9$ là số lẻ nên giá trị tứ phân vị thứ hai là: $Q_2 = 4$.

□ Tứ phân vị thứ nhất là trung vị của mẫu: 1; 3; 3; 4. Do đó $Q_1 = 3$.

□ Tứ phân vị thứ ba là trung vị của mẫu: 7; 9; 10; 20. Do đó $Q_3 = 9,5$.

□ Khoảng tứ phân vị của mẫu là: $\Delta_Q = 9,5 - 3 = 6,5$.

Ý nghĩa của khoảng biến thiên và khoảng tứ phân vị

Khoảng biến thiên đặc trưng cho độ phân tán của toàn bộ mẫu số liệu.

Khoảng tứ phân vị đặc trưng cho độ phân tán của một nửa các số liệu, có giá trị thuộc đoạn từ Q_1 đến Q_3 trong mẫu.

Khoảng tứ phân vị không bị ảnh hưởng bởi các giá trị rất lớn hoặc rất bé trong mẫu.

Giá trị ngoại lệ

Khoảng tứ phân vị được dùng để xác định các giá trị ngoại lệ trong mẫu, đó là các giá trị quá nhỏ hay quá lớn so với đa số các giá trị của mẫu.

Cụ thể, phần tử x trong mẫu là giá trị ngoại lệ nếu $x < Q_1 - 1,5 \cdot \Delta_Q$ hoặc $x > Q_3 + 1,5 \cdot \Delta_Q$

Trong Ví dụ 1, $Q_3 + 1,5 \cdot \Delta_Q = 9,5 + 1,5 \cdot 6,5 = 19,25$ và $Q_1 - 1,5 \cdot \Delta_Q = 3 - 1,5 \cdot 6,5 = -6,75$

Nên mẫu có một giá trị ngoại lệ là 20.

Sự xuất hiện của các giá trị ngoại lệ làm cho số trung bình và phạm vi của mẫu thay đổi lớn. Do đó, khi mẫu có giá trị ngoại lệ, người ta thường sử dụng trung vị và khoảng tứ phân vị đo mức độ tập trung và mức độ phân tán của đa số các phần tử trong mẫu số liệu.

2. Phương sai và độ lệch chuẩn

Giả sử ta có một mẫu số liệu là $x_1; x_2; \dots; x_n$

□ **Phương sai** của mẫu số liệu này, kí hiệu là S^2 , được tính bởi công thức:

$$S^2 = \frac{1}{n} \left[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2 \right]$$

trong đó $\bar{x}$ là số trung bình của mẫu số liệu.

□ Căn bậc hai của phương sai được gọi là **độ lệch chuẩn**, kí hiệu là S .

Chú ý

Có thể biến đổi công thức tính phương sai ở trên thành:

$$S^2 = \frac{1}{n} (x_1^2 + x_2^2 + \dots + x_n^2) - \bar{x}^2.$$

Trong thống kê, người ta cũng quan tâm đến phương sai hiệu chỉnh, kí hiệu là $\hat{S}^2$, được tính bởi công thức:

$$\hat{S}^2 = \frac{1}{n-1} \left[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2 \right]$$

Ý nghĩa của phương sai và độ lệch chuẩn

Phương sai là trung bình cộng của các bình phương độ lệch từ mỗi giá trị của mẫu số liệu đến số trung bình.

Phương sai và độ lệch chuẩn được dùng để đo mức độ phân tán của các số liệu trong mẫu quanh số trung bình. Phương sai và độ lệch chuẩn càng lớn thì các giá trị của mẫu càng cách xa nhau (có độ phân tán lớn).

Giả sử mẫu số liệu được cho dưới dạng bảng tần số:

Giá trị	x_1	x_2	$\dots$	x_k
Tần số	n_1	n_2	$\dots$	n_k

Khi đó, công thức tính phương sai trở thành:

$$S^2 = \frac{1}{n} \left[n_1 (x_1 - \bar{x})^2 + n_2 (x_2 - \bar{x})^2 + \dots + n_k (x_k - \bar{x})^2 \right]$$

trong đó $n = n_1 + n_2 + \dots + n_k$

Có thể biến đổi công thức tính phương sai trên thành:

$$S^2 = \frac{1}{n} (n_1 \cdot x_1^2 + n_2 \cdot x_2^2 + \dots + n_k \cdot x_k^2) - \bar{x}^2.$$

☞ CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CHO MẪU SỐ LIỆU GHEP NHÓM

1. Số trung bình và một của mẫu số liệu ghép nhóm

Số liệu ghép nhóm

Mẫu số liệu ghép nhóm thường được trình bày dưới dạng bảng thống kê có dạng như sau:

Bảng 1: Bảng tần số ghép nhóm

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

Chú ý

- Bảng trên gồm k nhóm $[u_j; u_{j+1})$ với $1 \leq j \leq k$, mỗi nhóm gồm một số giá trị được ghép theo một tiêu chí xác định.
- Cỡ mẫu $n = n_1 + n_2 + \dots + n_k$.
- Giá trị chính giữa mỗi nhóm được dùng làm **giá trị đại diện** cho nhóm ấy. Ví dụ nhóm $[u_1; u_2)$ có giá trị đại diện là $\frac{1}{2}(u_1 + u_2)$.
- Hiệu $u_{j+1} - u_j$ được gọi là độ dài của nhóm $[u_j; u_{j+1})$.

Một số quy tắc ghép nhóm của mẫu số liệu

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

Số trung bình

Giả sử mẫu số liệu được cho dưới dạng bảng tần số ghép nhóm:

Nhóm	$[u_1; u_2)$	$[u_2; u_3)$	...	$[u_k; u_{k+1})$
Giá trị đại diện	c_1	c_2	...	c_k
Tần số	n_1	n_2	...	n_k

Số trung bình của mẫu số liệu ghép nhóm, kí hiệu $\bar{x}$, được tính như sau:

$$\bar{x} = \frac{n_1 c_1 + n_2 c_2 + \dots + n_k c_k}{n}$$

trong đó $n = n_1 + n_2 + \dots + n_k$.

Ý nghĩa của số trung bình của mẫu số liệu ghép nhóm

Số trung bình của mẫu số liệu ghép nhóm là giá trị xấp xỉ cho số trung bình của mẫu số liệu gốc. Nó thường dùng để đo xu thế trung tâm của mẫu số liệu.

Mốt

Bảng tần số ghép nhóm

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

Nhóm chứa một của mẫu số liệu ghép nhóm là nhóm có tần số lớn nhất.

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

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

Chú ý

Nếu không có nhóm kề trước của nhóm chứa một thì $n_{m-1} = 0$. Nếu không có nhóm kề sau của nhóm chứa một thì $n_{m+1} = 0$.

Ý nghĩa của mốt của mẫu số liệu ghép nhóm

- Mốt của mẫu số liệu không ghép nhóm là giá trị có khả năng xuất hiện cao nhất khi lấy mẫu. Mốt của mẫu số liệu sau khi ghép nhóm M_o xấp xỉ với mốt của mẫu số liệu không ghép nhóm. Các giá trị nằm xung quanh M_o thường có khả năng xuất hiện cao hơn các giá trị khác.
- Một mẫu số liệu ghép nhóm có thể có nhiều nhóm chứa một và nhiều mốt.

2. Trung vị và tứ phân vị của mẫu số liệu ghép nhóm

Trung vị

Công thức xác định trung vị của mẫu số liệu ghép nhóm:

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

Khi đó:

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

Ý nghĩa của trung vị của mẫu số liệu ghép nhóm

Từ dữ liệu ghép nhóm nói chung không thể xác định chính xác trung vị của mẫu số liệu gốc. Trung vị của mẫu số liệu ghép nhóm là giá trị xấp xỉ cho mẫu số liệu gốc và có thể lấy làm giá trị đại diện cho mẫu số liệu.

Tứ phân vị**Công thức xác định tứ phân vị của mẫu số liệu ghép nhóm**

Tứ phân vị thứ hai của mẫu số liệu ghép nhóm, kí hiệu Q_2 , cũng chính là trung vị của mẫu số liệu ghép nhóm.

Để tìm tứ phân vị thứ nhất của mẫu số liệu ghép nhóm, kí hiệu Q_1 , ta thực hiện như sau:

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

Khi đó:

$$Q_1 = u_m + \frac{\frac{n}{4} - C}{n_m} \cdot (u_{m+1} - u_m)$$

Tương tự, để tìm tứ phân vị thứ ba của mẫu số liệu ghép nhóm, kí hiệu Q_3 , ta thực hiện như sau:

- Giả sử nhóm $[u_j; u_{j+1})$ chứa tứ phân vị thứ ba;
- n_j là tần số của nhóm chứa tứ phân vị thứ ba;
- $C = n_1 + n_2 + \dots + n_{j-1}$.

Khi đó:

$$Q_3 = u_j + \frac{\frac{3n}{4} - C}{n_j} \cdot (u_{j+1} - u_j).$$

Ý nghĩa của tứ phân vị của mẫu số liệu ghép nhóm

Ba điểm tứ phân vị chia mẫu số liệu đã sắp xếp theo thứ tự không giảm thành bốn phần đều nhau. Giống như với trung vị, nói chung không thể xác định chính xác các điểm tứ phân vị của mẫu số liệu ghép nhóm.

Bộ ba tứ phân vị của mẫu số liệu ghép nhóm là giá trị xấp xỉ cho tứ phân vị của mẫu số liệu gốc và được sử dụng làm giá trị đo xu thế trung tâm của mẫu số liệu.

Tứ phân vị thứ nhất và thứ ba đo xu thế trung tâm của nửa dưới (các dữ liệu nhỏ hơn Q_2) và nửa trên (các dữ liệu lớn hơn Q_2) của mẫu số liệu.

BÀI TẬP THAM KHẢO

✓ TRẮC NGHIỆM NHIỀU PHƯƠNG ÁN LỰA CHỌN

Câu 1. Bảng sau thống kê khối lượng một số quả măng cụt được lựa chọn ngẫu nhiên trong một thùng hàng.

Khối lượng (gam)	[80;82)	[82;84)	[84;86)	[86;88)	[88;90)
Số quả	17	20	25	16	12

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

- A. 10 gam. B. 12 gam. C. 2 gam. D. 20 gam.

✗ **Lời giải**

Chọn A

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là $90 - 80 = 10$ gam.

Câu 2. 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	20	60	40	32	7

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

- A. $3 m^3$. B. $15 m^3$. C. $18 m^3$. D. $20 m^3$.

✗ **Lời giải**

Chọn B

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là $18 - 3 = 15 m^3$

Câu 3. Người ta thống kê tốc độ của một số xe ô tô đi chuyển qua một trạm kiểm soát trên đường cao tốc trong một khoảng thời gian ở bảng sau.

Tốc độ (km/h)	[75;80)	[80;85)	[85;90)	[90;95)	[95;100)
Số xe	15	22	28	34	19

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là

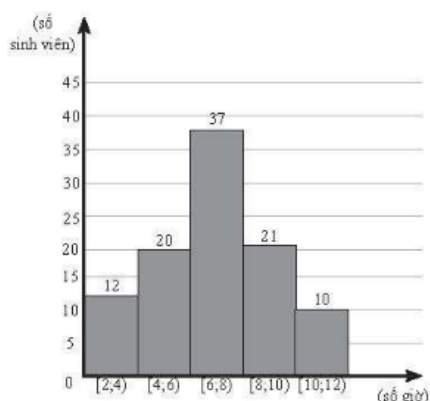
- A. 75 km/h. B. 25 km/h. C. 100 km/h. D. 5 km/h.

✗ **Lời giải**

Chọn B

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là $100 - 75 = 25$ km/h.

Câu 4. Kết quả điều tra về số giờ làm thêm trong một tuần của 100 sinh viên được cho ở biểu đồ sau.



Khoảng biến thiên của mẫu số liệu được cho trong biểu đồ trên là:

- A. 2. B. 3. C. 10. D. 37.

✗ **Lời giải**

Chọn C

Số giờ làm thêm	[2;4)	[4;6)	[6;8)	[8;10)	[10;12)
Số sinh viên	12	20	37	21	10

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là $12 - 2 = 10$ giờ.

Câu 5. Thâm niên công tác của các công nhân hai nhà máy A và B được cho trong bảng sau

Thâm niên công tác (năm)	[0;5)	[5;10)	[10;15)	[15;20)	[20;25)
Số công nhân nhà máy A	35	13	12	12	8
Số công nhân nhà máy B	19	26	24	11	0

Sử dụng khoảng biến thiên, hãy cho biết thâm niên công tác các công nhân của nhà máy nào có độ phân tán lớn hơn?

- A. Không so sánh được.
 B. Nhà máy A có thâm niên công tác của các công nhân phân tán lớn hơn nhà máy B .
 C. Nhà máy A có thâm niên công tác của các công nhân phân tán nhỏ hơn nhà máy B .
 D. Nhà máy A có thâm niên công tác của các công nhân phân tán bằng nhà máy B .

Lời giải

Chọn B

Khoảng biến thiên của mẫu số liệu ghép nhóm về thâm niên công tác của các công nhân của nhà máy A là $25 - 0 = 25$ năm.

Khoảng biến thiên của mẫu số liệu ghép nhóm về thâm niên công tác của các công nhân của nhà máy B là $20 - 0 = 20$ năm.

Do vậy, nhà máy A có thâm niên công tác của các công nhân phân tán lớn hơn nhà máy B .

Câu 6. Kết quả khảo sát cân nặng của 40 quả cam Canh ở mỗi lô hàng 1 và lô hàng 2 được cho ở bảng sau

Cân nặng (gam)	[100;110)	[110;120)	[120;130)	[130;140)	[140;150)
Số quả cam ở lô hàng 1	0	10	11	19	0
Số quả cam ở lô hàng 2	3	15	12	7	3

Sử dụng khoảng biến thiên, hãy cho biết cân nặng của 40 quả cam Canh của lô hàng nào có độ phân tán lớn hơn.

- A. Không so sánh được.
 B. Lô hàng 2 có cân nặng của 40 quả cam Canh phân tán lớn hơn lô hàng 1.
 C. Lô hàng 1 có cân nặng của 40 quả cam Canh phân tán lớn hơn lô hàng 2.
 D. Lô hàng 1 có cân nặng của 40 quả cam Canh phân tán bằng lô hàng 2.

Lời giải

Chọn B

Khoảng biến thiên của mẫu số liệu ghép nhóm về cân nặng của 40 quả cam Canh của lô hàng 1 là $140 - 110 = 30$ gam.

Khoảng biến thiên của mẫu số liệu ghép nhóm về cân nặng của 40 quả cam Canh của lô hàng 2 là $150 - 100 = 50$ gam.

Do vậy, lô hàng 2 có cân nặng của 40 quả cam Canh phân tán lớn hơn lô hàng 1.

Câu 7. Bạn Lan thống kê lại chiều cao (đơn vị: cm) của các học sinh nữ lớp 10B và lớp 10C ở bảng sau.

Chọn đáp án có khẳng định đúng.

Chiều cao(cm)	[150;155)	[155;160)	[160;165)	[165;170)	[170;175)
Số học sinh nữ lớp 10B	0	5	13	7	0
Số học sinh nữ lớp 10C	2	10	9	3	1

- A. Chiều cao của các bạn nữ lớp 10B đồng đều hơn chiều cao của các bạn nữ lớp 10C.
 B. Chiều cao của các bạn nữ lớp 10C đồng đều hơn chiều cao của các bạn nữ lớp 10B.
 C. Chiều cao của các bạn nữ lớp 10B và lớp 10C đồng đều như nhau.
 D. Khoảng biến thiên của mẫu số liệu ghép nhóm về chiều cao của các bạn nữ lớp 10C nhỏ hơn khoảng biến thiên của mẫu số liệu ghép nhóm về chiều cao của các bạn nữ lớp 10B.

Lời giải

Chọn A

Ta có

Khoảng biến thiên của mẫu số liệu ghép nhóm về chiều cao của các bạn nữ lớp 10B là $170 - 155 = 15$

Khoảng biến thiên của mẫu số liệu ghép nhóm về chiều cao của các bạn nữ lớp 10C là $175 - 150 = 25$

Vì $15 < 25$ nên mẫu số liệu ghép nhóm về chiều cao của các bạn nữ lớp 10B có độ phân tán ít hơn so với mẫu số liệu ghép nhóm về chiều cao của các bạn nữ lớp 10C, hay nói cách khác chiều cao của các bạn nữ lớp 10B đồng đều hơn chiều cao của các bạn nữ lớp 10C.

Câu 8. Một hãng xe ô tô thống kê lại số lần gặp sự cố về động cơ của 100 chiếc xe cùng loại sau 2 năm sử dụng đầu tiên ở bảng sau. Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm này? (Làm tròn các kết quả đến hàng phần trăm).

Số lần gặp sự cố	$[0,5;2,5)$	$[2,5;4,5)$	$[4,5;6,5)$	$[6,5;8,5)$	$[8,5;10,5)$
Số xe	17	33	25	20	5

A. 5,32.

B. 3,52.

C. 2,53.

D. 5,23.

Lời giải

Chọn B

Do cỡ mẫu $n = 100$

Gọi $x_1; x_2; \dots; x_{100}$ là mẫu số liệu gốc gồm số lần gặp sự cố của 100 chiếc xe cùng loại sau 2 năm sử dụng.

Ta có $x_1, \dots, x_{17} \in [0,5;2,5); x_{18}, \dots, x_{50} \in [2,5;4,5); x_{51}, \dots, x_{75} \in [4,5;6,5); x_{76}, \dots, x_{95} \in [6,5;8,5); x_{96}, \dots, x_{100} \in [8,5;10,5)$.

Nên tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_{25} + x_{26}) \in [2,5;4,5)$. Do đó tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là

$$Q_1 = 2,5 + \frac{\frac{100}{4} - 17}{33} \cdot (4,5 - 2,5) \approx 2,98$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{75} + x_{76}) \in [2,5;4,5)$.

Mà $x_{75} \in [4,5;6,5); x_{76} \in [6,5;8,5)$. Nên $Q_3 = 6,5$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = Q_3 - Q_1 \approx 6,5 - 2,98 = 3,52$

Câu 9. Kiểm tra điện lượng của một số viên pin tiểu do một hãng sản xuất thu được kết quả sau. Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm này? (Làm tròn các kết quả đến hàng phần trăm).

Điện lượng (nghìn mAh)	$[0,9;0,95)$	$[0,95;1,0)$	$[1,0;1,05)$	$[1,05;1,1)$	$[1,1;1,15)$
Số viên pin	10	20	35	15	5

A. 0,06.

B. 0,08.

C. 0,07.

D. 0,09.

Lời giải

Chọn C

Do cỡ mẫu $n = 10 + 20 + 35 + 15 + 5 = 85$

Gọi $x_1; x_2; \dots; x_{85}$ là mẫu số liệu gốc gồm điện lượng của 85 viên pin tiểu.

Ta có $x_1, \dots, x_{10} \in [0,9;0,95); x_{11}, \dots, x_{30} \in [0,95;1,0); x_{31}, \dots, x_{65} \in [1,0;1,05); x_{66}, \dots, x_{80} \in [1,05;1,1); x_{81}, \dots, x_{85} \in [1,1;1,15)$.

Nên tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_{21} + x_{22}) \in [0,95;1,0)$.

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

$$Q_1 = 0,95 + \frac{\frac{85}{4} - 10}{20} \cdot (1,0 - 0,95) \approx 0,98$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{64} + x_{65}) \in [1,0; 1,05)$.

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

$$Q_3 = 1,0 + \frac{\frac{3.85}{4} - 30}{35} \cdot (1,05 - 1,0) \approx 1,05$$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = Q_3 - Q_1 \approx 1,05 - 0,98 = 0,07$

Câu 10. Một người thống kê lại thời gian thực hiện các cuộc gọi điện thoại của người đó trong một tuần ở bảng sau. Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm này?

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

A. 100.

B. 110.

C. 120.

D. 130.

Lời giải

Chọn C

Do cỡ mẫu $n = 8 + 10 + 7 + 5 + 2 + 1 = 33$

Nên tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_8 + x_9)$.

Mà $x_8 \in [0; 60)$; $x_9 \in [60; 120)$. Nên $Q_1 = 60$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{25} + x_{26})$.

Mà $x_{25} \in [120; 180)$; $x_{26} \in [180; 240)$. Nên $Q_3 = 180$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = Q_3 - Q_1 = 180 - 60 = 120$

Câu 11. Một phòng khám thống kê số bệnh nhân đến khám bệnh mỗi ngày trong tháng 4 năm 2022 ở bảng sau. Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm này? (Làm tròn các kết quả đến hàng phần chục).

Số bệnh nhân	[0,5; 10,5)	[10,5; 20,5)	[20,5; 30,5)	[30,5; 40,5)	[40,5; 50,5)
Số ngày	7	8	7	6	2

A. 20,3.

B. 20,2.

C. 20,4.

D. 20,5.

Lời giải

Chọn B

Do cỡ mẫu $n = 7 + 8 + 7 + 6 + 2 = 30$

Gọi $x_1; x_2; \dots; x_{85}$ là mẫu số liệu gốc gồm điện lượng của 85 viên pin tiểu.

Ta có $x_1, \dots, x_{10} \in [0,9; 0,95)$; $x_{11}, \dots, x_{30} \in [0,95; 1,0)$; $x_{31}, \dots, x_{65} \in [1,0; 1,05)$; $x_{66}, \dots, x_{80} \in [1,05; 1,1)$; $x_{81}, \dots, x_{85} \in [1,1; 1,15)$.

Nên tứ phân vị thứ nhất của mẫu số liệu gốc là $x_8 \in [10,5; 20,5)$.

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

$$Q_1 = 10,5 + \frac{\frac{30}{4} - 7}{8} \cdot (20,5 - 10,5) \approx 11,1$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{23} \in [30,5; 40,5)$.

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

$$Q_3 = 30,5 + \frac{\frac{3.30}{4} - 22}{6} \cdot (40,5 - 30,5) \approx 31,3$$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = Q_3 - Q_1 \approx 31,3 - 11,1 = 20,2$

Câu 12. Cân nặng của một số lợn con mới sinh thuộc hai giống A và B được cho ở bảng sau. Chọn đáp án có khẳng định đúng.

Cân nặng (kg)	[1,0;1,1)	[1,1;1,2)	[1,2;1,3)	[1,3;1,4)
Giống A	8	28	32	17
Giống B	13	14	24	14

A. Cân nặng của lợn con mới sinh thuộc giống A đồng đều hơn cân nặng của lợn con mới sinh thuộc giống B.

B. Cân nặng trung bình của lợn con mới sinh giống A nhỏ hơn giống B.

C. Cân nặng của lợn con mới sinh thuộc giống B đồng đều hơn cân nặng của lợn con mới sinh thuộc giống A.

D. Khoảng biến thiên của mẫu số liệu ghép nhóm về cân nặng của lợn con mới sinh thuộc giống A nhỏ hơn khoảng biến thiên của mẫu số liệu ghép nhóm về cân nặng của lợn con mới sinh thuộc giống B.

Lời giải

Chọn A

Đối với lợn con giống A

Gọi $x_1; x_2; \dots; x_{85}$ là mẫu số liệu được sắp xếp theo thứ tự không giảm.

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_{21} + x_{22}) \in [1,1;1,2)$ nên

$$Q_1(A) = 1,1 + \frac{\frac{85}{4} - 8}{28} \cdot (1,2 - 1,1) \approx 1,15$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{64} + x_{65}) \in [1,2;1,3)$ nên

$$Q_3(A) = 1,2 + \frac{\frac{3.85}{4} - 36}{32} \cdot (1,3 - 1,2) \approx 1,29$$

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

$$\Delta_Q(A) = Q_3(A) - Q_1(A) \approx 1,29 - 1,15 = 0,14$$

Đối với lợn con giống B

Gọi $y_1; y_2; \dots; y_{65}$ là mẫu số liệu được sắp xếp theo thứ tự không giảm.

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(y_{16} + y_{17}) \in [1,1;1,2)$ nên

$$Q_1(B) = 1,1 + \frac{\frac{65}{4} - 13}{14} \cdot (1,2 - 1,1) \approx 1,123$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(y_{49} + y_{50}) \in [1,2;1,3)$ nên

$$Q_3(B) = 1,2 + \frac{\frac{3.65}{4} - 27}{24} \cdot (1,3 - 1,2) \approx 1,29$$

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

$$\Delta_Q(B) = Q_3(B) - Q_1(B) \approx 1,29 - 1,123 = 0,167$$

Ta thấy $\Delta_Q(A) < \Delta_Q(B)$ nên cân nặng của lợn con mới sinh thuộc giống A đồng đều hơn cân nặng của lợn con mới sinh thuộc giống B.



Câu 13. Một mẫu số liệu ghép nhóm có khoảng tứ phân vị là 4,43 và tứ phân vị thứ 3 là $\frac{68}{3}$ thì giá trị ngoại lệ của mẫu số liệu ghép nhóm đó phải là bao nhiêu?

- A. $x > 29,31$ B. $x > 1,51$ C. $x < 51,23$ D. $x \leq 25,11$

🔗 **Lời giải**

Chọn A

Do tứ phân vị thứ 3 là $\frac{68}{3}$

Nên giá trị ngoại lệ $x > Q_3 + 1,5\Delta_Q = \frac{68}{3} + 1,5 \cdot 4,43 \approx 29,31$

Câu 14. Một mẫu số liệu ghép nhóm có tứ phân vị thứ 1 là 254,9 và tứ phân vị thứ 3 là 417,25 thì điều kiện giá trị ngoại lệ của mẫu số liệu ghép nhóm đó là:

- A. $\begin{cases} x \geq 12,1 \\ x \leq 1,35 \end{cases}$ B. $\begin{cases} x > 11,2 \\ x < 0,375 \end{cases}$ C. $\begin{cases} x > 660,775 \\ x < 11,375 \end{cases}$ D. $x > 11,375$

🔗 **Lời giải**

Chọn C

Gọi giá trị ngoại lệ của mẫu số liệu ghép nhóm là x

Ta có khoảng tứ phân vị $\Delta_Q = 417,25 - 254,9 = 162,35$

Nên giá trị ngoại lệ $\begin{cases} x > Q_3 + 1,5\Delta_Q = 417,25 + 1,5 \cdot 162,35 = \frac{26431}{40} \approx 660,775 \\ x < Q_1 - 1,5\Delta_Q = 254,9 - 1,5 \cdot 162,35 = \frac{91}{8} \approx 11,375 \end{cases}$

Vậy $\begin{cases} x > 660,775 \\ x < 11,375 \end{cases}$

Câu 15. Giáo viên chủ nhiệm thống kê chiều cao (đơn vị cm) của các bạn học sinh nữ của lớp 12A ở bảng sau:

Chiều cao	[150;155)	[155;160)	[160;165)	[165;170)	[170;175)
Số học sinh	2	4	10	0	1

Xác định khoảng biến thiên của chiều cao của các bạn học sinh nữ lớp 12A.

- A. 20 B. 10 C. 5 D. 25

🔗 **Lời giải**

Chọn A

Khoảng biến thiên của chiều cao của các bạn học sinh nữ lớp 12A là: $175 - 155 = 20$

Câu 16. Bảng dưới đây thống kê cân nặng (gam) của 100 quả xoài cát Hòa Lộc sau khi thu hoạch tại một khu vườn

Cân nặng	[200;205)	[205;300)	[300;305)	[305;310)
Số lượng xoài	10	15	45	30

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

- A. 35 B. 60 C. 55 D. 110

🔗 **Lời giải**

Chọn D

Khoảng biến thiên của cân nặng của số lượng xoài là: $310 - 200 = 110$

Câu 17. Thời gian và số ngày tập thể dục của bác T và bác H trong một tháng (30 ngày) được thống kê theo bảng dưới đây:

Thời gian tập (phút)	[15;20)	[25;30)	[30;35)
Số ngày tập của bác T	10	15	5

Số ngày tập của bác H	9	21	0
-----------------------	---	----	---

Mệnh đề nào sau đây đúng?

- A. Khoảng biến thiên thời gian tập của bác T bằng 10
 B. Khoảng biến thiên thời gian tập của bác H bằng 20
 C. Độ phân tán thời gian tập của bác T cao hơn độ phân tán thời gian tập của bác H
 D. Độ phân tán thời gian tập của bác T thấp hơn độ phân tán thời gian tập của bác H

Lời giải

Chọn C

Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian tập của bác T là: $35 - 15 = 20$ (phút).

Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian tập của bác H là: $30 - 15 = 15$ (phút).

Do đó căn cứ theo khoảng biến thiên thì thời gian tập của bác T có độ phân tán lớn hơn.

Câu 18. Bảng dưới đây biểu diễn mẫu số liệu ghép nhóm về số tiền (đơn vị: nghìn đồng) mà 12 khách hàng mua sách ở một cửa hàng trong một ngày.

Giá tiền	$[40; 50)$	$[50; 60)$	$[60; 70)$
Số lượng khách mua	2	6	4

Xác định khoảng tứ phân vị của mẫu số liệu ghép nhóm trên

- A. $\frac{65}{6}$ B. $\frac{55}{3}$ C. $\frac{12}{5}$ D. $\frac{312}{5}$

Lời giải

Chọn A

Cỡ mẫu $n = 3 + 6 + 3 = 12$

Gọi $x_1; x_2; x_3; \dots; x_{12}$ là mẫu số liệu gốc gồm số lượng khách hàng của khách hàng mua sách ở một cửa hàng trong một ngày được xếp theo thứ tự không giảm.

Ta có $x_1; x_2 \in [40; 50)$; $x_3; x_4; x_5; \dots; x_8 \in [50; 60)$; $x_9; \dots; x_{12} \in [60; 70)$

Tứ phân vị thứ nhất $\frac{1}{2}(x_3 + x_4)$, do $x_3; x_4 \in [50; 60)$.

$$\text{Nên } Q_1 = 50 + \frac{\frac{12}{2} - 2}{6} \cdot (60 - 50) = \frac{155}{3}$$

Tứ phân vị thứ ba $\frac{1}{2}(x_9 + x_{10})$, do $x_9; x_{10} \in [60; 70)$.

$$\text{Nên } Q_3 = 60 + \frac{\frac{3 \cdot 12}{4} - (2 + 6)}{4} \cdot (70 - 60) = \frac{125}{2}$$

$$\text{Khi đó khoảng tứ phân vị } \Delta Q = Q_3 - Q_1 = \frac{125}{2} - \frac{155}{3} = \frac{65}{6}$$

Câu 19. Bảng sau thống kê cân nặng của 50 quả xoài Thanh Ca được lựa chọn ngẫu nhiên sau khi thu hoạch ở một nông trường

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

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

- A. $\frac{9080}{143}$ B. $\frac{980}{143}$ C. $\frac{908}{143}$ D. $\frac{9080}{145}$

Lời giải

Chọn A

Cỡ mẫu $n = 50$.

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

Ta có $x_1, x_2, x_3 \in [250; 290)$; $x_4, \dots, x_{16} \in [290; 330)$; $x_{17}, \dots, x_{34} \in [330; 370)$;

$x_{35}, \dots, x_{45} \in [370; 410)$; $x_{46}, \dots, x_{50} \in [410; 450)$.

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

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

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

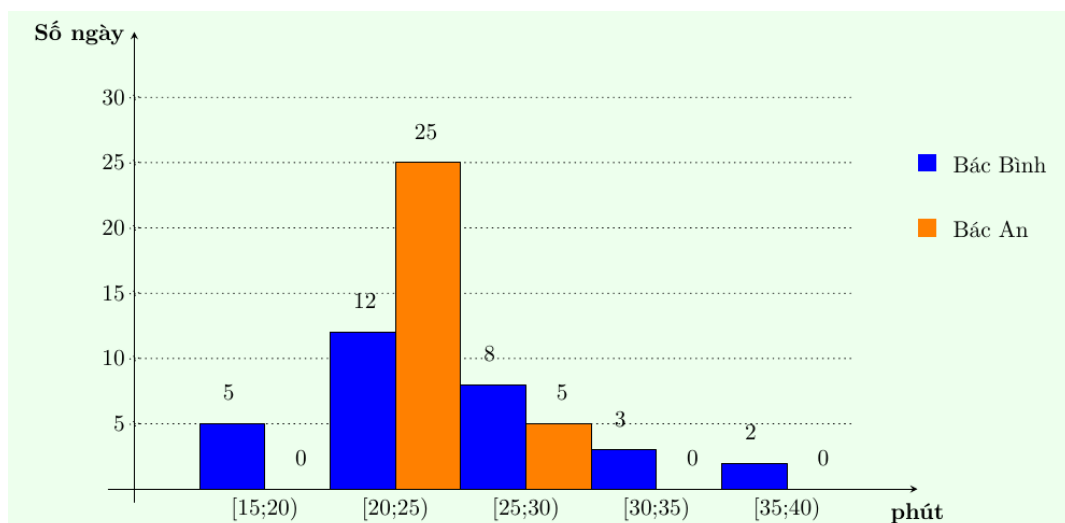
Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{38} \in [370; 410)$.

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

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

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = Q_3 - Q_1 = \frac{4210}{11} - \frac{4150}{13} = \frac{9080}{143}$.

Câu 20. Biểu đồ dưới đây thống kê thời gian tập thể dục buổi sáng mỗi ngày trong tháng 9/2022 của bác Bình và bác An.



Hỏi giữa bác Bình và bác An thì người nào có thời gian luyện tập đều?

- A. Bác Bình có thời gian tập đều hơn.
- B. Bác An có thời gian tập đều hơn.
- C. Không xác định được.
- D. Bác Bình và Bác An có thời gian tập đều như nhau.

Lời giải

Chọn B

Dựa vào biểu đồ, ta có bảng thống kê ghép nhóm như bên dưới

Xét thống kê của bác Bình, ta có cỡ mẫu $n = 30$.

Gọi $x_1; x_2; \dots; x_{30}$ là mẫu số liệu thời gian tập thể dục của bác Bình.

Ta có: $x_1, \dots, x_5 \in [15; 20)$; $x_6, \dots, x_{17} \in [20; 25)$; $x_{18}, \dots, x_{25} \in [25; 30)$; $x_{26}, x_{27}, x_{28} \in [30; 35)$; $x_{29}, x_{30} \in [35; 40)$.

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

$$Q_1 = 20 + \frac{\frac{30}{4} - 5}{12} \cdot (25 - 20) = \frac{505}{24}.$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{23} \in [25; 30)$. Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là

$$Q_3 = 25 + \frac{\frac{3 \cdot 30}{4} - (5+12)}{8} \cdot (30-25) = \frac{455}{16}.$$

Vậy khoảng tứ phân vị của mẫu số liệu của bác An là $\Delta_Q = Q_3 - Q_1 = \frac{355}{48}$.

Xét thống kê của bác An, ta có cỡ mẫu $n = 30$.

Gọi $x_1; x_2; \dots; x_{30}$ là mẫu số liệu thời gian tập thể dục của bác Bình.

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

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

$$Q_1 = 20 + \frac{\frac{30}{4} - 0}{25} \cdot (25 - 20) = \frac{43}{2}.$$

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

$$Q_3 = 20 + \frac{\frac{3 \cdot 30}{4} - 0}{25} \cdot (30 - 25) = \frac{49}{2}.$$

Vậy khoảng tứ phân vị của mẫu số liệu của bác An là $\Delta_Q = Q_3 - Q_1 = 3$.

Vậy khoảng tứ phân vị của mẫu số liệu của bác An nhỏ hơn của bác Bình. Suy ra bác An có thời gian tập đều hơn.

Câu 21. Bạn Trang thống kê lại chiều cao (đơn vị: cm) của các bạn học sinh nữ lớp 12C và lớp 12D ở bảng sau

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

Hãy so sánh độ phân tán của nửa giữa hai mẫu số liệu chiều cao của các học sinh nữ lớp 12C và 12D.

- A. Mẫu số liệu của lớp 12C ít phân tán hơn mẫu số liệu của lớp 12D.
 B. Mẫu số liệu của lớp 12D ít phân tán hơn mẫu số liệu của lớp 12C.
 C. Mẫu số liệu của lớp 12C và mẫu số liệu của lớp 12D không phân tán.
 D. Không xác định được.

Lời giải

Chọn A

Xét ở lớp 12C, ta có cỡ mẫu $n = 25$.

Gọi $x_1; x_2; \dots; x_{25}$ là mẫu số liệu gốc gồm chiều cao của 25 bạn nữ lớp 12C.

Ta có: $x_1, x_2 \in [155; 160); x_3, \dots, x_9 \in [160; 165); x_{10}, \dots, x_{21} \in [165; 170); x_{22}, x_{23}, x_{24} \in [170; 175); x_{25} \in [180; 185)$.

Tứ phân vị thứ nhất của mẫu số liệu lớp 12C là $\frac{x_6 + x_7}{2} \in [160; 165)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu lớp 12C là

$$Q_1 = 160 + \frac{\frac{25}{4} - 2}{7} \cdot (165 - 160) = \frac{4565}{28}.$$

Tứ phân vị thứ ba của mẫu số liệu lớp 12C là $\frac{x_{19} + x_{20}}{2} \in [165; 170)$.

Do đó, tứ phân vị thứ ba của mẫu số liệu lớp 12C là

$$Q_3 = 165 + \frac{\frac{3 \cdot 25}{4} - (2 + 7)}{12} \cdot (170 - 165) = \frac{2705}{16}.$$

Vậy khoảng tứ phân vị của mẫu số liệu lớp 12C là $\Delta_Q = Q_3 - Q_1 = \frac{675}{112} \approx 6,027$.

Xét ở lớp 12D, ta có cỡ mẫu $n = 25$.

Gọi $x_1; x_2; \dots; x_{25}$ là mẫu số liệu gốc gồm chiều cao của 25 bạn nữ lớp 12D.

Ta có: $x_1, \dots, x_5 \in [155; 160); x_6, \dots, x_{14} \in [160; 165); x_{15}, \dots, x_{22} \in [165; 170); x_{23}, x_{24} \in [170; 175); x_{25} \in [175; 180)$.

Tứ phân vị thứ nhất của mẫu số liệu lớp 12C là $\frac{x_6 + x_7}{2} \in [160; 165)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu lớp 12C là

$$Q_1 = 160 + \frac{\frac{25}{4} - 5}{9} \cdot (165 - 160) = \frac{5785}{36}.$$

Tứ phân vị thứ ba của mẫu số liệu lớp 12C là $\frac{x_{19} + x_{20}}{2} \in [165; 170)$.

Do đó, tứ phân vị thứ ba của mẫu số liệu lớp 12C là

$$Q_3 = 165 + \frac{\frac{3 \cdot 25}{4} - (5 + 9)}{8} \cdot (170 - 165) = \frac{5375}{32}.$$

Vậy khoảng tứ phân vị của mẫu số liệu lớp 12C là $\Delta_Q = Q_3 - Q_1 \approx 7,274$.

Vậy mẫu số liệu của lớp 12C ít phân tán hơn mẫu số liệu của lớp 12D.

Câu 22. 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

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

A. Phụ nữ ở khu vực B có độ tuổi kết hôn đồng đều hơn.

B. Phụ nữ ở khu vực A có độ tuổi kết hôn đồng đều hơn.

C. Phụ nữ ở khu vực A và khu vực B có độ tuổi kết hôn đồng đều như nhau.

D. Không xác định được.

Lời giải

Chọn A

Xét ở khu vực A, khoảng biến thiên của mẫu số liệu là $R = 34 - 19 = 15$.

Ta có cỡ mẫu $n = 100$.

Gọi $x_1; x_2; \dots; x_{100}$ là mẫu số liệu gốc gồm độ tuổi kết hôn của 100 phụ nữ ở khu vực A.

Ta có: $x_1, \dots, x_{10} \in [19; 22); x_{11}, \dots, x_{37} \in [22; 25); x_{38}, \dots, x_{68} \in [25; 28); x_{69}, \dots, x_{93} \in [28; 31); x_{94}, \dots, x_{100} \in [31; 34)$.

Tứ phân vị thứ nhất của mẫu số liệu khu vực A là $\frac{x_{25} + x_{26}}{2} \in [22; 25)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu khu vực A là

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

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

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

Vậy khoảng tứ phân vị của mẫu số liệu khu vực A là $\Delta_Q = Q_3 - Q_1 = \frac{388}{75} \approx 5,17$.

Xét ở khu vực B, khoảng biến thiên của mẫu số liệu là $R = 34 - 19 = 15$.

Ta có cỡ mẫu $n = 100$.

Gọi $x_1; x_2; \dots; x_{100}$ là mẫu số liệu gốc gồm độ tuổi kết hôn của 100 phụ nữ ở khu vực **B**.

Ta có:

$$x_1, \dots, x_{47} \in [19; 22); x_{48}, \dots, x_{87} \in [22; 25); x_{88}, \dots, x_{98} \in [25; 28); x_{99}, x_{100} \in [28; 31).$$

Tứ phân vị thứ nhất của mẫu số liệu khu vực B là $\frac{x_{25} + x_{26}}{2} \in [19; 22)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu khu vực B là

$$Q_1 = 19 + \frac{\frac{100}{4} - 47}{22 - 19} \cdot (22 - 19) = \frac{968}{47}.$$

Tứ phân vị thứ ba của mẫu số liệu khu vực B là $\frac{x_{75} + x_{76}}{2} \in [22; 25)$.

Do đó, tứ phân vị thứ ba của mẫu số liệu khu vực B là

$$Q_3 = 22 + \frac{\frac{3 \cdot 100}{4} - 47}{25 - 22} \cdot (25 - 22) = \frac{241}{10}.$$

Vậy khoảng tứ phân vị của mẫu số liệu khu vực B là $\Delta_Q = Q_3 - Q_1 = \frac{1647}{470} \approx 3,5$.

Ta có khoảng tứ phân vị của khu vực A lớn hơn ở khu vực B nên khu vực B có độ tuổi kết hôn đồng đều hơn.

Câu 23. Kết quả đo chiều cao của 100 cây keo 3 năm tuổi tại một nông trường được cho ở bảng sau

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

Khẳng định nào sau đây là đúng?

A. Cây keo cao 8,4 m là cây keo có chiều cao là giá trị ngoại lệ.

B. Cây keo cao 8,4 m là cây keo có chiều cao không là giá trị ngoại lệ.

C. Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = 0,5$.

D. Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là 7.

Lời giải

Chọn A

Khoảng biến thiên của mẫu số liệu là $R = 9,4 - 8,4 = 1$.

Ta có cỡ mẫu $n = 100$.

Gọi $x_1; x_2; \dots; x_{100}$ là mẫu số liệu gồm chiều cao của 100 cây keo.

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

$x_{87}, \dots, x_{100} \in [9,2; 9,4)$.

Tứ phân vị thứ nhất của mẫu số liệu là $\frac{x_{25} + x_{26}}{2} \in [8,8; 9,0)$.

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

$$Q_1 = 8,8 + \frac{\frac{100}{4} - (5+12)}{25} \cdot (9,0 - 8,8) = 8,864.$$

Tứ phân vị thứ ba của mẫu số liệu là $\frac{x_{75} + x_{76}}{2} \in [9,0; 9,2)$.

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

$$Q_3 = 9,0 + \frac{\frac{3 \cdot 100}{4} - (5+12+25)}{44} \cdot (9,2 - 9,0) = 9,15.$$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = Q_3 - Q_1 = 0,286$.

Vì $Q_1 - 1,5\Delta_Q = 8,435$ và $Q_3 + 1,5\Delta_Q = 9,579$ nên cây keo có chiều cao 8,4 m là giá trị ngoại lệ của mẫu số liệu ghép nhóm.

Câu 24. Bảng tần số ghép nhóm dưới đây thể hiện kết quả điều tra về tuổi thọ trung bình của nam giới và nữ giới ở 50 quốc gia.

Giới tính	Nam	Nữ
Nhóm (Tuổi thọ)		
[50; 55)	4	3
[55; 60)	7	4
[60; 65)	4	5
[65; 70)	6	3
[70; 75)	15	7
[75; 80)	12	14
[80; 85)	2	13
[85; 90)	0	1

Khẳng định nào sau đây là đúng?

- A. Tuổi thọ của nam giới đều hơn tuổi thọ của nữ giới.
- B. Khoảng tứ phân vị của mẫu số liệu nam giới là $\Delta_Q = 14$.
- C. Khoảng tứ phân vị của mẫu số liệu nữ giới là $\Delta_Q = 16$.
- D. Tứ phân vị thứ nhất của mẫu số liệu nam giới là 63.

Lời giải

Chọn A

Xét ở nam giới, ta có cỡ mẫu $n = 50$.

Gọi $x_1; x_2; \dots; x_{50}$ là mẫu số liệu gồm tuổi thọ của 50 nam giới.

Ta có: $x_1, \dots, x_4 \in [50; 55); x_5, \dots, x_{11} \in [55; 60); x_{12}, \dots, x_{15} \in [60; 65); x_{16}, \dots, x_{21} \in [65; 70); x_{22}, \dots, x_{36} \in [70; 75); x_{37}, \dots, x_{48} \in [75; 80); x_{49}, x_{50} \in [80; 85)$.

Tứ phân vị thứ nhất của mẫu số liệu là $x_{13} \in [60; 65)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu nam giới là

$$Q_1 = 60 + \frac{\frac{50}{4} - (4+7)}{4} \cdot (65 - 60) = \frac{495}{8}.$$

Tứ phân vị thứ ba của mẫu số liệu là $x_{38} \in [75; 80)$.

Do đó, tứ phân vị thứ ba của mẫu số liệu nam giới là

$$Q_3 = 75 + \frac{\frac{3 \cdot 50}{4} - (4 + 7 + 4 + 6 + 15)}{12} \cdot (80 - 75) = \frac{605}{8}.$$

Vậy khoảng tứ phân vị của mẫu số liệu nam giới là $\Delta_Q = Q_3 - Q_1 = \frac{55}{4} = 13,75$.

Xét ở nữ giới, ta có cỡ mẫu $n = 50$.

Gọi $x_1; x_2; \dots; x_{50}$ là mẫu số liệu gồm tuổi thọ của 50 nữ giới.

Ta có: $x_1, x_2, x_3 \in [50; 55); x_4, \dots, x_7 \in [55; 60); x_8, \dots, x_{12} \in [60; 65); x_{13}, x_{14}, x_{15} \in [65; 70); x_{16}, \dots, x_{22} \in [70; 75); x_{23}, \dots, x_{36} \in [75; 80); x_{37}, \dots, x_{49} \in [80; 85); x_{50} \in [85; 90)$.

Tứ phân vị thứ nhất của mẫu số liệu là $x_{13} \in [65; 70)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu nữ giới là

$$Q_1 = 65 + \frac{\frac{50}{4} - (3 + 4 + 5)}{3} \cdot (70 - 65) = \frac{395}{6}.$$

Tứ phân vị thứ ba của mẫu số liệu là $x_{38} \in [80; 85)$.

Do đó, tứ phân vị thứ ba của mẫu số liệu nữ giới là

$$Q_3 = 80 + \frac{\frac{3 \cdot 50}{4} - (3 + 4 + 5 + 3 + 7 + 14)}{13} \cdot (85 - 80) = \frac{2095}{26}.$$

Vậy khoảng tứ phân vị của mẫu số liệu nữ giới là $\Delta_Q = Q_3 - Q_1 = \frac{575}{39} \approx 14,74$.

Do khoảng tứ phân vị của mẫu số liệu của nam giới nhỏ hơn mẫu số liệu của nữ giới nên tuổi thọ của nam giới đều hơn tuổi thọ của nữ giới.

Câu 25. Bạn Chi rất thích nhảy hiện đại. Thời gian tập nhảy mỗi ngày trong thời gian gần đây của bạn Chi được thống kê lại ở bảng sau:

Cự li (m)	[19;19,5)	[19,5;20)	[20;20,5)	[20,5;21)	[21;21,5)
Tần số	13	45	24	12	6

Số trung bình của mẫu số liệu ghép nhóm là

A. 100.

B. 20,015.

C. 2001,5.

D. 2.

Lời giải

Chọn B

+ Cỡ mẫu: $n = 100$.

Cự li (m)	[19;19,5)	[19,5;20)	[20;20,5)	[20,5;21)	[21;21,5)
Giá trị đại diện	19,25	19,75	20,25	20,75	21,25
Tần số	13	45	24	12	6

+ Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x} = \frac{19,25 \cdot 13 + 19,75 \cdot 45 + 20,25 \cdot 24 + 20,75 \cdot 12 + 21,25 \cdot 6}{100} = 20,015.$$

Câu 26. 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 trong 50 ngày giao dịch liên tiếp.

Giá đóng cửa	[120;122)	[122;124)	[124;126)	[126;128)	[128;130)
Số ngày giao dịch của cổ phiếu A	8	9	12	10	11

Số trung bình của mẫu số liệu ghép nhóm là

A. 50.

B. 12528.

C. 125,28.

D. 10.

Lời giải

Chọn C



+ Cỡ mẫu: $n = 50$.

Giá đóng cửa	[120;122)	[122;124)	[124;126)	[126;128)	[128;130)
Giá trị đại diện	121	123	125	127	129
Số ngày giao dịch của cổ phiếu A	8	9	12	10	11

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

$$\bar{x} = \frac{121.8 + 123.9 + 125.12 + 127.10 + 129.11}{50} = 125,28.$$

Câu 27. Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khỏe. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

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

Phương sai của mẫu số liệu ghép nhóm là (làm tròn đến hàng phần trăm)

A. 3,39.

B. 11,62.

C. 0,13.

D. 0,36.

Lời giải

Chọn C

+ Cỡ mẫu: $n = 20$.

Quãng đường (km)	[2, 7; 3, 0)	[3, 0; 3, 3)	[3, 3; 3, 6)	[3, 6; 3, 9)	[3, 9; 4, 2)
Giá trị đại diện	2,85	3,15	3,45	3,75	4,05
Số ngày	3	6	5	4	2

+ Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x} = \frac{2,85.3 + 3,15.6 + 3,45.5 + 3,75.4 + 4,05.2}{20} = 3,39.$$

+ Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{20} (2,85^2.3 + 3,15^2.6 + 3,45^2.5 + 3,75^2.4 + 4,05^2.2) - 3,39^2 \approx 0,13.$$

Câu 28. Bạn Chi rất thích nhảy hiện đại. Thời gian tập nhảy mỗi ngày trong thời gian gần đây của bạn Chi được thống kê lại ở bảng sau:

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

Phương sai của mẫu số liệu ghép nhóm là (làm tròn đến hàng phần trăm)

A. 31,77.

B. 31,25.

C. 31,44.

D. 32,25.

Lời giải

Chọn B

+ Cỡ mẫu: $n = 18$.

Thời gian (phút)	[20; 25)	[25; 30)	[30; 35)	[35; 40)	[40; 45)
Giá trị đại diện	22,5	27,5	32,5	37,5	42,5
Số ngày	6	6	4	1	1

+ Số trung bình của mẫu số liệu ghép nhóm là: $\bar{x} = \frac{22,5.6 + 27,5.6 + 32,5.4 + 37,5.1 + 42,5.1}{18} = \frac{85}{3}.$

+ Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{18} (22,5^2.6 + 27,5^2.6 + 32,5^2.4 + 37,5^2.1 + 42,5^2.1) - \left(\frac{85}{3}\right)^2 = 31,25.$$

Câu 29. Dũng là học sinh rất giỏi chơi rubik, bạn có thể giải nhiều loại khối rubik khác nhau. Trong một lần tập luyện giải khối rubik 3×3 , bạn Dũng đã tự thống kê lại thời gian giải rubik trong 25 lần giải liên tiếp ở bảng sau::

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

Độ lệch chuẩn của mẫu số liệu ghép nhóm là (làm tròn đến hàng phần trăm)

A. 5,98.

B. 6.

C. 2,44.

D. 2,5.

Lời giải

Chọn C

+ Cỡ mẫu: $n = 25$.

Thời gian giải rubik (giây)	[8;10)	[10;12)	[12;14)	[14;16)	[16;18)
Giá trị đại diện	9	11	13	15	17
Số lần	4	6	8	4	3

+ Số trung bình của mẫu số liệu ghép nhóm là: $\bar{x} = \frac{9.4 + 11.6 + 13.8 + 15.4 + 17.3}{25} = 12,68$.

+ Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{25} (9^2.4 + 11^2.6 + 13^2.8 + 15^2.4 + 17^2.3) - 12,68^2 = \frac{3736}{625}$$

+ Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S = \sqrt{\frac{3736}{625}} \approx 2,44$.

Câu 30. Một bác tài xế thống kê lại độ dài quãng đường (đơn vị: km) bác đã lái xe mỗi ngày trong một tháng ở bảng sau:

Độ dài quãng đường (km)	[50;100)	[100;150)	[150;200)	[200;250)	[250;300)
Số ngày	5	10	9	4	2

Độ lệch chuẩn của mẫu số liệu ghép nhóm là (làm tròn đến hàng phần trăm)

A. 55,68.

B. 56,67.

C. 3100.

D. 3000.

Lời giải

Chọn A

+ Cỡ mẫu: $n = 30$.

Độ dài quãng đường (km)	[50;100)	[100;150)	[150;200)	[200;250)	[250;300)
Giá trị đại diện	75	125	175	225	275
Số ngày	5	10	9	4	2

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

$$\bar{x} = \frac{75.5 + 125.10 + 175.9 + 225.4 + 275.2}{30} = 155$$

+ Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{30} (75^2.5 + 125^2.10 + 175^2.9 + 225^2.4 + 275^2.2) - 155^2 = 3100$$

+ Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S = \sqrt{3100} \approx 55,68$.

Câu 31. Người ta ghi lại tiền lãi (đơn vị: triệu đồng) của một số nhà đầu tư (với số tiền đầu tư như nhau), khi đầu tư vào hai lĩnh vực A, B cho kết quả như sau

Tiền lãi	[5; 10)	[10; 15)	[15; 20)	[20; 25)	[25; 30)
Số nhà đầu tư vào lĩnh vực A	2	5	8	6	4
Số nhà đầu tư vào lĩnh vực B	8	4	2	5	6

Người ta có thể dùng phương sai và độ lệch chuẩn để so sánh mức độ rủi ro đầu tư các lĩnh vực có giá trị trung bình tiền lãi gần bằng nhau. Lĩnh vực nào có phương sai, độ lệch chuẩn tiền lãi cao hơn thì được coi là có độ rủi ro lớn hơn.

Theo quan điểm trên, độ rủi ro của cổ phiếu nào cao hơn?

A. Lĩnh vực A có độ rủi ro bằng lĩnh vực B.

- B. Lĩnh vực A có độ rủi ro cao hơn lĩnh vực B .
 C. Lĩnh vực A có độ rủi ro thấp hơn lĩnh vực B .
 D. Không so sánh được.

Lời giải

Chọn C

Lĩnh vực A

Tiền lãi	[5; 10)	[10; 15)	[15; 20)	[20; 25)	[25; 30)
Giá trị đại diện	7,5	12,5	17,5	22,5	27,5
Số nhà đầu tư vào lĩnh vực A	2	5	8	6	4

Lĩnh vực B

Tiền lãi	[5; 10)	[10; 15)	[15; 20)	[20; 25)	[25; 30)
Giá trị đại diện	7,5	12,5	17,5	22,5	27,5
Số nhà đầu tư vào lĩnh vực B	8	4	2	5	6

Giá trị trung bình của hai lĩnh vực A và B là

$$\bar{x}_A = \frac{1}{25} \cdot (2 \cdot 7,5 + 5 \cdot 12,5 + 8 \cdot 17,5 + 6 \cdot 22,5 + 4 \cdot 27,5) = 18,5$$

$$\bar{x}_B = \frac{1}{25} \cdot (8 \cdot 7,5 + 4 \cdot 12,5 + 2 \cdot 17,5 + 5 \cdot 22,5 + 6 \cdot 27,5) = 16,9$$

Về độ trung bình đầu tư vào lĩnh vực A lãi hơn lĩnh vực B .

Độ lệch chuẩn của hai lĩnh vực A và B là

$$s_A = \sqrt{\frac{1}{25} \cdot (2 \cdot 7,5^2 + 5 \cdot 12,5^2 + 8 \cdot 17,5^2 + 6 \cdot 22,5^2 + 4 \cdot 27,5^2) - 18,5^2} = 5,8$$

$$s_B = \sqrt{\frac{1}{25} \cdot (8 \cdot 7,5^2 + 4 \cdot 12,5^2 + 2 \cdot 17,5^2 + 5 \cdot 22,5^2 + 6 \cdot 27,5^2) - 16,9^2} = 8,04.$$

Như vậy độ lệch chuẩn của mẫu số liệu thu tiền được hàng tháng khi đầu tư vào lĩnh vực B cao hơn lĩnh vực A nên đầu tư vào lĩnh vực B rủi ro hơn.

Câu 32. Giá đóng cửa của một cổ phiếu là giá của cổ phiếu đó cuối một phiên giao dịch. Bảng sau thống kê giá đóng cửa (đơn vị: nghìn đồng) của hai mã cổ phiếu A và B trong 50 ngày giao dịch liên tiếp.

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

Người ta có thể dùng phương sai và độ lệch chuẩn để so sánh mức độ rủi ro của các loại cổ phiếu có giá trị trung bình gần bằng nhau. Cổ phiếu nào có phương sai, độ lệch chuẩn cao hơn thì được coi là có độ rủi ro lớn hơn.

Theo quan điểm trên, độ rủi ro của cổ phiếu nào cao hơn?

- A. Cổ phiếu A có độ rủi ro bằng cổ phiếu B .
 B. Cổ phiếu A có độ rủi ro cao hơn cổ phiếu B .
 C. Cổ phiếu A có độ rủi ro thấp hơn cổ phiếu B .
 D. Không so sánh được.

Lời giải

Chọn C

Ta có bảng thống kê giá đóng cửa theo giá trị đại diện

Giá đóng cửa	121	123	125	127	129
Cổ phiếu A	8	9	12	10	11
Cổ phiếu B	16	4	3	6	21

- Xét mẫu số liệu của cổ phiếu A

$$\text{Số trung bình của mẫu số liệu ghép nhóm là: } \bar{x}_1 = \frac{8.121 + 9.123 + 12.125 + 10.127 + 11.129}{50} = 125,28.$$

Phương sai của mẫu số liệu ghép nhóm là

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

$$\text{Độ lệch chuẩn của mẫu số liệu ghép nhóm là } S_1 = \sqrt{7,5216}$$

- Xét mẫu số liệu của cổ phiếu B

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_2 = \frac{16.121 + 4.123 + 3.125 + 6.127 + 21.129}{50} = 125,48.$$

Phương sai của mẫu số liệu ghép nhóm là

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

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

Vậy nếu đánh giá độ rủi ro theo phương sai và độ lệch chuẩn thì cổ phiếu A có độ rủi ro thấp hơn cổ phiếu B.

Câu 33. Kết quả 40 lần nhảy xa của hai vận động viên nam Dũng được lần lượt thống kê trong các bảng sau (đơn vị: mét)

Nhóm	Tần số
[6, 22; 6, 46)	3
[6, 46; 6, 70)	7
[6, 70; 6, 94)	5
[6, 94; 7, 18)	20
[7, 18; 7, 42)	5
	$n = 40$

Giá trị trung bình của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Dũng có giá trị gần với giá trị

A. 6,92.

B. 7,34.

C. 9,86

D. 5,85.

Lời giải

Chọn A

Ta có bảng thống kê

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
		$n = 40$

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 thứ nhất là

$$\bar{x}_D = \frac{3.6,34 + 7.6,58 + 5.6,82 + 20.7,06 + 5.7,30}{40} = \frac{276,88}{40} \approx 6,92(\text{m}).$$



Câu 34. Một siêu thị thống kê số tiền (đơn vị: chục nghìn đồng) mà 44 khách hàng mua hàng ở siêu thị đó trong một ngày. Số liệu được ghi lại trong bảng sau.

Nhóm	Giá trị đại diện	Tần số
[40; 45)	42,5	4
[45; 50)	47,5	14
[50; 55)	52,5	8
[55; 60)	57,5	10
[60; 75)	62,5	6
[65; 70)	67,5	2
		$n = 44$

Giá trị trung bình của mẫu số liệu ghép nhóm trên có giá trị gần nhất với giá trị nào dưới đây

A. 53,2.

B. 57,34.

C. 49,86

D. 65,85.

Lời giải

Chọn A

$$\text{Số trung bình: } \bar{x} = \frac{42,5 \cdot 4 + 47,5 \cdot 14 + 52,5 \cdot 8 + 57,5 \cdot 10 + 62,5 \cdot 6 + 67,5 \cdot 2}{44} = \frac{585}{11} \approx 53,2$$

Câu 35. Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khỏe. Quãng đường đi bộ mỗi ngày (đơn vị km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau

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

Phương sai của mẫu số liệu ghép nhóm là

A. 3,39.

B. 11,62.

C. 0,1314.

D. 0,36.

Lời giải

Chọn C

Xét mẫu số liệu ghép nhóm cho bởi bảng sau

Nhóm	[2,7; 3,0)	[3,0; 3,3)	[3,3; 3,6)	[3,6; 3,9)	[3,9; 4,2)
Giá trị đại diện	2,85	3,15	3,45	3,75	4,05
Tần số	3	6	5	4	2

Số trung bình của mẫu số liệu là

$$\bar{x} = \frac{1}{20} \cdot (2,85 \cdot 3 + 3,15 \cdot 6 + 3,45 \cdot 5 + 3,75 \cdot 4 + 4,05 \cdot 2) = 3,39.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{20} (3 \cdot 2,85^2 + 6 \cdot 3,15^2 + 5 \cdot 3,45^2 + 4 \cdot 3,75^2 + 2 \cdot 4,05^2) - 3,39^2 = 0,1314.$$

Câu 36. Bạn Chi rất thích nhảy hiện đại. Thời gian tập nhảy mỗi ngày trong thời gian gần đây của bạn Chi được thống kê lại ở bảng sau

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

Phương sai của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

A. 31,77.

B. 32.

C. 31.

D. 31,44.

Lời giải

Chọn D

Xét mẫu số liệu ghép nhóm cho bởi bảng sau

Nhóm	[20; 25)	[25; 30)	[30; 35)	[35; 40)	[40; 45)
Giá trị đại diện	22,5	27,5	32,5	37,5	42,5
Tần số	6	6	4	1	1

Số trung bình của mẫu số liệu là: $\bar{x} = \frac{1}{18} \cdot (22,5 \cdot 6 + 27,5 \cdot 6 + 32,5 \cdot 4 + 37,5 \cdot 1 + 42,5 \cdot 1) = \frac{85}{3}$.

Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{18} (6 \cdot 22,5^2 + 6 \cdot 27,5^2 + 4 \cdot 32,5^2 + 1 \cdot 37,5^2 + 1 \cdot 42,5^2) - \left(\frac{85}{3} \right)^2 = 31,25.$$

Vậy phương sai của mẫu số liệu ghép nhóm gần nhất với 31,44.

Câu 37. Một mẫu số liệu có bảng tần số ghép nhóm như sau:

Nhóm	[1; 5)	[5; 9)	[9; 13)	[13; 17)	[17; 21)
Tần số	4	8	13	6	4

Phương sai của mẫu số liệu là (kết quả làm tròn đến hàng phần trăm)

- A. 21,01. B. 20,01. C. 22. D. 23.

Lời giải

Chọn A

Xét mẫu số liệu ghép nhóm cho bởi bảng sau

Nhóm	[1; 5)	[5; 9)	[9; 13)	[13; 17)	[17; 21)
Giá trị đại diện	3	7	11	15	19
Tần số	4	8	13	6	4

Số trung bình của mẫu số liệu là: $\bar{x} = \frac{1}{35} \cdot (3 \cdot 4 + 7 \cdot 8 + 11 \cdot 13 + 15 \cdot 6 + 19 \cdot 4) \approx 10,77$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{35} (4 \cdot 3^2 + 8 \cdot 7^2 + 13 \cdot 11^2 + 6 \cdot 15^2 + 4 \cdot 19^2) - 10,77^2 \approx 21,01$$

Câu 38. Khảo sát thời gian tự học bài ở nhà của học sinh khối 9 ở trường X, ta thu được bảng sau:

Thời gian(phút)	[0; 30)	[30; 60)	[60; 90)	[90; 120)	[120; 150)
Số học sinh	9	10	9	15	7

Phương sai của mẫu số liệu ghép nhóm là

- A. 1602. B. 1601,64. C. 1601,9. D. 1603.

Lời giải

Chọn B

Thời gian(phút)	[0; 30)	[30; 60)	[60; 90)	[90; 120)	[120; 150)
Giá trị đại diện	15	45	75	105	135
Số học sinh	9	10	9	15	7

Thời gian trung bình tự học ở nhà của các em học sinh đó là:

$$\bar{x} = \frac{9 \cdot 15 + 10 \cdot 45 + 9 \cdot 75 + 15 \cdot 105 + 7 \cdot 135}{50} = 75,6 \text{ (phút)}.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{50} (9 \cdot 15^2 + 10 \cdot 45^2 + 9 \cdot 75^2 + 15 \cdot 105^2 + 7 \cdot 135^2) - 75,6^2 = 1601,64$$

Câu 39. Thời gian truy cập Internet mỗi buổi tối của một số học sinh được cho trong bảng sau:

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

Độ lệch chuẩn của mẫu số liệu là (kết quả làm tròn đến hàng phần trăm)

- A. 2,93. B. 8,56. C. 8,59. D. 3,01.

🔗 Lời giải

Chọn A

Thời gian (phút)	[9,5;12,5)	[12,5;15,5)	[15,5;18,5)	[18,5;21,5)	[21,5;24,5)
Giá trị đại diện	11	14	17	20	23
Số học sinh	3	12	15	24	2

Số trung bình của mẫu số liệu là

$$\bar{x} = \frac{1}{56} \cdot (3 \cdot 11 + 12 \cdot 14 + 15 \cdot 17 + 24 \cdot 20 + 2 \cdot 23) \approx 17,54$$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{56} (3 \cdot 11^2 + 12 \cdot 14^2 + 15 \cdot 17^2 + 24 \cdot 20^2 + 2 \cdot 23^2) - 17,54^2 \approx 8,56$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S = \sqrt{8,56} \approx 2,93$

Câu 40. Một câu lạc bộ thể dục thể thao đã ghi lại số giờ các thành viên của mình sử dụng cơ sở vật chất của câu lạc bộ để tập luyện trong một tháng như sau:

Thời gian (giờ)	[1;5)	[5;9)	[9;13)	[13;17)	[17;21)	[21;25)
Tần số (Số người)	10	14	31	2	5	23

Độ lệch chuẩn của mẫu số liệu là (kết quả làm tròn đến hàng phần trăm)

A. 6,9.

B. 9,6.

C. 6,96.

D. 7,96.

🔗 Lời giải

Chọn B

Thời gian (giờ)	[1;5)	[5;9)	[9;13)	[13;17)	[17;21)	[21;25)
Giá trị đại diện	3	7	11	15	19	23
Tần số (Số người)	10	14	31	2	5	23

Số trung bình của mẫu số liệu là: $\bar{x} = \frac{1}{85} \cdot (10 \cdot 3 + 14 \cdot 7 + 31 \cdot 11 + 2 \cdot 15 + 5 \cdot 19 + 23 \cdot 23) \approx 13,21$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{85} \cdot (10 \cdot 3^2 + 14 \cdot 7^2 + 31 \cdot 11^2 + 2 \cdot 15^2 + 5 \cdot 19^2 + 23 \cdot 23^2) - 13,21^2 \approx 48,43$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S = \sqrt{48,43} \approx 6,96$

Câu 41. Thống kê tổng số giờ nắng trong tháng 9 tại một trạm quan trắc đặt ở Cà Mau trong các năm từ 2002 đến 2021 được thống kê như sau:

Số giờ nắng	[80;98)	[98;116)	[116;134)	[134;152)	[152;170)
Số năm	3	6	3	5	3

Độ lệch chuẩn của mẫu số liệu là (kết quả làm tròn đến hàng phần nghìn)

A. 23,795.

B. 24,795.

C. 23,794.

D. 23,796.

🔗 Lời giải

Chọn A

Số giờ nắng	[80;98)	[98;116)	[116;134)	[134;152)	[152;170)
Giá trị đại diện	89	107	125	143	161
Số năm	3	6	3	5	3

Số trung bình của mẫu số liệu là

$$\bar{x} = \frac{1}{20} \cdot (3 \cdot 89 + 6 \cdot 107 + 3 \cdot 125 + 5 \cdot 143 + 3 \cdot 161) = 124,1$$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2 = \frac{1}{20} \cdot (3 \cdot 89^2 + 6 \cdot 107^2 + 3 \cdot 125^2 + 5 \cdot 143^2 + 3 \cdot 161^2) - 124,1^2 = 566,19$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S = \sqrt{566,19} \approx 23,795$

Câu 42. 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)
Số ngày giao dịch của cổ phiếu A	8	9	12	10	11
Số ngày giao dịch của cổ phiếu B	16	4	3	6	21

Người ta có thể dùng phương sai và độ lệch chuẩn để so sánh mức độ rủi ro của các loại cổ phiếu có giá trị trung bình gần bằng nhau. Cổ phiếu nào có phương sai, độ lệch chuẩn cao hơn thì được coi là có độ rủi ro lớn hơn. Chọn khẳng định đúng.

- A.** Giá đóng cửa của cổ phiếu A ít phân tán hơn giá đóng cửa của cổ phiếu B .
B. Giá đóng cửa của cổ phiếu B ít phân tán hơn giá đóng cửa của cổ phiếu A .
C. Giá đóng cửa của hai cổ phiếu A và B có độ phân tán như nhau.
D. Phương sai của mẫu số liệu của cổ phiếu A lớn hơn phương sai của mẫu số liệu của cổ phiếu B .

Lời giải

Chọn A

Ta có bảng thống kê theo giá trị đại diện

Giá đóng cửa	121	123	125	127	129
Số ngày giao dịch của cổ phiếu A	8	9	12	10	11
Số ngày giao dịch của cổ phiếu B	16	4	3	6	21

Xét mẫu số liệu của cổ phiếu A

Số trung bình của mẫu số liệu là

$$\bar{x}_A = \frac{1}{50} \cdot (8 \cdot 121 + 9 \cdot 123 + 12 \cdot 125 + 10 \cdot 127 + 11 \cdot 129) = 125,28$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_A^2 = \frac{1}{50} \cdot (8 \cdot 121^2 + 9 \cdot 123^2 + 12 \cdot 125^2 + 10 \cdot 127^2 + 11 \cdot 129^2) - 125,28^2 = 7,5216$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là

$$S_A = \sqrt{7,5216}$$

Xét mẫu số liệu của cổ phiếu B

Số trung bình của mẫu số liệu là

$$\bar{x}_B = \frac{1}{50} \cdot (16 \cdot 121 + 4 \cdot 123 + 3 \cdot 125 + 6 \cdot 127 + 21 \cdot 129) = 125,48$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_B^2 = \frac{1}{50} \cdot (16 \cdot 121^2 + 4 \cdot 123^2 + 3 \cdot 125^2 + 6 \cdot 127^2 + 21 \cdot 129^2) - 125,48^2 = 12,4096$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là

$$S_B = \sqrt{12,4096}$$

Ta có $S_A < S_B$ nên giá đóng cửa của cổ phiếu A ít phân tán hơn giá đóng cửa của cổ phiếu B .

Câu 43. Trong 30 ngày, một nhà đầu tư đã theo dõi giá cổ phiếu của hai công ty G và H vào phiên mở cửa mỗi ngày. Thông tin được ghi lại ở hai bảng dưới đây:

Giá cổ phiếu của công ty G

Giá (nghìn đồng)	[50; 52)	[52; 54)	[54; 56)	[56; 58)	[58; 60)
Tần số	3	7	9	8	3

Giá cổ phiếu của công ty H

Giá (nghìn đồng)	[40; 42)	[42; 44)	[44; 46)	[46; 48)	[48; 50)
Tần số	6	7	5	7	5

Chọn câu trả lời đúng nhất biết độ lệch chuẩn càng cao thì tỷ lệ rủi ro càng lớn:

- A. Công ty G rủi ro hơn.
 B. Công ty H rủi ro hơn.
 C. Cả hai đều rủi ro như nhau.
 D. Cả hai công ty đều không rủi ro.

Lời giải

Chọn B

Công ty G :

Bổ sung thêm các giá trị đại diện, ta có bảng sau

Giá (nghìn đồng)	[50; 52)	[52; 54)	[54; 56)	[56; 58)	[58; 60)	
Giá trị đại diện	51	53	55	57	59	
Tần số	3	7	9	8	3	$N = 30$

Giá trị trung bình của mẫu số liệu là

$$\bar{x} = \frac{51 \cdot 3 + 53 \cdot 7 + 55 \cdot 9 + 57 \cdot 8 + 59 \cdot 3}{30} \approx 55,1.$$

Trung bình cộng của các bình phương số liệu thống kê là

$$\overline{x^2} = \frac{51^2 \cdot 3 + 53^2 \cdot 7 + 55^2 \cdot 9 + 57^2 \cdot 8 + 59^2 \cdot 3}{30} \approx 3037,5.$$

Từ đó ta có độ lệch chuẩn của mẫu số liệu là $s = \sqrt{\overline{x^2} - (\bar{x})^2} \approx \sqrt{5,2} \approx 2,3$.

Công ty H

Giá (nghìn đồng)	[40; 42)	[42; 44)	[44; 46)	[46; 48)	[48; 50)	
Giá trị đại diện	41	43	45	47	49	
Tần số	6	7	5	7	5	$N = 30$

Bổ sung thêm các giá trị đại diện, ta có bảng sau

Giá trị trung bình của mẫu số liệu là

$$\bar{x} = \frac{41 \cdot 6 + 43 \cdot 7 + 45 \cdot 5 + 47 \cdot 7 + 49 \cdot 5}{30} \approx 44,9.$$

Trung bình cộng của các bình phương số liệu thống kê là

$$\overline{x^2} = \frac{41^2 \cdot 6 + 43^2 \cdot 7 + 45^2 \cdot 5 + 47^2 \cdot 7 + 49^2 \cdot 5}{30} \approx 2020,7.$$

Từ đó ta có độ lệch chuẩn của mẫu số liệu là $s = \sqrt{\overline{x^2} - (\bar{x})^2} \approx \sqrt{7,7} \approx 2,8$.

Từ kết quả trên, ta thấy công ty H rủi ro hơn

Câu 44. Anh An đầu tư số tiền bằng nhau vào hai lĩnh vực kinh doanh A, B . Anh An thống kê số tiền thu được mỗi tháng trong vòng 60 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	5	10	30	10	5
Số tháng đầu tư vào lĩnh vực B	20	5	10	5	20

Đáp án nào sau đây đúng?

- A. Đầu tư ở lĩnh vực A rủi ro hơn.
 B. Đầu tư ở lĩnh vực B rủi ro hơn.
 C. Độ lệch chuẩn ở lĩnh vực A lớn hơn 10.
 D. Đầu tư ở hai lĩnh vực A và B rủi ro như nhau.

Lời giải

Chọn B

Ta có

Giá trị đại diện	7,5	12,5	17,5	22,5	27,5
Số tháng đầu tư vào lĩnh vực A	5	10	30	10	5
Số tháng đầu tư vào lĩnh vực B	20	5	10	5	20

Số tiền trung bình thu được khi đầu tư vào các lĩnh vực A, B tương ứng là:

$$\bar{x}_A = \frac{1}{60} (5 \cdot 7,5 + \dots + 5 \cdot 27,5) = 17,5;$$

$$\bar{x}_B = \frac{1}{60} (20 \cdot 7,5 + \dots + 20 \cdot 27,5) = 17,5$$

Độ lệch chuẩn của số tiền thu được hàng tháng khi đầu tư vào các lĩnh vực A, B tương ứng là

$$s_A = \sqrt{\frac{1}{60} (5 \cdot 7,5^2 + \dots + 5 \cdot 27,5^2) - (17,5)^2} \approx 5;$$

$$s_B = \sqrt{\frac{1}{60} (20 \cdot 7,5^2 + \dots + 20 \cdot 27,5^2) - (17,5)^2} \approx 8$$

Như vậy, về trung bình đầu tư vào các lĩnh vực A, B số tiền thu được hàng tháng như nhau tuy nhiên độ lệch chuẩn của mẫu số liệu về số tiền thu được hàng tháng khi đầu tư vào lĩnh vực B cao hơn khi đầu tư vào lĩnh vực A. Người ta nói rằng, đầu tư vào lĩnh vực B là "rủi ro" hơn.

Câu 45. Một vận động viên luyện tập chạy cự li 100 m đã ghi lại kết quả luyện tập như sau.

Thời gian (giây)	[10,2; 10,4)	[10,4; 10,6)	[10,6; 10,8)	[10,8; 11)
Số vận động viên	3	7	8	2

Tìm phương sai của mẫu số liệu ghép nhóm (làm tròn kết quả đến chữ số thập phân thứ 2)

A. 1.03 .

B. 0.03 .

C. 2,90 .

D. 1.86 .

Lời giải

Chọn B

Ta có

Thời gian (giây)	[10,2; 10,4)	[10,4; 10,6)	[10,6; 10,8)	[10,8; 11)
Giá trị đại diện	10,3	10,5	10,7	10,9
Số vận động viên	3	7	8	2

Thời gian trung bình là

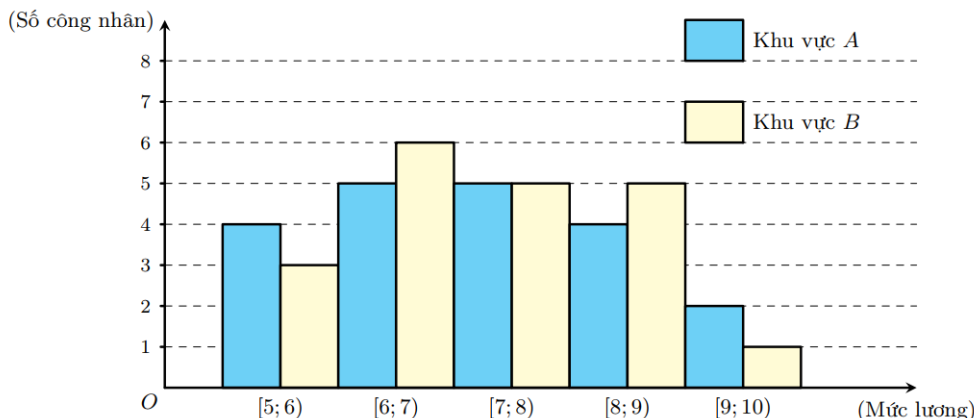
$$\bar{x} = \frac{1}{20} (10,3 \cdot 3 + 10,5 \cdot 7 + 10,7 \cdot 8 + 10,9 \cdot 2) = 10,59.$$

Phương sai

$$s^2 = \frac{1}{20} \cdot (10,3^2 \cdot 3 + 10,5^2 \cdot 7 + 10,7^2 \cdot 8 + 10,9^2 \cdot 2) - 10,59^2 = 0.03.$$

Câu 46. 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.

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



Tổng độ lệch chuẩn của mẫu số liệu ghép nhóm ở 2 khu vực gần bằng với số nào sau đây nhất.

A. 2.

B. 3.

C. 5.

D. 1.

~~Lời giải~~

Chọn A

Ta có

Mức lương	[5; 6)	[6; 7)	[7; 8)	[8; 9)	[9; 10)
Mức lương đại diện	5,5	6,5	7,5	8,5	9,5
Khu vực A	4	5	5	4	2
Khu vực B	3	6	5	5	1

Xét mẫu số liệu của khu vực A

Cỡ mẫu là $n_A = 4 + 5 + 5 + 4 + 2 = 20$.

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_A = \frac{4 \cdot 5,5 + 5 \cdot 6,5 + 5 \cdot 7,5 + 4 \cdot 8,5 + 2 \cdot 9,5}{20} = 7,25.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_A^2 = \frac{1}{20} (4 \cdot 5,5^2 + 5 \cdot 6,5^2 + 5 \cdot 7,5^2 + 4 \cdot 8,5^2 + 2 \cdot 9,5^2) - 7,25^2 = 1,5875.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_A = \sqrt{1,5875} \approx 1,2300$.

Xét mẫu số liệu của khu vực B

Cỡ mẫu là $n_B = 3 + 6 + 5 + 5 + 1 = 20$.

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_B = \frac{3 \cdot 5,5 + 6 \cdot 6,5 + 5 \cdot 7,5 + 5 \cdot 8,5 + 1 \cdot 9,5}{20} = 7,25.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_B^2 = \frac{1}{20} (3 \cdot 5,5^2 + 6 \cdot 6,5^2 + 5 \cdot 7,5^2 + 5 \cdot 8,5^2 + 1 \cdot 9,5^2) - 7,25^2 = 1,2875.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_B = \sqrt{1,2875} \approx 1,1347$.

Tổng: khoảng 2,3647.

Câu 47. Bộ phận kiểm tra chất lượng sản phẩm dùng máy để đo (chính xác đến 0,001mm) độ dày của một chi tiết máy. Kết quả đo một số sản phẩm được thống kê trong bảng sau:

Độ dày của chi tiết máy

Độ dày (mm)	[18; 19)	[19; 20)	[20; 21)	[21; 22)	[22; 23)
Tần số	3	7	23	25	2

Nhận xét nào sau đây **sai**?

A. Độ lệch chuẩn của mẫu lớn hơn 2.

- B.** Số trung bình của mẫu số liệu gần bằng với 20,77.
C. Độ dày của chi tiết máy không bị sai lệch nhiều.
D. Cỡ mẫu của mẫu số liệu là 60.

Lời giải

Chọn A

Ta có cỡ mẫu $n = 60$.

Số trung bình của mẫu số liệu là

$$\bar{x} = \frac{3 \cdot 18,5 + 7 \cdot 19,5 + 23 \cdot 20,5 + 25 \cdot 21,5 + 2 \cdot 22,5}{60} = \frac{623}{30} \approx 20,77.$$

Phương sai của mẫu số liệu là

$$S^2 = \frac{1}{60} (3 \cdot 18,5^2 + 7 \cdot 19,5^2 + 23 \cdot 20,5^2 + 25 \cdot 21,5^2 + 2 \cdot 22,5^2) - \left(\frac{623}{30} \right)^2 = \frac{179}{225}.$$

$$\text{Độ lệch chuẩn của mẫu số liệu là } S = \sqrt{\frac{179}{225}} = \frac{\sqrt{179}}{15} \approx 0,89.$$

Câu 48. Thống kê lợi nhuận hàng tháng (đơn vị: triệu đồng) trong 20 tháng của hai nhà đầu tư được cho như sau:

Lợi nhuận	[10; 20)	[20; 30)	[30; 40)	[40; 50)	[50; 60)
Số tháng	2	4	8	4	2

Bảng 1. Lợi nhuận theo tháng của nhà đầu tư nhỏ

Lợi nhuận	[510; 520)	[520; 530)	[530; 540)	[540; 550)	[550; 560)
Số tháng	4	3	6	3	4

Bảng 2. Lợi nhuận theo tháng của nhà đầu tư lớn

Đáp án nào dưới đây đúng nhất?

- A.** Không nên chỉ dựa vào độ lệch chuẩn để đánh giá mức độ rủi ro.
B. Lợi nhuận trung bình mỗi tháng của nhà đầu tư nhỏ cao hơn nhà đầu tư lớn.
C. Độ lệch chuẩn của nhà đầu tư lớn cao hơn 14.
D. Nhà đầu tư lớn có mức độ rủi ro cao hơn.

Lời giải

Chọn A

Chọn điểm đại diện cho các nhóm số liệu ta tính được các số đặc trưng như sau Lợi nhuận trung bình một tháng của các nhà đầu tư tương ứng là

$$\bar{x}_A = \frac{1}{20} (2 \cdot 15 + \dots + 2 \cdot 55) = 35$$

$$\bar{x}_B = \frac{1}{20} (4 \cdot 515 + \dots + 4 \cdot 555) = 535$$

Độ lệch chuẩn của lợi nhuận hàng tháng của hai nhà đầu tư tương ứng là

$$s_A = \sqrt{\frac{1}{20} (2 \cdot 15^2 + \dots + 2 \cdot 55^2) - (35)^2} \approx 10,95$$

$$s_B = \sqrt{\frac{1}{20} (4 \cdot 515^2 + \dots + 4 \cdot 555^2) - (535)^2} \approx 13,78.$$

Độ lệch chuẩn cho lợi nhuận hàng tháng của nhà đầu tư lớn cao hơn của nhà đầu tư nhỏ. Lợi nhuận trung bình của hai nhà đầu tư khác nhau rất nhiều, do đó ta không nên dùng độ lệch chuẩn để so sánh mức độ rủi ro của hai nhà đầu tư này

✓ TRẮC NGHIỆM ĐÚNG SAI

Câu 49. Thời gian chờ khám bệnh của hai phòng khám 1 và phòng khám 2 ở thành phố X được cho trong bảng sau:

Thời gian (phút)	[0;5)	[5;10)	[10;15)	[15;20)
Phòng khám số 1 (Số bệnh nhân)	3	12	15	18
Phòng khám số 2 (Số bệnh nhân)	5	10	12	0

Xét tính đúng/sai các mệnh đề sau:

(a) Tổng số bệnh nhân chờ khám bệnh ở phòng khám số 1 dưới 5 phút là 3.

(b) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian chờ khám bệnh của phòng khám số 1 là $R_1 = 15$.

(c) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian chờ khám bệnh của phòng khám số 2 là $R_2 = 20$.

(d) Thời gian chờ khám bệnh ở phòng khám số 2 phân tán hơn thời gian chờ khám bệnh ở phòng khám số 1.

Lời giải

(a) Tổng số bệnh nhân chờ khám bệnh ở phòng khám số 1 dưới 5 phút là 3.

Chọn ĐÚNG.

(b) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian chờ khám bệnh của phòng khám số 1 là $R_1 = 15$.

Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian chờ khám bệnh của phòng khám số 1 là $R_1 = 20 - 0 = 20$.

Chọn SAI.

(c) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian chờ khám bệnh của phòng khám số 2 là $R_2 = 20$.

Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian chờ khám bệnh của phòng khám số 2 là $R_2 = 15 - 0 = 15$.

Chọn SAI.

(d) Thời gian chờ khám bệnh ở phòng khám số 2 phân tán hơn thời gian chờ khám bệnh ở phòng khám số 1.

Vì $R_1 > R_2$ nên thời gian khám bệnh ở phòng khám số 1 phân tán hơn thời gian chờ khám bệnh ở phòng khám số 2.

Chọn SAI.

Câu 50. Thống kê thu nhập theo tháng (đơn vị: triệu đồng) của một nhóm người chạy Grab được cho trong bảng sau:

Thu nhập (triệu đồng)	[3;5)	[5;7)	[7;9)	[9;11)
Số người	5	10	5	2

Xét tính đúng/sai các mệnh đề sau:

(a) Cỡ mẫu là $n = 22$

(b) Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là $Q_1 = 10$.

(c) Tứ phân vị thứ ba của mẫu số liệu ghép nhóm là $Q_3 = 5$.

(d) Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = 5$.

Lời giải

(a) Cỡ mẫu là $n = 22$

Chọn ĐÚNG.

(b) Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là $Q_1 = 10$.

Gọi $x_1, \dots, x_{22}$ là thu nhập của 22 người đã được sắp xếp theo thứ tự tăng dần.

Tứ phân vị thứ nhất của mẫu số liệu gốc là x_6 nên nhóm chứa tứ phân vị thứ nhất là nhóm $[5; 7)$.

$$\text{Do đó } Q_1 = 5 + \frac{\frac{1.22}{4} - 5}{10} \cdot 2 = 5,1.$$

Chọn SAI.

(c) Tứ phân vị thứ ba của mẫu số liệu ghép nhóm là $Q_3 = 5$.

Tứ phân vị thứ ba của mẫu số liệu gốc là x_{17} nên nhóm chứa tứ phân vị thứ nhất là nhóm $[7; 9)$.

$$\text{Do đó } Q_3 = 7 + \frac{\frac{3.22}{4} - 15}{5} \cdot 2 = 7,6.$$

Chọn SAI.

(d) Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = 5$.

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = 7,6 - 5,1 = 2,5$.

Chọn SAI.

Câu 51. Người ta theo dõi sự thay đổi cân nặng, được tính bằng hiệu cân nặng trước và sau ba tháng áp dụng chế độ ăn kiêng của một số người cho kết quả sau:

Thay đổi cân nặng (kg)	$[-1; 0)$	$[0; 1)$	$[1; 2)$	$[2; 3)$	$[3; 4)$
Số người nam	6	4	2	3	1
Số người nữ	5	6	3	1	0

Xét tính đúng/sai các mệnh đề sau:

(a) Số người thay đổi cân nặng theo chiều hướng giảm cân là 11.

(b) Khoảng biến thiên của mẫu số liệu ghép nhóm về sự thay đổi cân nặng của nam là $R_1 = 5$.

(c) Khoảng biến thiên của mẫu số liệu ghép nhóm về sự thay đổi cân nặng của nữ là $R_2 = 3$.

(d) Nếu biết nữ tăng cân nhiều nhất là 2 kg và giảm cân được nhiều nhất 1 kg thì khoảng biến thiên của mẫu số liệu gốc là 1.

Lời giải

(a) Số người thay đổi cân nặng theo chiều hướng giảm cân là 11.

Số người thay đổi theo chiều hướng giảm cân là $5 + 6 = 11$.

Chọn ĐÚNG.

(b) Khoảng biến thiên của mẫu số liệu ghép nhóm về sự thay đổi cân nặng của nam là $R_1 = 5$.

Khoảng biến thiên của mẫu số liệu ghép nhóm về sự thay đổi cân nặng của nam là $R_1 = 4 - (-1) = 5$.

Chọn ĐÚNG.

(c) Khoảng biến thiên của mẫu số liệu ghép nhóm về sự thay đổi cân nặng của nữ là $R_2 = 3$.

Khoảng biến thiên của mẫu số liệu ghép nhóm về sự thay đổi cân nặng của nữ là $R_2 = 3 - (-1) = 4$.

Chọn SAI.

(d) Nếu biết nữ tăng cân nhiều nhất là 2 kg và giảm cân được nhiều nhất 1 kg thì khoảng biến thiên của mẫu số liệu gốc là 1.

Nếu biết nữ tăng cân nhiều nhất là 2 kg và giảm cân được nhiều nhất 1 kg thì khoảng biến thiên của mẫu số liệu gốc là $R_2 = 2 - (-1) = 3$.

Chọn SAI.

Câu 52. Thu nhập theo tháng (đơn vị: triệu đồng) của 20 người lao động ở ba nhà máy như sau:

Thu nhập	[5; 8)	[8; 11)	[11; 14)	[14; 17)	[17; 20)	[20; 23)
Số người của nhà máy A	2	5	4	4	5	0
Số người của nhà máy B	0	6	4	3	7	0
Số người của nhà máy C	1	5	8	6	0	0

Xét tính đúng/sai các mệnh đề sau:

(a) Trong 20 người lao động ở nhà máy A, hiệu số thu nhập của hai người lao động bất kì không vượt quá 15 triệu đồng

(b) Trong 20 người lao động ở nhà máy B, hiệu số thu nhập của hai người lao động bất kì không vượt quá 18 triệu đồng

(c) Nếu dựa vào khoảng biến thiên thì thu nhập của người lao động ở nhà máy A phân tán hơn so với người lao động ở nhà máy B

(d) Nếu dựa vào khoảng biến thiên thì thu nhập của người lao động ở nhà máy C phân tán hơn so với người lao động ở nhà máy A

Lời giải

Ta có khoảng biến thiên thu nhập của người lao động ở nhà máy A là $20 - 5 = 15$ triệu đồng.

Ta có khoảng biến thiên thu nhập của người lao động ở nhà máy B là $20 - 8 = 12$ triệu đồng.

Ta có khoảng biến thiên thu nhập của người lao động ở nhà máy C là $17 - 5 = 12$ triệu đồng.

(a) Trong 20 người lao động ở nhà máy A, hiệu số thu nhập của hai người lao động bất kì không vượt quá 15 triệu đồng.

Chọn ĐÚNG.

(b) Trong 20 người lao động ở nhà máy B, hiệu số thu nhập của hai người lao động bất kì không vượt quá 18 triệu đồng.

Chọn SAI.

(c) Nếu dựa vào khoảng biến thiên thì thu nhập của người lao động ở nhà máy A phân tán hơn so với người lao động ở nhà máy B.

Chọn ĐÚNG.

(d) Nếu dựa vào khoảng biến thiên thì thu nhập của người lao động ở nhà máy C phân tán hơn so với người lao động ở nhà máy A.

Chọn SAI.

Câu 53. Điểm kiểm tra cuối khóa môn Tiếng Anh của hai lớp ở một trung tâm ngoại ngữ được thống kê trong các bảng sau

Bảng a. Điểm của lớp A

Bảng b. Điểm của lớp B

Điểm	Số học viên	Điểm	Số học viên
[50; 60)	8	[50; 60)	15
[60; 70)	20	[60; 70)	20
[70; 80)	50	[70; 80)	30
[80; 90)	17	[80; 90)	20
[90; 100)	5	[90; 100)	15

Xét tính đúng/sai các mệnh đề sau:

(a) Nhóm chứa tứ phân vị thứ nhất của mẫu số liệu ở Bảng a là [60; 70).

(b) Nhóm chứa tứ phân vị thứ ba của mẫu số liệu ở Bảng b là [90; 100).

(c) Khoảng tứ phân vị của mẫu số liệu ở Bảng a là 10.

(d) Khoảng tứ phân vị của mẫu số liệu ở Bảng b là 20.

Lời giải

(a) Nhóm chứa tứ phân vị thứ nhất của mẫu số liệu ở Bảng a là $[60; 70)$.

Ta có $n = 100$. Mà $8 < \frac{n}{4} = \frac{100}{4} = 25 < 8 + 20$.

Do đó, nhóm chứa tứ phân vị thứ nhất là $[60; 70)$.

Chọn ĐÚNG.

(b) Nhóm chứa tứ phân vị thứ ba của mẫu số liệu ở Bảng b là $[90; 100)$.

Ta có $n = 100$. Mà $65 < \frac{3n}{4} = \frac{3 \cdot 100}{4} = 75 < 85$.

Do đó, nhóm chứa tứ phân vị thứ ba là $[80; 90)$.

Chọn SAI.

(c) Khoảng tứ phân vị của mẫu số liệu ở Bảng a là 10.

Ta có $n = 100$.

Mà $8 < \frac{n}{4} = \frac{100}{4} = 25 < 8 + 20$. Do đó, nhóm chứa tứ phân vị thứ nhất là $[60; 70)$.

$$\Rightarrow Q_1 = 60 + \frac{\frac{100}{4} - 8}{20} \cdot (70 - 60) = 68,5.$$

Lại có $28 < \frac{3n}{4} = \frac{3 \cdot 100}{4} = 75 < 78$. Do đó, nhóm chứa tứ phân vị thứ ba là $[70; 80)$.

$$\Rightarrow Q_3 = 70 + \frac{\frac{3 \cdot 100}{4} - (8 + 20)}{50} \cdot (80 - 70) = 79,4.$$

Vậy khoảng tứ phân vị của mẫu số liệu trên là $\Delta_Q = Q_3 - Q_1 = 79,4 - 68,5 = 10,9$.

Chọn SAI.

(d) Khoảng tứ phân vị của mẫu số liệu ở Bảng b là 20.

Ta có $n = 100$.

Mà $15 < \frac{n}{4} = \frac{100}{4} = 25 < 35$. Do đó, nhóm chứa tứ phân vị thứ nhất là $[60; 70)$.

$$\Rightarrow Q_1 = 60 + \frac{\frac{100}{4} - 15}{20} \cdot (70 - 60) = 65.$$

Lại có $65 < \frac{3n}{4} = \frac{3 \cdot 100}{4} = 75 < 85$. Do đó, nhóm chứa tứ phân vị thứ ba là $[80; 90)$.

$$\Rightarrow Q_3 = 80 + \frac{\frac{3 \cdot 100}{4} - 65}{20} \cdot (90 - 80) = 85.$$

Vậy khoảng tứ phân vị của mẫu số liệu trên là $\Delta_Q = Q_3 - Q_1 = 85 - 65 = 20$.

Chọn ĐÚNG.

Câu 54. Cho mẫu số liệu thống kê chiều cao của các học sinh lớp 12A1, 12A2 và 12A3 của một trường THPT như bảng sau

Chiều cao (đơn vị: cm)	[150; 155)	[155; 160)	[160; 165)	[165; 170)	[170; 175)	[175; 180)
Số học sinh lớp 12A1	1	13	18	5	3	0
Số học sinh lớp 12A2	0	12	20	7	1	0
Số học sinh lớp 12A3	1	8	12	15	3	1

Xét tính đúng/sai các mệnh đề sau:

(a) Nếu dựa vào khoảng tứ phân vị thì mẫu số liệu thống kê chiều cao của học sinh lớp 12A1 phân tán hơn so với lớp 12A2.

(b) Nếu dựa vào khoảng tứ phân vị thì mẫu số liệu thống kê chiều cao của học sinh lớp 12A2 phân tán hơn so với lớp 12A3.

(c) Ở lớp 12A2 có một học sinh có chiều cao là 173 cm, chiều cao của học sinh đó là giá trị ngoại lệ của mẫu số liệu của lớp 12A2.

(d) Ở lớp 12A3 có một học sinh có chiều cao là 177 cm, chiều cao của học sinh đó là giá trị ngoại lệ của mẫu số liệu của lớp 12A3.

Lời giải

Xét mẫu số liệu thống kê chiều cao của học sinh lớp 12A1

Ta có $n = 40$.

Mà $1 < \frac{n}{4} = \frac{40}{4} = 10 < 14$. Do đó, nhóm chứa tứ phân vị thứ nhất là $[155; 160)$.

$$\Rightarrow Q_1 = 155 + \frac{\frac{40}{4} - 1}{13} \cdot (160 - 155) = \frac{2060}{13}.$$

Lại có $14 < \frac{3n}{4} = \frac{3 \cdot 40}{4} = 30 < 32$. Do đó, nhóm chứa tứ phân vị thứ ba là $[160; 165)$.

$$\Rightarrow Q_3 = 160 + \frac{\frac{3 \cdot 40}{4} - 14}{18} \cdot (165 - 160) = \frac{1480}{9}.$$

$$\text{Vậy khoảng tứ phân vị của mẫu số liệu trên là } \Delta_Q' = Q_3 - Q_1 = \frac{1480}{9} - \frac{2060}{13} = \frac{700}{117}.$$

Xét mẫu số liệu thống kê chiều cao của học sinh lớp 12A2

Ta có $n = 40$.

Mà $0 < \frac{n}{4} = \frac{40}{4} = 10 < 12$. Do đó, nhóm chứa tứ phân vị thứ nhất là $[155; 160)$.

$$\Rightarrow Q_1 = 155 + \frac{\frac{40}{4} - 0}{12} \cdot (155 - 150) = \frac{955}{6}.$$

Lại có $12 < \frac{3n}{4} = \frac{3 \cdot 40}{4} = 30 < 32$. Do đó, nhóm chứa tứ phân vị thứ ba là $[160; 165)$.

$$\Rightarrow Q_3 = 160 + \frac{\frac{3 \cdot 40}{4} - 12}{20} \cdot (165 - 160) = \frac{329}{2}.$$

$$\text{Vậy khoảng tứ phân vị của mẫu số liệu trên là } \Delta_Q'' = Q_3 - Q_1 = \frac{329}{2} - \frac{955}{6} = \frac{16}{3}.$$

Xét mẫu số liệu thống kê chiều cao của học sinh lớp 12A3

Ta có $n = 40$.

Mà $9 < \frac{n}{4} = \frac{40}{4} = 10 < 21$. Do đó, nhóm chứa tứ phân vị thứ nhất là $[160; 165)$.

$$\Rightarrow Q_1 = 160 + \frac{\frac{40}{4} - 9}{12} \cdot (160 - 155) = \frac{1925}{12}.$$

Lại có $21 < \frac{3n}{4} = \frac{3 \cdot 40}{4} = 30 < 36$. Do đó, nhóm chứa tứ phân vị thứ ba là $[165; 170)$.

$$\Rightarrow Q_3 = 165 + \frac{\frac{3 \cdot 40}{4} - 21}{15} \cdot (170 - 165) = 168.$$

Vậy khoảng tứ phân vị của mẫu số liệu trên là $\Delta_Q''' = Q_3 - Q_1 = 168 - \frac{1925}{12} = \frac{91}{12}$.

(a) Nếu dựa vào khoảng tứ phân vị thì mẫu số liệu thống kê chiều cao của học sinh lớp 12A1 phân tán hơn so với lớp 12A2.

Ta có $\Delta_Q' > \Delta_Q'' \left(\frac{700}{117} > \frac{16}{3} \right)$. Do đó, mẫu số liệu thống kê chiều cao của học sinh lớp 12A1 phân tán hơn so với lớp 12A2.

Chọn ĐÚNG.

(b) Nếu dựa vào khoảng tứ phân vị thì mẫu số liệu thống kê chiều cao của học sinh lớp 12A2 phân tán hơn so với lớp 12A3.

Ta có $\Delta_Q'' < \Delta_Q''' \left(\frac{16}{3} < \frac{91}{12} \right)$. Do đó, mẫu số liệu thống kê chiều cao của học sinh lớp 12A3 phân tán hơn so với lớp 12A2.

Chọn SAI.

(c) Ở lớp 12A2 có một học sinh có chiều cao là 173 cm, chiều cao của học sinh đó là giá trị ngoại lệ của mẫu số liệu của lớp 12A2.

Xét mẫu số liệu lớp 12A2, ta có $\Delta_Q'' = \frac{16}{3}$. Khi đó, giá trị ngoại lệ là các giá trị $x > Q_3 + 1,5\Delta_Q'' \Leftrightarrow x > \frac{329}{2} + 1,5 \cdot \frac{16}{3} \Leftrightarrow x > 172,5$. Do đó, giá trị 173 cm là giá trị ngoại lệ của mẫu số liệu lớp 12A2.

Chọn ĐÚNG.

(d) Ở lớp 12A3 có một học sinh có chiều cao là 177 cm, chiều cao của học sinh đó là giá trị ngoại lệ của mẫu số liệu của lớp 12A3.

Xét mẫu số liệu lớp 12A3, ta có $\Delta_Q''' = \frac{91}{12}$. Khi đó, giá trị ngoại lệ là các giá trị $x > Q_3 + 1,5\Delta_Q''' \Leftrightarrow x > 168 + 1,5 \cdot \frac{91}{12} \Leftrightarrow x > 179,375$. Do đó, giá trị 177 cm là không là giá trị ngoại lệ của mẫu số liệu lớp 12A3.

Chọn SAI.

Câu 55. Để chuẩn bị mở một trung tâm thể dục thể thao, ông A đã tiến hành điều tra tuổi thọ (đơn vị tính bằng năm) của máy chạy bộ do ba hãng X, Y, Z sản xuất. Bảng số liệu sau biểu thị các mẫu số liệu mà anh thu được trong quá trình nghiên cứu

Tuổi thọ	[2; 4)	[4; 6)	[6; 8)	[8; 10)	[10; 12)
Số máy của hãng X	7	20	36	20	17
Số máy của hãng Y	0	20	35	35	10
Số máy của hãng Z	5	20	30	30	15

Xét tính đúng/sai các mệnh đề sau:

(a) Khoảng biến thiên của máy chạy bộ do hãng X sản xuất bằng 10.

(b) Khoảng biến thiên của máy chạy bộ do hãng Y sản xuất không vượt quá 6.

(c) Khoảng biến thiên của máy chạy bộ do hãng Z sản xuất thì lớn hơn khoảng biến thiên của máy chạy bộ do hãng Y sản xuất.

(d) Nếu dựa vào khoảng biến thiên thì máy chạy bộ do hãng Z sản xuất có tuổi thọ phân tán nhất trong ba hãng.

Lời giải

Ta có khoảng biến thiên của máy chạy bộ do hãng X sản xuất là

$$R_X = 12 - 2 = 10.$$

Ta có khoảng biến thiên của máy chạy bộ do hãng Y sản xuất là

$$R_Y = 12 - 4 = 8.$$

Ta có khoảng biến thiên của máy chạy bộ do hãng Z sản xuất là

$$R_Z = 12 - 2 = 10.$$

Ta có $R_Z = R_X = 10 > 8 = R_Y$

(a) Khoảng biến thiên của máy chạy bộ do hãng X sản xuất bằng 10.

Chọn ĐÚNG.

(b) Khoảng biến thiên của máy chạy bộ do hãng Y sản xuất không vượt quá 6.

Chọn SAI.

(c) Khoảng biến thiên của máy chạy bộ do hãng Z sản xuất thì lớn hơn khoảng biến thiên của máy chạy bộ do hãng Y sản xuất.

Chọn ĐÚNG.

(d) Nếu dựa vào khoảng biến thiên thì máy chạy bộ do hãng Z sản xuất có tuổi thọ phân tán nhất trong ba hãng.

Chọn SAI.

Câu 56. Điểm kiểm tra cuối học kỳ 2 môn Toán (thang điểm 10) của hai lớp 12A và 12B ở một trường THPT được thống kê trong các bảng sau

Bảng A. Điểm của lớp 12A

Bảng B. Điểm của lớp 12B

Điểm	Số học sinh	Điểm	Số học sinh
[6, 5; 7)	5	[6, 5; 7)	3
[7; 7, 5)	4	[7; 7, 5)	6
[7, 5; 8)	9	[7, 5; 8)	7
[8; 8, 5)	10	[8; 8, 5)	12
[8, 5; 9)	7	[8, 5; 9)	8
[9; 9, 5)	4	[9; 9, 5)	3
[9, 5; 10)	2	[9, 5; 10)	2

Xét tính đúng/sai các mệnh đề sau:

(a) Nhóm chứa tứ phân vị thứ nhất của mẫu số liệu ở Bảng A là [7; 7, 5)

(b) Nhóm chứa tứ phân vị thứ ba của mẫu số liệu ở Bảng B là [8, 5; 9)

(c) Khoảng tứ phân vị của mẫu số liệu ở Bảng A thì lớn hơn 2

(d) Khoảng tứ phân vị của mẫu số liệu ở Bảng B nhỏ hơn khoảng tứ phân vị của mẫu số liệu ở bảng A

Lời giải

(a) Nhóm chứa tứ phân vị thứ nhất của mẫu số liệu ở Bảng A là [7; 7, 5).

Nhóm chứa tứ phân vị thứ nhất của mẫu số liệu ở Bảng A là [7; 7, 5).

Gọi $x_1 \leq x_2 \leq \dots \leq x_{41}$ là số điểm kiểm tra Toán ở học kỳ 2 của lớp 12A.

Do đó tứ phân vị thứ nhất của mẫu số liệu là $x_{11} \in [7,5; 8)$.

Chọn SAI.

(b) Nhóm chứa tứ phân vị thứ ba của mẫu số liệu ở Bảng B là $[8,5; 9)$.

Nhóm chứa tứ phân vị thứ ba của mẫu số liệu ở Bảng B là $[8,5; 9)$.

Gọi $y_1 \leq y_2 \leq \dots \leq y_{41}$ là số điểm kiểm tra Toán ở học kỳ 2 của lớp 12B.

Do đó tứ phân vị thứ ba của mẫu số liệu là $y_{31} \in [8,5; 9)$.

Chọn ĐÚNG.

(c) Khoảng tứ phân vị của mẫu số liệu ở Bảng A thì lớn hơn 2.

Khoảng tứ phân vị của mẫu số liệu ở Bảng A thì lớn hơn 2.

Xét mẫu số liệu ở bảng A. Ta có

Ta có nhóm chứa tứ phân vị thứ nhất là $[7; 7,5)$.

$$\Rightarrow Q_1 = 7 + \frac{\frac{41}{4} - 9}{9} \cdot (8 - 7,5) = 7,07.$$

Nhóm chứa tứ phân vị thứ ba là $[8,5; 9)$.

$$\Rightarrow Q_3 = 8,5 + \frac{\frac{3 \cdot 41}{4} - 28}{7} \cdot (9 - 8,5) = 8,70.$$

Vậy khoảng tứ phân vị của mẫu số liệu trên là $\Delta_Q = Q_3 - Q_1 = 8,70 - 7,07 = 1,63$.

Chọn SAI.

(d) Khoảng tứ phân vị của mẫu số liệu ở Bảng B nhỏ hơn khoảng tứ phân vị của mẫu số liệu ở bảng A.

Khoảng tứ phân vị của mẫu số liệu ở Bảng B nhỏ hơn khoảng tứ phân vị của mẫu số liệu ở bảng A.

Xét mẫu số liệu ở bảng B. Ta có

Nhóm chứa tứ phân vị thứ nhất là $[7,5; 8)$.

$$\Rightarrow Q_1 = 7,5 + \frac{\frac{41}{4} - 9}{7} \cdot (8 - 7,5) = 7,59.$$

Nhóm chứa tứ phân vị thứ ba là $[8,5; 9)$.

$$\Rightarrow Q_3 = 8,5 + \frac{\frac{3 \cdot 41}{4} - 28}{8} \cdot (9 - 8,5) = 8,67.$$

Vậy khoảng tứ phân vị của mẫu số liệu trên là $\Delta_Q = Q_3 - Q_1 = 8,67 - 7,59 = 1,08$.

Chọn ĐÚNG.

Câu 57. Số khách hàng mua bảo hiểm ở từng độ tuổi được thống kê trong bảng số liệu ghép nhóm sau:

Độ tuổi	$[20; 30)$	$[30; 40)$	$[40; 50)$	$[50; 60)$	$[60; 70)$
Số khách hàng	4	6	10	7	3

Xét tính đúng/sai các mệnh đề sau:

(a) Giá trị đại diện của nhóm $[20; 30)$ là 25.

(b) Giá trị đại diện của nhóm $[60; 70)$ là 70.

(c) Công thức tính độ tuổi bình quân của khách hàng mua bảo hiểm trong bảng thống kê là

$$\bar{x} = \frac{20 \cdot 4 + 30 \cdot 6 + 40 \cdot 10 + 50 \cdot 7 + 60 \cdot 3}{30}.$$

(d) Độ tuổi trung bình (làm tròn đến hàng phần trăm) của khách hàng mua bảo hiểm là 44,67.

🔗 Lời giải

(a) giá trị đại diện của nhóm $[20; 30)$ là 25.

Giá trị đại diện của nhóm $[20; 30)$ là $\frac{20+30}{2} = 25$.

Chọn ĐÚNG.

(b) giá trị đại diện của nhóm $[60; 70)$ là 70.

Giá trị đại diện của nhóm $[60; 70)$ là $\frac{60+70}{2} = 65$.

Chọn SAI.

(c) Công thức tính độ tuổi bình quân của khách hàng mua bảo hiểm trong bảng thống kê là $\bar{x} = \frac{20.4 + 30.6 + 40.10 + 50.7 + 60.3}{30}$.

Công thức tính độ tuổi bình quân của khách hàng mua bảo hiểm trong bảng thống kê là

$$\bar{x} = \frac{25.4 + 35.6 + 45.10 + 55.7 + 65.3}{30}$$

Chọn SAI.

(d) độ tuổi trung bình (làm tròn đến hàng phần trăm) của khách hàng mua bảo hiểm là 44,67.

Độ tuổi trung bình của khách hàng mua bảo hiểm là

$$\bar{x} = \frac{25.4 + 35.6 + 45.10 + 55.7 + 65.3}{30} \approx 44,67.$$

Chọn ĐÚNG.

Câu 58. Xét mẫu dữ liệu ghép nhóm được cho ở bảng sau với $n = n_1 + n_2 + \dots + n_k$.

Nhóm	$[u_1; u_2)$	$[u_2; u_3)$	...	$[u_k; u_{k+1})$
Giá trị đại diện	c_1	c_2	...	c_k
Tần số	n_1	n_2	...	n_k

Xét tính đúng/sai các mệnh đề sau:

(a) Giá trị đại diện của $[u_2; u_3)$ là $c_2 = \frac{u_2 + u_3}{2}$.

(b) Giá trị trung bình của mẫu số liệu ghép nhóm là $\bar{x} = \frac{1}{n}(c_1 + c_2 + \dots + c_k)$.

(c) Phương sai của mẫu số liệu ghép nhóm là $s^2 = \frac{1}{n}[n_1(c_1 - \bar{x})^2 + n_2(c_2 - \bar{x})^2 + \dots + n_k(c_k - \bar{x})^2]$.

(d) Phương sai của mẫu số liệu ghép nhóm là $s^2 = \frac{1}{n}(n_1c_1^2 + n_2c_2^2 + \dots + n_kc_k^2) - \bar{x}$.

🔗 Lời giải

(a) giá trị đại diện của $[u_2; u_3)$ là $c_2 = \frac{u_2 + u_3}{2}$.

Chọn ĐÚNG.

(b) giá trị trung bình của mẫu số liệu ghép nhóm là $\bar{x} = \frac{1}{n}(c_1 + c_2 + \dots + c_k)$.

Giá trị trung bình của mẫu số liệu ghép nhóm là $\bar{x} = \frac{1}{n}(n_1c_1 + n_2c_2 + \dots + n_kc_k)$.

Chọn SAI.

(c) phương sai của mẫu số liệu ghép nhóm là $s^2 = \frac{1}{n}[n_1(c_1 - \bar{x})^2 + n_2(c_2 - \bar{x})^2 + \dots + n_k(c_k - \bar{x})^2]$.

Chọn ĐÚNG.

(d) phương sai của mẫu số liệu ghép nhóm là $s^2 = \frac{1}{n}(n_1c_1^2 + n_2c_2^2 + \dots + n_kc_k^2) - \bar{x}^2$.

Phương sai của mẫu số liệu ghép nhóm là

$$s^2 = \frac{1}{n}(n_1c_1^2 + n_2c_2^2 + \dots + n_kc_k^2) - \bar{x}^2.$$

Chọn SAI.

Câu 59. Cho mẫu số liệu dưới dạng bảng sau

Số câu trả lời đúng	[16; 21)	[21; 26)	[26; 31)	[31; 36)	[36; 41)	
Tần số	4	6	8	18	4	$N = 40$

Xét tính đúng/sai các mệnh đề sau:

(a) Giá trị đại diện của lớp [36; 41) là 38,5.

(b) Công thức tính số trung bình là $\bar{x} = \frac{18,5 \cdot 4 + 23,5 \cdot 6 + 28,5 \cdot 8 + 33,5 \cdot 18 + 38,5 \cdot 4}{40}$.

(c) Số trung bình là 30.

(d) Phương sai của mẫu số liệu là $S^2 = 32,75$.

Lời giải

(a) giá trị đại diện của lớp [36; 41) là 38,5.

Chọn ĐÚNG.

(b) công thức tính số trung bình là $\bar{x} = \frac{18,5 \cdot 4 + 23,5 \cdot 6 + 28,5 \cdot 8 + 33,5 \cdot 18 + 38,5 \cdot 4}{40}$.

Chọn ĐÚNG.

(c) số trung bình là 30.

số trung bình là $\bar{x} = \frac{18,5 \cdot 4 + 23,5 \cdot 6 + 28,5 \cdot 8 + 33,5 \cdot 18 + 38,5 \cdot 4}{40} = 30$.

Chọn ĐÚNG.

(d) phương sai của mẫu số liệu là $S^2 = 32,75$.

Phương sai của mẫu số liệu là

$$S^2 = \frac{1}{40} \left[4(18,5 - 30)^2 + 6(23,5 - 30)^2 + 8(28,5 - 30)^2 + 18(33,5 - 30)^2 + 4(38,5 - 30)^2 \right] = 32,75$$

Chọn ĐÚNG.

Câu 60. Số lượng khách hàng nữ mua hàng thời trang trong một ngày của một cửa hàng được thống kê trong bảng tần số ghép nhóm sau:

Khoảng tuổi	[20; 30)	[30; 40)	[40; 50)	[50; 60)	[60; 70)
Số khách hàng nữ	3	9	6	4	2

Xét tính đúng/sai các mệnh đề sau:

(a) Giá trị đại diện của nhóm [30; 40) là 35.

(b) Cỡ mẫu là $n = 20$.

(c) Số trung bình của mẫu số liệu ghép nhóm là 42,08 (làm tròn đến hàng phần trăm).

(d) Phương sai của mẫu số liệu ghép nhóm là 129,99 (làm tròn đến hàng phần trăm).

Lời giải

Khoảng tuổi	[20; 30)	[30; 40)	[40; 50)	[50; 60)	[60; 70)
Giá trị đại diện	25	35	45	55	65
Số khách hàng nữ	3	9	6	4	2

(a) Giá trị đại diện của nhóm [30; 40) là 35.

Giá trị đại diện của nhóm [30; 40) là 35.

Ta có giá trị đại diện của nhóm [30; 40) là: $\frac{30+40}{2} = 35$.

Chọn ĐÚNG.

(b) Cỡ mẫu là $n = 20$.

Cỡ mẫu là: $n = 3 + 9 + 6 + 4 + 2 = 24$.

Chọn SAI.

(c) Số trung bình của mẫu số liệu ghép nhóm là 42,08 (làm tròn đến hàng phần trăm).

$$\bar{x} = \frac{25.3 + 35.9 + 45.6 + 55.4 + 65.2}{24} = \frac{1010}{24} \approx 42,08$$

Chọn ĐÚNG.

(d) Phương sai của mẫu số liệu ghép nhóm là 129,99 (làm tròn đến hàng phần trăm).

$$S^2 = \frac{1}{24} (3.25^2 + 9.35^2 + 6.45^2 + 4.55^2 + 2.65^2) - \left(\frac{1010}{24} \right)^2 \approx 128,99$$

Chọn SAI.

Câu 61. Thời gian (phút) để học sinh hoàn thành một câu hỏi thi được cho như sau:

Thời gian (phút)	[0, 5; 10, 5)	[10, 5; 20, 5)	[20, 5; 30, 5)	[30, 5; 40, 5)	[40, 5; 50, 5)
Số học sinh lớp 11A	2	10	6	4	3
Số học sinh lớp 11B	3	8	10	2	4

Xét tính đúng/sai các mệnh đề sau:

(a) Phương sai của mẫu số liệu lớp 11A là: 133,44 (làm tròn đến hàng phần trăm).

(b) Độ lệch chuẩn của mẫu số liệu lớp 11A là: 11,77 (làm tròn đến hàng phần trăm).

(c) Độ lệch chuẩn của mẫu số liệu lớp 11B là: 11,55 (làm tròn đến hàng phần trăm).

(d) Nếu so sánh theo độ lệch chuẩn thì thời gian để học sinh hoàn thành một câu hỏi thi của lớp 11A ít phân tán hơn lớp 11B

Lời giải

Thời gian (phút)	[0, 5; 10, 5)	[10, 5; 20, 5)	[20, 5; 30, 5)	[30, 5; 40, 5)	[40, 5; 50, 5)
Giá trị đại diện	5,5	15,5	25,5	35,5	45,5
Số học sinh lớp 11A	2	10	6	4	3
Số học sinh lớp 11B	3	8	10	2	4

(a) Phương sai của mẫu số liệu lớp 11A là: 133,44 (làm tròn đến hàng phần trăm).

Phương sai của mẫu số liệu lớp 11A là: 133,44 (làm tròn đến hàng phần trăm).

Xét mẫu số liệu của lớp 11A:

Cỡ mẫu là $n_1 = 2 + 10 + 6 + 4 + 3 = 25$

$$\text{Số trung bình: } \bar{x}_1 = \frac{5.2 + 15.10 + 25.6 + 35.4 + 45.3}{25} = 23,9$$

Phương sai:

$$S_1^2 = \frac{1}{25} (2.5,5^2 + 10.15,5^2 + 6.25,5^2 + 4.35,5^2 + 3.45,5^2) - 23,9^2 = 133,44$$

Chọn ĐÚNG.

(b) Độ lệch chuẩn của mẫu số liệu lớp 11A là: 11,77 (làm tròn đến hàng phần trăm).

$$S_1 = \sqrt{133,44} \approx 11,55$$

Chọn SAI.

(c) Độ lệch chuẩn của mẫu số liệu lớp 11B là: 11,55 (làm tròn đến hàng phần trăm).

Xét mẫu số liệu của lớp 11B:

$$\text{Cỡ mẫu là } n_2 = 3 + 8 + 10 + 2 + 4 = 27$$

$$\text{Số trung bình: } \bar{x}_2 = \frac{5,5 \cdot 3 + 15,5 \cdot 8 + 25,5 \cdot 10 + 35,5 \cdot 2 + 45,5 \cdot 4}{27} = \frac{648,5}{27} \approx 24,02$$

$$S_2^2 = \frac{1}{27} (3 \cdot 5,5^2 + 8 \cdot 15,5^2 + 10 \cdot 25,5^2 + 2 \cdot 35,5^2 + 4 \cdot 45,5^2) - 24,02^2 \approx 138,47$$

$$S_2 \approx \sqrt{138,47} \approx 11,77$$

Chọn SAI.

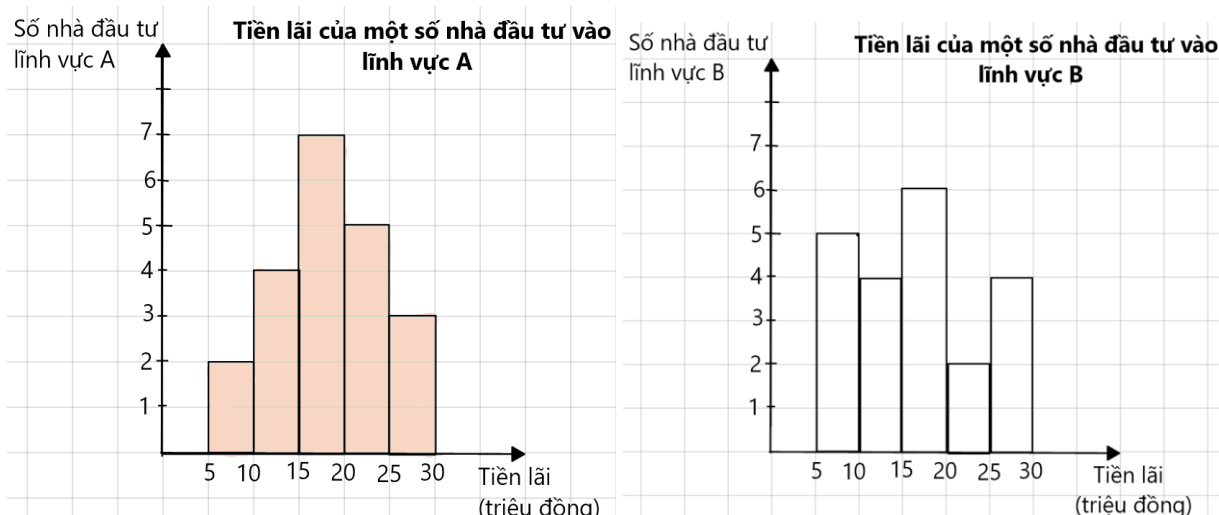
(d) Nếu so sánh theo độ lệch chuẩn thì thời gian để học sinh hoàn thành một câu hỏi thi của lớp 11A ít phân tán hơn lớp 11B.

$$\text{Ta có: } S_1 < S_2$$

Nên nếu so sánh theo độ lệch chuẩn thì thời gian để học sinh hoàn thành một câu hỏi thi của lớp 11A ít phân tán hơn lớp 11B.

Chọn ĐÚNG.

Câu 62. Người ta ghi lại tiền lãi (đơn vị: triệu đồng) của một số nhà đầu tư (với số tiền đầu tư như nhau), khi đầu tư và hai lĩnh vực A, B cho kết quả bằng biểu đồ dưới đây



Xét tính đúng/sai các mệnh đề sau:

(a) Độ lệch chuẩn của mẫu số liệu số nhà đầu tư vào lĩnh vực A là: 5,83 (làm tròn đến hàng phần trăm).

(b) Độ lệch chuẩn của mẫu số liệu số nhà đầu tư vào lĩnh vực B là: 7,01 (làm tròn đến hàng phần trăm).

(c) Về trung bình, đầu tư vào lĩnh vực B đem lại tiền lãi cao hơn lĩnh vực A.

(d) Nếu so sánh theo độ lệch chuẩn thì tiền lãi của các nhà đầu tư trong lĩnh vực A có xu hướng phân tán rộng hơn so với tiền lãi của các nhà đầu tư trong lĩnh vực B.

Lời giải

Từ biểu đồ ta có bảng thống kê sau:

Tiền lãi	[5;10)	[10;15)	[15;20)	[20;25)	[25;30)
Giá trị đại diện	7,5	12,5	17,5	22,5	27,5
Số nhà đầu tư vào lĩnh vực A	2	4	7	5	3
Số nhà đầu tư vào lĩnh vực B	5	4	6	2	4

(a) Độ lệch chuẩn của mẫu số liệu số nhà đầu tư vào lĩnh vực A là: 5,83 (làm tròn đến hàng phần trăm).

Xét mẫu số liệu của số nhà đầu tư vào lĩnh vực A:

Cỡ mẫu là $n_1 = 2 + 4 + 7 + 5 + 3 = 21$

$$\text{Số trung bình: } \bar{x}_1 = \frac{7,5.2 + 12,5.4 + 17,5.7 + 22,5.5 + 27,5.3}{21} = \frac{255}{14}$$

Phương sai:

$$S_1^2 = \frac{1}{21} (2,7.5^2 + 4.12,5^2 + 7.17,5^2 + 5.22,5^2 + 3.27,5^2) - \left(\frac{255}{14} \right)^2 = \frac{5000}{147}$$

$$S_1 = \sqrt{\frac{5000}{147}} \approx 5,83$$

Chọn ĐÚNG.

(b) Độ lệch chuẩn của mẫu số liệu số nhà đầu tư vào lĩnh vực B là: 7,01 (làm tròn đến hàng phần trăm).

Xét mẫu số liệu của số nhà đầu tư vào lĩnh vực B:

Cỡ mẫu là $n_2 = 5 + 4 + 6 + 2 + 4 = 21$

$$\text{Số trung bình: } \bar{x}_2 = \frac{7,5.5 + 12,5.4 + 17,5.6 + 22,5.2 + 27,5.4}{21} = \frac{695}{42}$$

$$S_2^2 = \frac{1}{21} (5.7,5^2 + 4.12,5^2 + 6.17,5^2 + 2.22,5^2 + 4.27,5^2) - \left(\frac{695}{42} \right)^2 = \frac{21650}{441}$$

$$S_2 = \sqrt{\frac{21650}{441}} \approx 7,01$$

Chọn ĐÚNG.

(c) Về trung bình, đầu tư vào lĩnh vực B đem lại tiền lãi cao hơn lĩnh vực A.

$$\text{Số trung bình: } \bar{x}_1 = \frac{7,5.2 + 12,5.4 + 17,5.7 + 22,5.5 + 27,5.3}{21} = \frac{255}{14} \approx 18,21$$

$$\text{Số trung bình: } \bar{x}_2 = \frac{7,5.5 + 12,5.4 + 17,5.6 + 22,5.2 + 27,5.4}{21} = \frac{695}{42} \approx 16,55$$

Về trung bình, đầu tư vào lĩnh vực A đem lại tiền lãi cao hơn lĩnh vực

B.

Chọn SAI.

(d) Nếu so sánh theo độ lệch chuẩn thì tiền lãi của các nhà đầu tư trong lĩnh vực A có xu hướng phân tán rộng hơn so với tiền lãi của các nhà đầu tư trong lĩnh vực

B.

Ta có: $S_1 < S_2$

Vậy nếu so sánh theo độ lệch chuẩn thì tiền lãi của các nhà đầu tư trong lĩnh vực B có xu hướng phân tán rộng hơn so với tiền lãi của các nhà đầu tư trong lĩnh vực A.

Chọn SAI.

Câu 63. Bác sĩ A điều trị 18 bệnh nhân mỡ máu bằng cách xét nghiệm Cholesterol toàn phần trong buổi sáng điều trị như sau:

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

Xét tính đúng/sai các mệnh đề sau:

(a) Khoảng tứ phân vị của mẫu số liệu trên bằng 1, độ lệch chuẩn của mẫu số liệu trên bằng 0,61 do bác sĩ A điều trị.

(b) Bảng tần số ghép nhóm với nhóm đầu tiên là $[3,7;4,14)$ và độ dài mỗi nhóm bằng 0,44 do bác sĩ A điều trị được thống kê dưới đây

Chỉ số Cholesterol toàn phần do bác sĩ A điều trị	$[3,7;4,14)$	$[4,14;4,58)$	$[4,58;5,02)$	$[5,02;5,46)$	$[5,46;5,9)$
Số bệnh nhân	4	3	5	3	3

(c) Giá trị độ lệch chuẩn của mẫu số liệu ghép nhóm đầu tiên là $[3,7;4,14)$ và độ dài mỗi nhóm bằng 0,44 do bác sĩ A điều trị là 0,58

(d) Biết rằng bác sĩ B cũng điều trị 18 bệnh nhân trên với với nhóm đầu tiên là $[3,7;4,14)$ và độ dài mỗi nhóm bằng 0,44 được thống kê dưới đây:

Chỉ số Cholesterol toàn phần do bác sĩ B điều trị	$[3,7;4,14)$	$[4,14;4,58)$	$[4,58;5,02)$	$[5,02;5,46)$	$[5,46;5,9)$
Số bệnh nhân	3	4	3	4	4

Lời giải

(a) Khoảng tứ phân vị của mẫu số liệu trên bằng 1, độ lệch chuẩn của mẫu số liệu trên bằng 0,61 do bác sĩ A điều trị.

Sắp xếp lại bảng số liệu theo thứ tự không giảm như sau:

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

Gọi $x_1; x_2; \dots; x_{18}$ là mẫu số liệu gốc của 18 bệnh nhân mỡ máu bằng cách xét nghiệm Cholesterol toàn phần trong một ngày theo thứ tự không giảm.

$$\text{Trung vị } Q_2 = \frac{1}{2}(x_9 + x_{10}) = \frac{1}{2}(4,6 + 4,7) = 4,65.$$

Tứ phân vị thứ nhất của trung vị của nửa số liệu bên trái Q_2 là $Q_1 = 4,2$.

Tứ phân vị thứ nhất của trung vị của nửa số liệu bên phải Q_2 là $Q_3 = 5,2$.

Khoảng tứ phân vị của mẫu số liệu trên $\Delta Q = Q_3 - Q_1 = 5,2 - 4,2 = 1$.

$$\text{Số trung bình của mẫu số liệu trên do bác sĩ A điều trị bằng } \bar{x} = \frac{2,3,8 + 4,0 + 4,1 + 4,2 + 4,3 + 4,4 + 4,5 + 4,6 + 4,7 + 4,8 + 5,0 + 5,1 + 5,2 + 5,3 + 2,5,6 + 5,8}{18} = \frac{212}{45}$$

Phương sai của mẫu số liệu trên do bác sĩ A điều trị bằng

$$S^2 = \frac{x_1^2 + x_2^2 + x_3^2 + \dots + x_{18}^2}{18} - \bar{x}^2 = \frac{3023}{8100}.$$

Độ lệch chuẩn của mẫu số liệu trên do bác sĩ A điều trị bằng $\sigma = \sqrt{S^2} = 0,61$.

Chọn ĐÚNG.

(b) Bảng tần số ghép nhóm với nhóm đầu tiên là $[3,7;4,14)$ và độ dài mỗi nhóm bằng 0,44 do bác sĩ A điều trị được thống kê dưới đây

Chỉ số Cholesterol toàn phần do bác sĩ A điều trị	$[3,7;4,14)$	$[4,14;4,58)$	$[4,58;5,02)$	$[5,02;5,46)$	$[5,46;5,9)$
Số bệnh nhân	4	4	4	3	3

Chọn SAI.

(c) Giá trị độ lệch chuẩn của mẫu số liệu ghép nhóm đầu tiên là $[3,7;4,14)$ và độ dài mỗi nhóm bằng 0,44 do bác sĩ A điều trị là 0,58

Chỉ số Cholesterol toàn phần do bác sĩ A điều trị	$[3,7;4,14)$	$[4,14;4,58)$	$[4,58;5,02)$	$[5,02;5,46)$	$[5,46;5,9)$
Giá trị đại diện	3,92	4,36	4,8	5,24	5,68
Số bệnh nhân	4	4	4	3	3

$$\text{Số trung bình của mẫu số liệu trên do bác sĩ A điều trị bằng } \bar{x}_A = \frac{4.3,92 + 4.4,36 + 4.4,8 + 3.5,24 + 3.5,68}{18} = \frac{709}{150}$$

Phương sai của mẫu số liệu trên do bác sĩ A điều trị bằng

$$S_A^2 = \frac{4.3,92^2 + 4.4,36^2 + 4.4,8^2 + 3.5,24^2 + 3.5,68^2}{18} - \left(\frac{709}{150}\right)^2 = \frac{2783}{7500}.$$

Độ lệch chuẩn của mẫu số liệu trên do bác sĩ A điều trị bằng $\sigma_A = \sqrt{S_A^2} = 0,609$.

Chọn SAI.

(d) Biết rằng bác sĩ B cũng điều trị 18 bệnh nhân trên với với nhóm đầu tiên là $[3,7;4,14)$ và độ dài mỗi nhóm bằng 0,44 được thống kê dưới đây

Chỉ số Cholesterol toàn phần do bác sĩ B điều trị	$[3,7;4,14)$	$[4,14;4,58)$	$[4,58;5,02)$	$[5,02;5,46)$	$[5,46;5,9)$
Giá trị đại diện	3,92	4,36	4,8	5,24	5,68
Số bệnh nhân	3	4	3	4	4

Số trung bình của mẫu số liệu trên do bác sĩ B điều trị bằng

$$\bar{x}_B = \frac{3.3,92 + 4.4,36 + 3.4,8 + 4.5,24 + 4.5,68}{18} = \frac{1091}{225}$$

Phương sai của mẫu số liệu trên do bác sĩ B điều trị bằng

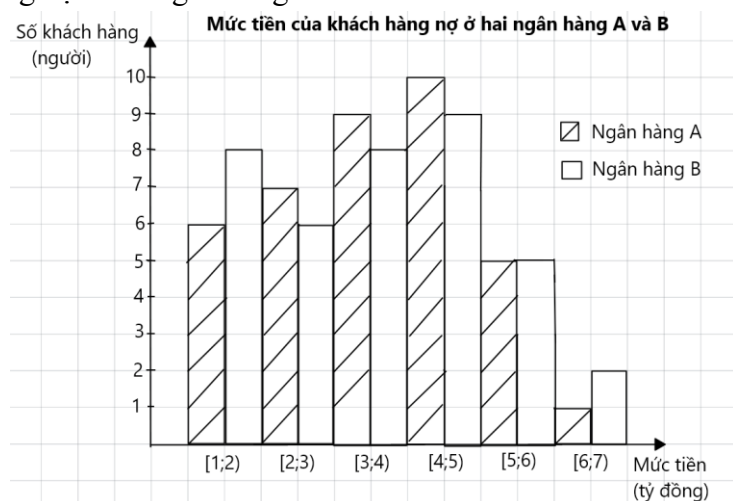
$$S_B^2 = \frac{3.3,92^2 + 4.4,36^2 + 3.4,8^2 + 4.5,24^2 + 4.5,68^2}{18} - \left(\frac{1091}{225}\right)^2 \approx 0,3848.$$

Độ lệch chuẩn của mẫu số liệu trên do bác sĩ B điều trị bằng $\sigma_B = \sqrt{S_B^2} = 0,62$.

Vì $\sigma_A < \sigma_B$ nên so sánh về độ lệch chuẩn thì chỉ số Cholesterol toàn phần bác sĩ A điều trị ít phân tán hơn bác sĩ B điều trị.

Chọn ĐÚNG.

Câu 64. Biểu đồ dưới đây biểu thị kết quả thu thập được về mức tiền (đơn vị: tỷ đồng) của một số khách hàng nợ ở hai ngân hàng A và B.



Xét tính đúng/sai các mệnh đề sau:

(a) Bảng giá trị đại diện cho mỗi nhóm và bảng tần số ghép nhóm cho mẫu số liệu tương ứng với biểu đồ trên

Mức tiền (tỷ đồng)	$[1;2)$	$[2;3)$	$[3;4)$	$[4;5)$	$[5;6)$	$[6;7)$
Mức tiền đại diện (tỷ đồng)	1,5	2,5	3,5	4,5	5,5	6,5
Số khách hàng ngân hàng A	6	7	9	10	5	1
Số khách hàng ngân hàng B	8	6	8	9	5	2

(b) Độ lệch chuẩn của mẫu số liệu ghép nhóm của ngân hàng A bằng $\frac{661}{361}$.

(c) Độ lệch chuẩn của mẫu số liệu ghép nhóm của ngân hàng B bằng $\frac{3221}{1444}$.

(d) Người ta dùng độ lệch chuẩn để so sánh mức độ rủi ro của số tiền khách hàng nợ ngân hàng. Ngân hàng nào có độ lệch chuẩn cao hơn thì có độ rủi ro lớn hơn. Theo quan điểm trên, độ rủi ro của ngân hàng A cao hơn ngân hàng B

Lời giải

(a) Bảng giá trị đại diện cho mỗi nhóm và bảng tần số ghép nhóm cho mẫu số liệu tương ứng với biểu đồ trên

Mức tiền (tỷ đồng)	[1; 2)	[2; 3)	[3; 4)	[4; 5)	[5; 6)	[6; 7)
Mức tiền đại diện (tỷ đồng)	1,5	2,5	3,5	4,5	5,5	6,5
Số khách hàng ngân hàng A	6	7	9	10	5	1
Số khách hàng ngân hàng B	8	6	8	9	5	2

Chọn ĐÚNG.

(b) Độ lệch chuẩn của mẫu số liệu ghép nhóm của ngân hàng A bằng $\frac{661}{361}$.

Số trung bình của mẫu số liệu ngân hàng A bằng $\bar{x}_A = \frac{6.1,5 + 7.2,5 + 9.3,5 + 10.4,5 + 5.5,5 + 1.6,5}{38} = \frac{137}{38}$

Phương sai của mẫu số liệu ngân hàng A bằng

$$S_A^2 = \frac{6.1,5^2 + 7.2,5^2 + 9.3,5^2 + 10.4,5^2 + 5.5,5^2 + 1.6,5^2}{38} - \left(\frac{137}{38}\right)^2 = \frac{661}{361}.$$

Độ lệch chuẩn của mẫu số liệu ngân hàng A bằng $\sigma_A = \sqrt{S_A^2} = \frac{\sqrt{661}}{19}$.

Chọn SAI.

(c) Độ lệch chuẩn của mẫu số liệu ghép nhóm của ngân hàng B bằng $\frac{3221}{1444}$.

Số trung bình của mẫu số liệu ngân hàng B bằng $\bar{x}_B = \frac{8.1,5 + 6.2,5 + 8.3,5 + 9.4,5 + 5.5,5 + 2.6,5}{38} = \frac{68}{19}$

Phương sai của mẫu số liệu ngân hàng B bằng

$$S_B^2 = \frac{8.1,5^2 + 6.2,5^2 + 8.3,5^2 + 9.4,5^2 + 5.5,5^2 + 2.6,5^2}{38} - \left(\frac{68}{19}\right)^2 = \frac{3221}{1444}.$$

Độ lệch chuẩn của mẫu số liệu ngân hàng B bằng $\sigma_B = \sqrt{S_B^2} = \sqrt{\frac{3221}{1444}}$.

Chọn SAI.

(d) Người ta dùng độ lệch chuẩn để so sánh mức độ rủi ro của số tiền khách hàng nợ ngân hàng. Ngân hàng nào có độ lệch chuẩn cao hơn thì có độ rủi ro lớn hơn. Theo quan điểm trên, độ rủi ro của ngân hàng A cao hơn ngân hàng B

Vì $\sigma_A < \sigma_B$ nên rủi ro của ngân hàng A thấp hơn rủi ro của ngân hàng B khi cho khách hàng vay nợ.

Chọn SAI.

✓ TRẮC NGHIỆM TRẢ LỜI NGẮN

Câu 65. Cho bảng thống kê thời gian tập thể dục buổi sáng mỗi ngày trong tháng 9/2022 của Bác An và Bác Bình. Gọi R_A, R_B lần lượt là khoảng biến thiên của mẫu số liệu về thời gian tập thể dục của Bác An và Bác Bình, khi đó $R_A + R_B$ bằng

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



Bác Bình	0	20	5	5	0
----------	---	----	---	---	---

✎ **Lời giải**✓ **Trả lời: 40**

Ta có $R_A = 40 - 15 = 25$; $R_B = 35 - 20 = 15$ suy ra $R_A + R_B = 25 + 15 = 40$.

Câu 66. Một hãng xe ô tô thống kê lại số lần gặp sự cố về động cơ của 100 chiếc xe cùng loại sau hai năm sử dụng đầu tiên được cho ở bảng sau

Số lần gặp sự cố	[1;2]	[3;4]	[5;6]	[7;8]	[9;10]
Số xe	17	33	25	20	5

Khoảng biến thiên của bảng số liệu trên là

✎ **Lời giải**✓ **Trả lời: 10**

Ta hiệu chỉnh lại bảng số liệu như sau

Số lần gặp sự cố	[0,5; 2,5)	[2,5; 4,5)	[4,5; 6,5)	[6,5; 8,5)	[8,5; 10,5)
Số xe	17	33	25	20	5

Khi đó $R = 10,5 - 0,5 = 10$

Câu 67. Bảng tần số ghép nhóm dưới đây thể hiện kết quả điều tra về tuổi thọ trung bình của nam giới ở 50 quốc gia.

Độ tuổi	[50;55)	[55;60)	[60;65)	[65;70)	[70;75)	[75;80)	[80;85)	[85;90)
Tần số	4	7	4	6	16	12	2	0

Hãy xác định khoảng biến thiên của tuổi thọ trung bình của nam giới trong mẫu số liệu ghép nhóm trên.

✎ **Lời giải**✓ **Trả lời: 35**

Do nhóm số liệu [85;90) có tần số là 0 nên ta sẽ chỉ xét đến nhóm số liệu [80;85).

Do đó: $R = 85 - 50 = 35$.

Câu 68. Bạn Trang thống kê lại chiều cao (đơn vị: cm) của các bạn học sinh nữ lớp 12C ở bảng sau:

Chiều cao	[155;160)	[160;165)	[165;170)	[170;175)	[175;180)	[180;185)
Tần số	2	7	12	3	0	1

Bạn Trang nhận xét như sau: Chênh lệch chiều cao của các bạn trong lớp không vượt quá $a(cm)$. Hãy xác định giá trị của a để nhận xét của Trang là đúng.

✎ **Lời giải**✓ **Trả lời: 30**

Ta có: $R = 185 - 155 = 30$.

Vậy giá trị của $a = 30$

Câu 69. Cô Hà thống kê lại đường kính thân gỗ của một số cây xoan đào 6 năm tuổi được trồng ở một lâm trường ở bảng sau.

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

Hãy xác định khoảng tứ phân vị của mẫu số liệu ghép nhóm trên (làm tròn đến hàng phần trăm).

✎ **Lời giải**✓ **Trả lời: 7,36**

Cỡ mẫu $n = 53$.

Gọi $x_1; x_2; \dots; x_{53}$ là mẫu số liệu gốc gồm các đường kính thân gỗ của một số cây xoan đào được xếp theo thứ tự không giảm.

Ta có: $x_1, \dots, x_5 \in [40; 45); x_6, \dots, x_{25} \in [45; 50); x_{26}, \dots, x_{43} \in [50; 55); x_{44}, \dots, x_{50} \in [55; 60)$.

Tứ phân vị thứ nhất của mẫu số liệu gốc là $x_{27} \in [50; 55)$.

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

$$Q_1 = 50 + \frac{\frac{53}{4} - 25}{18} \cdot (55 - 50) = \frac{3365}{72}.$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{40} + x_{41}) \in [50; 55)$.

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

$$Q_3 = 50 + \frac{\frac{3 \cdot 53}{4} - 25}{18} \cdot (55 - 50) = \frac{3895}{72}.$$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = \frac{3895}{72} - \frac{3365}{72} = \frac{265}{36} \approx 7,36$.

Câu 70. Kết quả đo chiều cao của 100 cây keo 3 năm tuổi tại một nông trường được cho ở bảng sau:

Chiều cao (m)	$[8, 4; 8, 6)$	$[8, 6; 8, 8)$	$[8, 9; 9, 0)$	$[9, 0; 9, 2)$	$[9, 2; 9, 4)$
Tần số	5	12	25	44	14

Hãy xác định khoảng tứ phân vị của mẫu số liệu ghép nhóm trên (làm tròn đến hàng phần trăm).

Lời giải

✓ **Trả lời: 0,23**

Cỡ mẫu $n = 100$.

Gọi $x_1; x_2; \dots; x_{100}$ là mẫu số liệu gốc gồm các chiều cao của 100 cây keo 3 năm tuổi được xếp theo thứ tự không giảm.

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

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_{50} + x_{51}) \in [9, 0; 9, 2)$.

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

$$Q_1 = 9,0 + \frac{\frac{100}{4} - 42}{44} \cdot (9,2 - 9,0) = \frac{1963}{220}.$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{75} \in [9, 0; 9, 2)$.

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

$$Q_3 = 9,0 + \frac{\frac{3 \cdot 100}{4} - 42}{44} \cdot (9,2 - 9,0) = \frac{183}{20}.$$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = \frac{183}{20} - \frac{1963}{220} = \frac{5}{22} \approx 0,23$.

Câu 71. Lương tháng của một số nhân viên một văn phòng được ghi lại như sau (đơn vị: triệu đồng):

Lương tháng (triệu đồng)	$[6; 8)$	$[8; 10)$	$[10; 12)$	$[12; 14)$
Số nhân viên	3	6	8	7

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

Lời giải

✓ **Trả lời: 3,29**

Gọi $x_1; x_2; x_3; \dots; x_{24}$ lần lượt là số nhân viên theo thứ tự không gian.

Do $x_1, \dots, x_3 \in [6; 8); x_4, \dots, x_9 \in [8; 10); x_{10}, \dots, x_{17} \in [10; 12); x_{18}, \dots, x_{24} \in [12; 14)$

Tứ phân vị thứ hai của dãy số liệu là $\frac{1}{2}(x_{12} + x_{13})$ thuộc nhóm $[10; 12)$ nên tứ phân vị thứ hai của mẫu số

$$\text{liệu là } Q_2 = 10 + \frac{\frac{24}{2} - 9}{8}(12 - 10) = 10,75.$$

Tứ phân vị thứ nhất của dãy số liệu là $\frac{1}{2}(x_6 + x_7)$ thuộc nhóm $[8; 10)$ nên tứ phân vị thứ nhất của mẫu số

$$\text{liệu là } Q_1 = 8 + \frac{\frac{24}{4} - 3}{6}(10 - 8) = 9.$$

Tứ phân vị thứ ba của dãy số liệu là $\frac{1}{2}(x_{18} + x_{19})$ thuộc nhóm $[12; 14)$ nên tứ phân vị thứ ba của mẫu số liệu

$$\text{là } Q_3 = 12 + \frac{\frac{3 \cdot 24}{4} - 17}{7}(14 - 12) = 12,29$$

Khoảng tứ phân vị là $\Delta Q = Q_3 - Q_1 = 12,29 - 9 = 3,29$.

Câu 72. Kết quả đo chiều cao của 100 cây keo 3 năm tuổi tại một nông trường được cho ở bảng sau:

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

Trong 100 cây keo trên có 1 cây cao 8,4 m. Hỏi giá trị ngoại lệ của mẫu số liệu là bao nhiêu?

✂ **Lời giải**

✓ **Trả lời: 8,4**

Cỡ mẫu $n = 100$.

Gọi $x_1; x_2; \dots; x_{100}$ là mẫu số liệu gốc về chiều cao của 100 cây keo 3 năm tuổi tại một nông trường được xếp theo thứ tự không giảm.

Ta có $x_1; \dots; x_5 \in [8,4; 8,6)$, $x_6; \dots; x_{17} \in [8,6; 8,8)$, $x_{18}; \dots; x_{42} \in [8,8; 9,0)$,

$x_{43}; \dots; x_{86} \in [9,0; 9,2)$, $x_{87}; \dots; x_{100} \in [9,2; 9,4)$.

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

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

$$Q_1 = 8,8 + \frac{\frac{100}{4} - (5 + 12)}{25} \cdot (9,0 - 8,8) = 8,86.$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{75} + x_{76}) \in [9,0; 9,2)$.

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

$$Q_3 = 9,0 + \frac{\frac{3 \cdot 100}{4} - (5 + 12 + 25)}{44} \cdot (9,2 - 9,0) = 9,15$$

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

$$\Delta_Q = Q_3 - Q_1 = 9,15 - 8,86 = 0,29.$$

Trong 100 cây keo trên có 1 cây cao 8,4m thuộc nhóm $[8,4; 8,6)$.

Vì $Q_1 - 1,5\Delta_Q = 8,86 - 1,5 \cdot 0,29 = 8,43 > 8,4$ nên chiều cao của cây keo cao 8,4m là giá trị ngoại lệ của mẫu số liệu ghép nhóm.

Câu 73. Thống kê số thẻ vàng của mỗi câu lạc bộ trong giải ngoại hạng Anh mùa giải 2021 – 2022 cho kết quả sau:

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

(Theo *premierleague.com*)

Hãy tìm khoảng biến thiên của mẫu số liệu ghép nhóm có độ dài bằng nhau với nhóm đầu tiên là $[40; 50)$?

Lời giải

✓ **Trả lời: 70**

Bảng số liệu ghép nhóm:

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

Khoảng biến thiên: $R = 110 - 40 = 70$.

Câu 74. Bạn Lộc thống kê lại chiều cao (đơn vị: cm) của các bạn nam lớp 12A và lớp 12B ở bảng sau.

Chiều cao (cm)	$[155; 160)$	$[160; 165)$	$[165; 170)$	$[175; 180)$	$[180; 185)$
Lớp 12A	3	5	12	1	0
Lớp 12B	6	10	7	0	2

Khoảng biến thiên chiều cao của học sinh nam lớp 12B có độ phân tán lớn hơn khoảng biến thiên chiều cao của học sinh nam lớp 12A bao nhiêu cm ?

Lời giải

✓ **Trả lời: 5**

Khoảng biến thiên của chiều cao của các bạn nam lớp 12A là: $180 - 155 = 25 (cm)$.

Khoảng biến thiên của chiều cao của các bạn nam lớp 12B là: $185 - 155 = 30 (cm)$.

Vậy khoảng biến thiên chiều cao của học sinh nam lớp 12B có độ phân tán lớn hơn khoảng biến thiên chiều cao của học sinh nam lớp 12A là: $30 - 25 = 5 (cm)$.

Câu 75. Thống kê số thẻ vàng của mỗi câu lạc bộ trong giải ngoại hạng Anh mùa giải 2021 – 2022 cho kết quả như sau:

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

Tính khoảng biến của mẫu số liệu ghép nhóm dãy số liệu trên thành các nhóm có độ dài bằng nhau với nhóm đầu tiên là $[40; 50)$.

Lời giải

✓ **Trả lời: 16**

Bảng số liệu ghép nhóm:

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

Cỡ mẫu $n = 20$.

Gọi $x_1; x_2; \dots; x_{20}$ là số thẻ vàng của mỗi câu lạc bộ trong giải ngoại hạng Anh mùa giải 2021 – 2022, các giá trị này đã được xếp theo thứ tự không giảm.

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{x_5 + x_6}{2} \in [50; 60)$.

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

$$Q_1 = 50 + \frac{\frac{20}{4} - 2}{5} \cdot 10 = 56.$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{x_{15} + x_{16}}{2} \in [70; 80)$.

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

$$Q_3 = 70 + \frac{\frac{3.20}{4} - (2+5+7)}{5} \cdot 10 = 72.$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 16.$

Câu 76. Bảng tần số ghép nhóm dưới đây thể hiện kết quả điều tra về tuổi thọ trung

Nhóm (Tuổi thọ)	[50;55)	[55;60)	[60;65)	[65;70)	[70;75)	[75;80)	[80;85)	[85;90)
Nam	4	7	4	6	15	12	2	0
Nữ	3	4	5	3	7	14	13	1

Tính khoảng tứ phân vị của mẫu số liệu ghép nhóm có tuổi thọ trung bình đồng đều nhất.

Lời giải

✓ **Trả lời: 3,75**

Cỡ mẫu $n = 50.$

Gọi $x_1; x_2; \dots; x_{50}$ là mẫu số liệu gốc về tuổi thọ trung bình của nam giới ở 50 quốc gia được xếp theo thứ tự không giảm.

Ta có: $x_1; x_2; \dots; x_4 \in [50;55); x_5; x_6; \dots; x_{11} \in [55;60); x_{12}; x_{13}; \dots; x_{15} \in [60;65); x_{16}; x_{17}; \dots; x_{21} \in [65;70); x_{22}; x_{23}; \dots; x_{36} \in [70;75); x_{37}; x_{38}; \dots; x_{48} \in [75;80); x_{49}; x_{50} \in [80;85)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $x_{13} \in [60;65).$

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

$$Q_1 = 60 + \frac{\frac{50}{4} - (4+7)}{4} (65 - 60) = 71,875.$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{38} \in [75;80).$

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

$$Q_3 = 75 + \frac{\frac{3.50}{4} - (4+7+4+6+15)}{12} (80 - 75) = 75,625.$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 3,75.$

Gọi $y_1; y_2; \dots; y_{50}$ là mẫu số liệu gốc về tuổi thọ trung bình của nữ giới ở 50 quốc gia được xếp theo thứ tự không giảm.

Ta có: $y_1; y_2; y_3 \in [50;55); y_4; \dots; y_7 \in [55;60); y_8; \dots; y_{12} \in [60;65); y_{13}; \dots; y_{15} \in [65;70); y_{16}; \dots; y_{22} \in [70;75); y_{23}; \dots; y_{36} \in [75;80); y_{37}; \dots; y_{49} \in [80;85); y_{50} \in [85;90).$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $y_{13} \in [65;70).$

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

$$Q'_1 = 65 + \frac{\frac{50}{4} - (3+4+5)}{3} (70 - 65) = \frac{395}{6}.$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $y_{38} \in [80;85).$

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

$$Q'_3 = 80 + \frac{\frac{3.50}{4} - (3+4+5+3+7+14)}{13} (85 - 80) = \frac{2095}{26}.$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_{Q'} = Q'_3 - Q'_1 = \frac{575}{39}.$

Suy ra $\Delta_{Q'} > \Delta_Q$ nên độ tuổi trung bình của nam giới đồng đều hơn.

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm có tuổi thọ trung bình đồng đều nhất là 3,75.



Câu 77. Bảng thống kê sau cho biết điện năng tiêu thụ của 30 hộ ở một khu dân cư trong một tháng như sau (đơn vị: kW):

50	47	30	65	63	70	38	34	48	53	33	39	32	40	50
55	50	61	37	37	43	35	65	60	31	33	41	45	55	59

Chuyển mẫu số liệu trên sang mẫu số liệu ghép nhóm gồm 8 nhóm có độ dài bằng nhau và bằng 5 như sau:

Giá trị	[30;35)	[35;40)	[40;45)	[45;50)	[50;55)	[55;60)	[60;65)	[65;70]
Số lượng	6	5	3	x	4	3	y	z

Tính giá trị của $S = 43x + 47y + 1000z$

✎ **Lời giải**

✓ **Trả lời: 3270**

Giá trị nhỏ nhất là 30, giá trị lớn nhất là 70 nên khoảng biến thiên là $70 - 30 = 40$.

Tổng độ dài của 8 nhóm là 40 nên ta được các nhóm như sau: $[30;35), [35;40), [40;45), [45;50), [50;55), [55;60), [60;65), [65;70]$.

Đếm số giá trị thuộc mỗi nhóm ta có mẫu số liệu ghép nhóm như sau:

Giá trị	[30;35)	[35;40)	[40;45)	[45;50)	[50;55)	[55;60)	[60;65)	[65;70]
Số lượng	6	5	3	3	4	3	3	3

Vậy $S = 43x + 47y + 1000z = 43 \cdot 3 + 47 \cdot 3 + 1000 \cdot 3 = 3270$.

Câu 78. Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một cửa hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

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

Số trung bình của mẫu số liệu trên bằng bao nhiêu?

✎ **Lời giải**

✓ **Trả lời: 9,4**

Bảng tần số ghép nhóm theo giá trị đại diện là

Doanh thu	[5;7)	[7;9)	[9;11)	[11;13)	[13;15)
Giá trị đại diện	6	8	10	12	14
Số ngày	2	7	7	3	1

Số trung bình: $\bar{x} = \frac{2 \cdot 6 + 7 \cdot 8 + 7 \cdot 10 + 3 \cdot 12 + 1 \cdot 14}{20} = 9,4$

Câu 79. Điều tra 42 học sinh của một lớp 12 về số giờ tự học ở nhà, người ta có bảng thống kê sau:

Số giờ tự học	[1;2)	[2;3)	[3;4)	[4;5)	[5;6)
Số học sinh	8	10	12	9	3

Tính phương sai của mẫu số liệu ghép nhóm trên.

✎ **Lời giải**

✓ **Trả lời: 1,431**

Chọn giá trị đại diện cho mẫu số liệu, ta có:

Số giờ tự học	[1;2)	[2;3)	[3;4)	[4;5)	[5;6)
Giá trị đại diện	1,5	2,5	3,5	4,5	5,5
Số học sinh	8	10	12	9	3

Số giờ học trung bình là:

$\bar{x} = \frac{8 \cdot 1,5 + 10 \cdot 2,5 + 12 \cdot 3,5 + 9 \cdot 4,5 + 3 \cdot 5,5}{42} = \frac{68}{21} \approx 3,238$.

Phương sai là:

$S^2 = \frac{1}{42} [8 \cdot (1,5)^2 + 10 \cdot (2,5)^2 + 12 \cdot (3,5)^2 + 9 \cdot (4,5)^2 + 3 \cdot (5,5)^2] - \left(\frac{68}{21} \right)^2 = \frac{2525}{1764} \approx 1,431$.

Câu 80. Thống kê điểm trắc nghiệm môn Tiếng Anh của 40 học sinh, người ta có bảng sau:

Điểm	[20;30)	[30;40)	[40;50)	[50;60)	[60;70)	[70;80)	[80;90)	[90;100)
Số học sinh	3	5	5	8	7	5	3	4

Tính phương sai của mẫu số liệu ghép nhóm trên.

Lời giải

✓ **Trả lời: 404,75**

Chọn giá trị đại diện cho mẫu số liệu, ta có:

Điểm	[20;30)	[30;40)	[40;50)	[50;60)	[60;70)	[70;80)	[80;90)	[90;100)
Giá trị đại diện	25	35	45	55	65	75	85	95
Số học sinh	3	5	5	8	7	5	3	4

Điểm trung bình là:

$$\bar{x} = \frac{3 \cdot 25 + 5 \cdot 35 + 5 \cdot 45 + 8 \cdot 55 + 7 \cdot 65 + 5 \cdot 75 + 3 \cdot 85 + 4 \cdot 95}{40} = 59,5.$$

Phương sai là:

$$S^2 = \frac{1}{40} \left[3 \cdot (25)^2 + 5 \cdot (35)^2 + 5 \cdot (45)^2 + 8 \cdot (55)^2 + 7 \cdot (65)^2 + 5 \cdot (75)^2 + 3 \cdot (85)^2 + 4 \cdot (95)^2 \right] - (59,5)^2 = 404,75$$

Câu 81. Thống kê doanh thu (đơn vị: triệu đô la) của 20 công ty sản xuất ô tô trong năm 2023, người ta có bảng sau:

Doanh thu	[0;20)	[20;40)	[40;60)	[60;80)	[80;100)
Số công ty	5	5	6	2	2

Tính độ lệch chuẩn của mẫu số liệu ghép nhóm trên.

Lời giải

✓ **Trả lời: 24,88**

Chọn giá trị đại diện cho mẫu số liệu, ta có:

Doanh thu	[0;20)	[20;40)	[40;60)	[60;80)	[80;100)
Giá trị đại diện	10	30	50	70	90
Số công ty	5	5	6	2	2

Điểm trung bình là:

$$\bar{x} = \frac{5 \cdot 10 + 5 \cdot 30 + 6 \cdot 50 + 2 \cdot 70 + 2 \cdot 90}{20} = 41.$$

Phương sai là:

$$S^2 = \frac{1}{20} \left[5 \cdot (10)^2 + 5 \cdot (30)^2 + 6 \cdot (50)^2 + 2 \cdot (70)^2 + 2 \cdot (90)^2 \right] - (41)^2 = 619.$$

$$\text{Độ lệch chuẩn: } S = \sqrt{619} \approx 24,88.$$

Câu 82. Điều tra về số tiền mua sách (đơn vị: nghìn đồng) trong một năm của 50 học sinh trong một trường THPT, người ta có bảng sau:

Số tiền mua sách	[0;200)	[200;400)	[400;600)	[600;800)	[800;1000)
Số học sinh	29	11	3	4	3

Tính độ lệch chuẩn của mẫu số liệu ghép nhóm trên.

Lời giải

✓ **Trả lời: 242,29**

Chọn giá trị đại diện cho mẫu số liệu, ta có:



Số tiền mua sách	[0; 200)	[200; 400)	[400; 600)	[600; 800)	[800; 1000)
Giá trị đại diện	100	300	500	700	900
Số học sinh	29	11	3	4	3

Điểm trung bình là:

$$\bar{x} = \frac{29 \cdot 100 + 11 \cdot 300 + 3 \cdot 500 + 4 \cdot 700 + 3 \cdot 900}{50} = 264.$$

Phương sai là:

$$S^2 = \frac{1}{50} \left[29 \cdot (100)^2 + 11 \cdot (300)^2 + 3 \cdot (500)^2 + 4 \cdot (700)^2 + 3 \cdot (900)^2 \right] - (264)^2 = 58704.$$

Độ lệch chuẩn: $S = \sqrt{58704} \approx 242,29$.

Câu 83. Tốc độ của 20 xe hơi khi đi qua một trạm kiểm tra tốc độ (đơn vị: km/h) được thống kê lại như sau.

Hãy tính phương sai của mẫu số liệu ghép nhóm với nhóm đầu tiên là [42; 46) và độ dài mỗi nhóm bằng 4.

(làm tròn đến hàng phần mười)

42	43,4	43,4	46,5	46,7	46,8	47,5	47,7	48,1	48,4
50,8	52,1	52,7	53,9	54,8	55,6	57,5	59,6	60,3	61,1

Lời giải

✓ **Trả lời: 26,6**

Ta lập được bảng số liệu ghép nhóm theo giá trị đại diện như sau:

Tốc độ (km/h)	[42; 46)	[46; 50)	[50; 54)	[54; 58)	[58; 62)
Giá trị đại diện	44	48	52	56	60
Số xe	3	7	4	3	3

Số trung bình của mẫu số liệu ghép nhóm là:

$$\bar{x} = \frac{3 \cdot 44 + 7 \cdot 48 + 4 \cdot 52 + 3 \cdot 56 + 3 \cdot 60}{20} = 51,2$$

Phương sai của mẫu số liệu ghép nhóm là:

$$S^2 = \frac{1}{20} (3 \cdot 44^2 + 7 \cdot 48^2 + 4 \cdot 52^2 + 3 \cdot 56^2 + 3 \cdot 60^2) - (51,2)^2 \approx 26,6$$

Câu 84. Một giống cây xoan đào được trồng tại hai địa điểm A và B. Người ta thống kê đường kính thân của một số cây xoan đào 5 năm tuổi ở bảng sau. Gọi phương sai đường kính thân của một số cây xoan đào 5 năm tuổi ở địa điểm A và địa điểm B lần lượt là S_A^2 và S_B^2 . Tính $S_A^2 - S_B^2$ bằng bao nhiêu?

Đường kính (cm)	[30; 32)	[32; 34)	[34; 36)	[36; 38)	[38; 40)
Số cây trồng ở địa điểm A	25	38	20	10	9
Số cây trồng ở địa điểm B	22	27	19	14	14

Lời giải

✓ **Trả lời: 1,91**

Ta lập bảng theo giá trị đại diện như sau:

Đường kính (cm)	[30; 32)	[32; 34)	[34; 36)	[36; 38)	[38; 40)
Giá trị đại diện	31	33	35	37	39
Số cây trồng ở địa điểm A	25	38	20	10	9
Số cây trồng ở địa điểm B	22	27	19	14	14

Cỡ mẫu: $n_A = 25 + 38 + 20 + 10 + 9 = 100$; $n_B = 22 + 27 + 19 + 14 + 14 = 100$

Đường kính trung bình của thân cây xoan đào trồng tại địa điểm A là:

$$\bar{x}_A = \frac{25 \cdot 31 + 38 \cdot 33 + 20 \cdot 35 + 10 \cdot 37 + 9 \cdot 39}{100} = 33,72$$

Phương sai của mẫu số liệu ghép nhóm về đường kính của thân cây xoan đào trồng tại địa điểm A là:

$$S_A^2 = \frac{1}{100} (25 \cdot 31^2 + 38 \cdot 33^2 + 20 \cdot 35^2 + 10 \cdot 37^2 + 9 \cdot 39^2) - (33,72)^2 \approx 5,40$$

Đường kính trung bình của thân cây xoan đào trồng tại địa điểm B là:

$$\bar{x}_B = \frac{22 \cdot 31 + 27 \cdot 33 + 19 \cdot 35 + 18 \cdot 37 + 14 \cdot 39}{100} = 34,5$$

Phương sai của mẫu số liệu ghép nhóm về đường kính của thân cây xoan đào trồng tại địa điểm B là:

$$S_B^2 = \frac{1}{100} (22 \cdot 31^2 + 27 \cdot 33^2 + 19 \cdot 35^2 + 18 \cdot 37^2 + 14 \cdot 39^2) - (34,5)^2 = 7,31$$

$$\text{Vậy } |S_A^2 - S_B^2| = |5,40 - 7,31| = 1,91$$

Câu 85. Thời gian hoàn thành một bài viết chính tả của một số học sinh lớp 4 hai trường X và Y được ghi lại ở bảng sau. Gọi độ lệch chuẩn thời gian hoàn thành một bài viết chính tả của một số học sinh lớp 4 ở trường X và Y lần lượt là S_X và S_Y . Tính $S_X - S_Y$ bằng bao nhiêu?

Thời gian (Phút)	[6;7)	[7;8)	[8;9)	[9;10)	[10;11)
Số học sinh trường X	8	10	13	10	9
Số học sinh trường Y	4	12	17	14	3

Lời giải

✓ **Trả lời: 0,29**

Ta lập bảng theo giá trị đại diện như sau:

Thời gian (Phút)	[6;7)	[7;8)	[8;9)	[9;10)	[10;11)
Giá trị đại diện	6,5	7,5	8,5	9,5	10,5
Số học sinh trường X	8	10	13	10	9
Số học sinh trường Y	4	12	17	14	3

Cỡ mẫu $n_X = 8 + 10 + 13 + 10 + 9 = 50$, $n_Y = 4 + 12 + 17 + 14 + 3 = 50$.

Thời gian trung bình hoàn thành một bài viết chính tả của học sinh trường X là:

$$\bar{x}_X = \frac{8 \cdot 6,5 + 10 \cdot 7,5 + 13 \cdot 8,5 + 10 \cdot 9,5 + 9 \cdot 10,5}{50} = 8,54$$

Phương sai của mẫu số liệu ghép nhóm của trường X là:

$$S_X^2 = \frac{1}{50} [8 \cdot (6,5)^2 + 10 \cdot (7,5)^2 + 13 \cdot (8,5)^2 + 10 \cdot (9,5)^2 + 9 \cdot (10,5)^2] - (8,54)^2 = 1,76$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm của trường X là:

$$S_X = \sqrt{S_X^2} = \sqrt{1,76} \approx 1,33$$

Thời gian trung bình hoàn thành một bài viết chính tả của học sinh trường Y là:

$$\bar{x}_Y = \frac{4 \cdot 6,5 + 12 \cdot 7,5 + 17 \cdot 8,5 + 14 \cdot 9,5 + 3 \cdot 10,5}{50} = 8,5$$

Phương sai của mẫu số liệu ghép nhóm của trường Y là:

$$S_Y^2 = \frac{1}{50} [4 \cdot (6,5)^2 + 12 \cdot (7,5)^2 + 17 \cdot (8,5)^2 + 14 \cdot (9,5)^2 + 3 \cdot (10,5)^2] - (8,5)^2 = 1,08$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm của trường Y là:

$$S_Y = \sqrt{S_Y^2} = \sqrt{1,08} \approx 1,04$$

$$\text{Vậy } S_X - S_Y = 1,33 - 1,04 = 0,29$$

Câu 86. Kết quả 40 lần nhảy xa của vận động viên HUY được thống kê trong bảng sau (đơn vị: mét):

Nhóm	Tần số
[6, 22; 6, 46)	3
[6, 46; 6, 70)	7
[6, 70; 6, 94)	5
[6, 94; 7, 18)	20
[7, 18; 7, 42)	5
	$n = 40$

Độ lệch chuẩn của mẫu số liệu ghép nhóm trên là bao nhiêu? (làm tròn kết quả đến hàng phần trăm)

Lời giải

✓ **Trả lời: 0,26**

Ta có 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
		$n = 40$

Số trung bình cộng của mẫu số liệu trên là

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

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 số liệu trên (làm tròn kết quả đến hàng phần trăm) là:

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

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

Câu 87. Anh An đầu tư số tiền sử dụng vào hai lĩnh vực kinh doanh A, B . Anh An thống kê số tiền thu được mỗi tháng trong vòng 60 tháng theo từng 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	20	5	10	5	20
Số tháng đầu tư vào lĩnh vực B	5	10	30	10	5

So sánh giá trị trung bình và độ lệch chuẩn của số tiền thu được mỗi tháng khi bắt đầu tư vào từng lĩnh vực A, B . Kí hiệu $\Delta = S_A - S_B$ nếu $\Delta > 0$ thì đầu tư vào lĩnh vực A “rủi ro” hơn, ngược lại nếu $\Delta < 0$ thì đầu tư vào lĩnh vực B “rủi ro” hơn. Tính $\Delta = S_A - S_B$ (làm tròn kết quả đến hàng phần trăm).

Lời giải

✓ **Trả lời: 3,42**

Chọn giá trị đại diện cho các nhóm dữ liệu có:

Giá trị đại diện	7,5	12,5	17,5	22,5	27,5
Số tháng đầu tư vào lĩnh vực B	20	5	10	5	20
Số tháng đầu tư vào lĩnh vực A	5	10	30	10	5

Xét mẫu số liệu của lĩnh vực A :

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_A = \frac{20 \cdot 7,5 + 5 \cdot 12,5 + 10 \cdot 17,5 + 5 \cdot 22,5 + 20 \cdot 27,5}{60} = 17,5 \text{ (triệu đồng)}$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_A^2 = \frac{1}{60} (20 \cdot 7,5^2 + 5 \cdot 12,5^2 + 10 \cdot 17,5^2 + 5 \cdot 22,5^2 + 20 \cdot 27,5^2) - (17,5)^2 = 70,83$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_A = \sqrt{S_A^2} = \sqrt{70,83} \approx 8,42$.

Xét mẫu số liệu của lĩnh vực B :

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_B = \frac{5 \cdot 7,5 + 10 \cdot 12,5 + 30 \cdot 17,5 + 10 \cdot 22,5 + 5 \cdot 27,5}{60} = 17,5 \text{ (triệu đồng)}$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_B^2 = \frac{1}{60} (5 \cdot 7,5^2 + 10 \cdot 12,5^2 + 30 \cdot 17,5^2 + 10 \cdot 22,5^2 + 5 \cdot 27,5^2) - (17,5)^2 = 25$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_B = \sqrt{S_B^2} = \sqrt{25} = 5$.

$$\Rightarrow \Delta = S_A - S_B = 8,42 - 5 = 3,42$$

Câu 88. 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

So sánh giá trị trung bình và độ lệch chuẩn của số tiền thu được mỗi tháng khi bắt đầu tư vào từng lĩnh vực A, B. Kí hiệu $\Delta = S_A - S_B$ nếu $\Delta > 0$ thì đầu tư vào lĩnh vực A “rủi ro” hơn, ngược lại nếu $\Delta < 0$ thì đầu tư vào lĩnh vực B “rủi ro” hơn. Tính $\Delta = S_A - S_B$ (làm tròn kết quả đến hàng phần trăm).

Lời giải

✓ **Trả lời: -0,8**

Ta có bảng thống kê giá đóng cửa theo giá trị đại diện:

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

Xét mẫu số liệu của cổ phiếu A :

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_A = \frac{8 \cdot 121 + 9 \cdot 123 + 12 \cdot 125 + 10 \cdot 127 + 11 \cdot 129}{50} = 125,28.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_A^2 = \frac{1}{50} (8 \cdot 121^2 + 9 \cdot 123^2 + 12 \cdot 125^2 + 10 \cdot 127^2 + 11 \cdot 129^2) - (125,28)^2 \approx 7,5.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_A = \sqrt{S_A^2} = \sqrt{7,52} \approx 2,7$.

Xét mẫu số liệu của cổ phiếu B :

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x}_B = \frac{16 \cdot 121 + 4 \cdot 123 + 3 \cdot 125 + 6 \cdot 127 + 21 \cdot 129}{50} = 125,28.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_B^2 = \frac{1}{50} (16 \cdot 121^2 + 4 \cdot 123^2 + 3 \cdot 125^2 + 6 \cdot 127^2 + 21 \cdot 129^2) - (125,28)^2 \approx 12,4$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_B = \sqrt{S_B^2} = \sqrt{12,4} \approx 3,5$.

$$\Rightarrow \Delta = S_A - S_B = 2,7 - 3,5 = -0,8$$

**ĐỀ THAM KHẢO KIỂM TRA THƯỜNG XUYÊN****TRƯỜNG CÔNG VÕ**
BÀI TẬP THAM KHẢO**ĐỀ KIỂM TRA THƯỜNG XUYÊN 3****Môn thi: TOÁN – Lớp 12**

Thời gian làm bài: ... phút (không kể thời gian giao đề)

PHẦN I. CÂU TRẮC NGHIỆM NHIỀU PHƯƠNG ÁN LỰA CHỌN (Mỗi câu hỏi, học sinh chỉ chọn một phương án)**Câu 1.** Đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

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

Khoảng biến thiên của mẫu số liệu ghép nhóm là

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

Câu 2. Bạn Chi rất thích nhảy hiện đại. Thời gian tập nhảy mỗi ngày trong thời gian gần đây của bạn Chi được thống kê lại ở bảng sau:

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

Khoảng biến thiên của mẫu số liệu ghép nhóm là

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

Câu 3. Một vườn thú ghi lại tuổi thọ (đơn vị: năm) của 20 con hổ và thu được kết quả như sau:

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

Nhóm chứa tứ phân vị thứ nhất là

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

Câu 4. Cho mẫu số liệu ghép nhóm về tuổi thọ (đơn vị tính là năm) của một loại bóng đèn mới như sau.

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

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

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

Câu 5. Khảo sát thời gian tập thể dục trong ngày của một số học sinh khối 11 thu được mẫu số liệu ghép nhóm sau:

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

Nhóm chứa trung vị là

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

Câu 6. Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một cửa hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

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

Số trung bình của mẫu số liệu trên thuộc khoảng nào trong các khoảng dưới đây?

- A.** $[7; 9)$. **B.** $[9; 11)$. **C.** $[11; 13)$. **D.** $[13; 15)$.

Câu 7. Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khỏe. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

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

Phương sai của mẫu số liệu ghép nhóm là

- A. 3,39. B. 11,62. C. 0,1314. D. 0,36.

Câu 8. Dũng là học sinh rất giỏi chơi rubik, bạn có thể giải nhiều loại khối rubik khác nhau. Trong một lần tập luyện giải khối rubik 3×3 , bạn Dũng đã tự thống kê lại thời gian giải rubik trong 25 lần giải liên tiếp ở bảng sau:

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

Độ lệch chuẩn của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

- A. 5,98. B. 6. C. 2,44. D. 2,5.

Câu 9. Thời gian (phút) truy cập Internet mỗi buổi tối của một số học sinh được cho trong bảng sau:

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

Tổng các giá trị đại diện của 3 nhóm ở giữa là

- A. 49 B. 51 C. 53 D. 55

PHẦN II. CÂU TRẮC NGHIỆM ĐÚNG SAI (Trong mỗi ý a), b), c), d) ở mỗi câu, học sinh chọn Đúng hoặc Sai.)

Câu 10. 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

Các mệnh đề sau đúng hay sai?

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

Câu 11. 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) Khoảng biến thiên của mẫu số liệu: 375,9(mm)
b) Khoảng tứ phân vị của mẫu số liệu là: $\Delta_Q = \frac{1827}{100}$
c) Chia mẫu số liệu trên thành 4 nhóm, nhóm đầu tiên là [140; 240). Khoảng biến thiên của mẫu số liệu ghép nhóm: 400(mm)

d) Chia mẫu số liệu trên thành 4 nhóm, nhóm đầu tiên là $[140; 240)$. Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q' = \frac{1000}{7}$

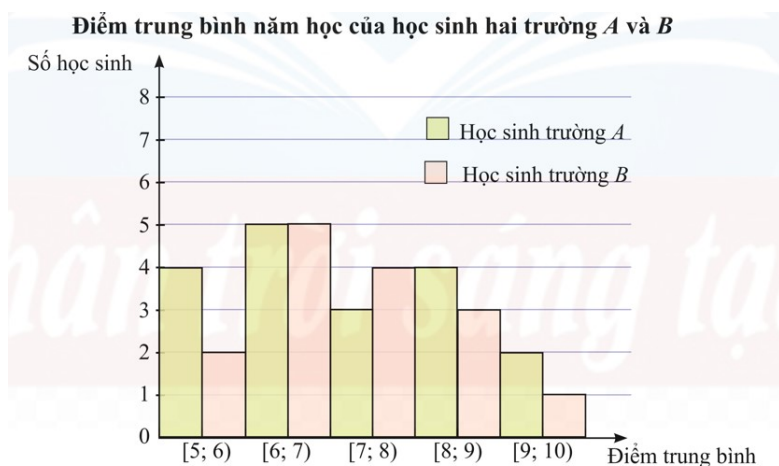
Câu 12. Kết quả đo chiều cao của 100 cây keo 3 năm tuổi tại một nông trường được cho ở bảng sau:

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

Các mệnh đề sau đúng hay sai?

- a) Khoảng biên thiên của mẫu số liệu ghép nhóm: $1(m)$
- b) Tứ phân vị thứ ba của mẫu số liệu ghép nhóm là: $Q_3 = 10,5$
- c) Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = 2,06$
- d) Trong 100 cây keo trên có 1 cây cao $8,4m$. Thì chiều cao của cây keo này là giá trị ngoại lệ

Câu 13. Biểu đồ sau mô tả kết quả điều tra về điểm trung bình năm học của học sinh hai trường A và B.



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: 1,73
- c) Nếu so sánh theo khoảng tứ phân vị của mẫu số liệu ghép nhóm thì học sinh trường B có điểm trung bình đồng đều hơn
- d) Nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì học sinh trường A có điểm trung bình đồng đều hơn

Câu 14. 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 A ta có phương sai của mẫu số liệu ghép nhóm là 7,5216
- b) Xét mẫu số liệu của cổ phiếu B ta có số trung bình của mẫu số liệu ghép nhóm là 115,28.
- c) Xét mẫu số liệu của cổ phiếu B ta có độ lệch chuẩn của mẫu số liệu ghép nhóm là $\sqrt{15,4096}$

d) Người ta có thể dùng phương sai và độ lệch chuẩn để so sánh mức độ rủi ro của các loại cổ phiếu có giá trị trung bình gần bằng nhau. Cổ phiếu nào có phương sai, độ lệch chuẩn cao hơn thì được coi là có độ rủi ro lớn hơn. Theo quan điểm trên, thì cổ phiếu A có độ rủi ro thấp hơn cổ phiếu B .

Câu 15. 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 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 .

PHẦN III. CÂU TRẮC NGHIỆM TRẢ LỜI NGẮN (Học sinh trả lời đáp số)

Câu 16. Điểm thi môn Toán (thang điểm 100, điểm được làm tròn đến 1) của 60 thí sinh được cho mẫu số liệu ghép nhóm như sau:

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

Tính khoảng tứ phân vị của mẫu số liệu ghép nhóm

Câu 17. Thời gian sử dụng điện thoại trong một ngày của 30 sinh viên được ghi lại ở bảng sau (đơn vị: phút).

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

Trong các số nguyên 1, 3, 250, 290, 296, 298 có bao nhiêu số là giá trị ngoại lệ?

Câu 18. Biểu đồ dưới đây biểu diễn số lượt khách hàng đặt bàn qua hình thức trực tuyến mỗi ngày trong quý III năm 2022 của một nhà hàng. Cột thứ nhất biểu diễn số ngày có từ 1 đến dưới 6 lượt đặt bàn; cột thứ hai biểu diễn số ngày có từ 6 đến dưới 11 lượt đặt bàn;...

Trong có số nguyên: 1; 2; 3; 4; 5; 6; 21; 22; 23; 24; 25, có bao nhiêu số là giá trị ngoại lệ?

