CHUYÊN ĐỀ THỐNG KÊ CỦA MẪU SỐ LIỆU GHÉP NHÓM

I. KIẾN THỨC CẦN NHỚ:

1. SỐ TRUNG BÌNH CỦA MẪU SỐ LIỆU GHÉP NHÓM

Số trung bình của mẫu số liệu ghép nhóm kí hiệu là $\bar{x}$

$$\bar{x}=\frac{m_1x_1+...+m_kx_k}{n}$$

Trong đó, $n=m_1+...+m_k$ là cỡ mẫu và $x_i=\frac{a_i+a_{i+1}}{2}$ (với $i=1,...,k$) là giá trị đại diện của nhóm $[a_i;a_{i+1})$

2. TRUNG VỊ CỦA MẪU SỐ LIỆU GHÉP NHÓM

Để tính trung vị của mẫu số liệu ghép nhóm, ta làm như sau:

Bước 1. Xác định nhóm chứa trung vị. Giả sử đó là nhóm thứ *p*: $[a_p;a_{p+1})$.

Bước 2. Trung vị là $M_e=a_p+\frac{\frac{n}{2}-(m_1+...+m_{p-1})}{m_p}.(a_{p+1}-a_p)$, trong đó *n* là cỡ mẫu, m_p là tần số nhóm *p*. Với $p=1$, ta quy ước $m_1+...+m_{p-1}=0$.

3. TỨ PHÂN VỊ CỦA MẪU SỐ LIỆU GHÉP NHÓM

Để tính tứ phân vị thứ nhất Q_1 của mẫu số liệu ghép nhóm, trước hết ta xác định nhóm chứa Q_1, giả sử đó là nhóm thứ $p:[a_p;a_{p+1})$. Khi đó

$Q_1=a_p+\frac{\frac{n}{4}-(m_1+...+m_{p-1})}{m_p}.(a_{p+1}-a_p)$, trong đó n là cỡ mẫu, m_p là tần số nhóm p, với $p=1$, ta quy ước $m_1+...+m_{p-1}=0$.

Để tính tứ phân vị thứ ba Q_3 của mẫu số liệu ghép nhóm, trước hết ta xác định nhóm chứa Q_3, giả sử đó là nhóm thứ $p:[a_p;a_{p+1})$. Khi đó

$Q_3=a_p+\frac{\frac{3n}{4}-(m_1+...+m_{p-1})}{m_p}.(a_{p+1}-a_p)$, trong đó *n* là cỡ mẫu, m_p là tần số nhóm p, với $p=1$, ta quy ước $m_1+...+m_{p-1}=0$.

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

4. MỐT CỦA MẪU SỐ LIỆU GHÉP NHÓM

Để tìm môt của mẫu số liệu ghép nhóm, ta thự hiện theo các bước sau:

Bước 1. Xác định nhóm có tần số lớn nhất (gọi là nhóm chứa mốt), giả sử là nhóm *j*: $[a_j;a_{j+1})$.

Bước 2. Mốt được xác định là $M_0=a_j+\frac{m_j-m_{j-1}}{(m_j-m_{j-1})+(m_j-m_{j+1})}.h$, trong đó m_j là tần số nhóm *j* (quy ước $m_0=m_{k+1}=0$) và *h* là độ dài của nhóm.

5. KHOẢNG TỨ PHÂN VỊ

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

Ý nghĩa. Khoảng tứ phân vị của mẫu số liệu ghép nhóm xấp xỉ cho khoảng tứ phân vị của mẫu số liệu gốc. Khoảng tứ phân vị cũng được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm. Khoảng tứ phân vị càng lớn thì mẫu số liệu càng phân tán.

6. PHƯƠNG SAI VÀ ĐỘ LỆCH CHUẨN

Phương sai của mẫu số liệu ghép nhóm, kí hiệu là s^2, là một số được tính theo công thức sau:

$$s^2 = \frac{m_1\left(x_1 - \overline{x}\right)^2 + \ldots + m_k\left(x_k - \overline{x}\right)^2}{n}$$

trong đó, $n = m_1 + \ldots + m_k ; x_i = \frac{a_i + a_{i+1}}{2}$ với $i = 1,2,\ldots,k$ là giá trị đại diện cho nhóm $\left[a_i; a_{i+1}\right)$ và $\overline{x} = \frac{m_1 x_1 + \ldots + m_k x_k}{n}$ là số trung bình của mẫu số liệu ghép nhóm.

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

Ý nghĩa. Phương sai, độ lệch chuẩn của mẫu số liệu ghép nhóm là các xấp xỉ cho phương sai, độ lệch chuẩn của mẫu số liệu gốc. Chúng được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm xung quanh số trung bình của mẫu số liệu đó. Phương sai, độ lệch chuẩn càng lớn thì mẫu số liệu càng phân tán.

7. SỬ DỤNG PHƯƠNG SAI, ĐỘ LỆCH CHUẨN ĐO ĐỘ RỦI RO

Để so sánh độ phân tán của hai mẫu số liệu khi đơn vị đo trên hai mẫu số liệu khác nhau hoặc giá trị trung bình của hai mẫu số liệu này khác nhau rất nhiều người ta dùng hệ số biến thiên CV (Coefficient of Variation). Hệ số biến thiên được tính theo công thức:

$$CV = \frac{s}{\overline{x}}$$

trong đó s là độ lệch chuẩn và $\overline{x}$ là số trung bình của mẫu số liệu.

II. CÁC DẠNG BÀI TẬP MINH HỌA:

1. BÀI TẬP MẪU SỐ LIỆU GHÉP NHÓM

PHẦN 1: Trắc nghiệm

Câu 1. (NB) Mỗi nhóm số liệu ghép nhóm là tập hợp gồm:

A. các giá trị của số liệu được ghép nhóm theo nhiều tiêu chí xác định.
B. các giá trị của số liệu được ghép nhóm theo hai tiêu chí xác định.
C. các giá trị của số liệu được ghép nhóm theo một tiêu chí xác định.
D. các giá trị của số liệu được ghép nhóm theo ba tiêu chí xác định.

Lời giải

Chọn C

Mỗi nhóm số liệu là tập hợp gồm các giá trị của số liệu được ghép nhóm theo một tiêu chí xác định.

Câu 2.(NB) Cho mẫu số liệu ghép nhóm về thời gian (phút) đi từ nhà đến nơi làm việc của các nhân viên một công ty như sau:

Thời gian	$[15;20)$	$[20;25)$	$[25;30)$	$[30;35)$	$[35;40)$	$[40;45)$	$[45;50)$
Số nhân viên	6	14	25	37	21	13	9

Mẫu số liệu được chia thành bao nhiêu nhóm?

A. 6 nhóm **B.** 5 nhóm **C.** 7 nhóm **D.** 8 nhóm

Lời giải

Chọn C

Mẫu số liệu được chia thành 7 nhóm

Câu 3. (TH) Cho mẫu số liệu ghép nhóm về thống kê nhiệt độ tại một địa điểm trong 40 ngày, ta có bảng số liệu sau:

Nhiệt độ $\left(^{\circ}C\right)$	$[19;22)$	$[22;25)$	$[25;28)$	$[28;31)$
Số ngày	7	15	12	6

Có bao nhiêu ngày có nhiệt độ từ $28^{\circ}C$ đến dưới $31^{\circ}C$

A. 4 **B.** 5 **C.** 6 **D.** 7

Lời giải

Chọn C

Có 6 ngày có nhiệt độ từ $28^{\circ}C$ đến dưới $31^{\circ}C$

Câu 4. (TH) Tìm hiểu thời gian xem tivi trong tuần trước (đơn vị: giờ) của một số học sinh thu được kết quả sau:

Thời gian (giờ)	$[0;5)$	$[5;10)$	$[10;15)$	$[15;20)$	$[20;25)$
Số học sinh	8	16	4	2	2

Giá trị đại diện của nhóm $[20;25)$ là

A. 22,5 . **B.** 23. **C.** 20 . **D.** 5 .

Lời giải

Giá trị đại diện của nhóm $[20;25)$ là $\frac{20+25}{2}=22,5$

Câu 5. (TH) Quãng đường (km) từ nhà đến nơi làm việc của 40 công nhân một nhà máy được ghi lại như sau:

5	3	10	20	25	11	13	7	12	31
19	10	12	17	18	11	32	17	16	2
7	9	7	8	3	5	12	15	18	3
12	14	2	9	6	15	15	7	6	12

Ghép nhóm dãy số liệu trên thành các khoảng có độ rộng bằng nhau, khoảng đầu tiên là $[0;5)$. Tần số của nhóm $[10;15)$ là:

A. 12 **B.** 11 **C.** 10 **D.** 9

Lời giải

Chọn D

Ta chia thành các nhóm có độ dài là 5. Ta sẽ chọn đầu mút phải của nhóm cuối cùng là 35. Ta có bảng ghép nhóm sau:

Quảng đường (km)	$[0;5)$	$[5;10)$	$[10;15)$	$[15;20)$	$[20;25)$	$[25;30)$	$[30;35)$

Số công nhân	6	10	11	9	1	1	2

Vậy tần số của lớp $[10;15)$ là 11

PHẦN 2. Trả lời Đúng – Sai

Câu 1: Cho bảng số liệu thống kê thành tích chạy $50m$ của học sinh lớp 10A – trường THPT X như sau:

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

Lập bảng tần số ghép nhóm với nhóm đầu tiên là $[6,0;\ 6,5)$ và độ dài mỗi nhóm bằng 5 ta được:

a) (NB) Cỡ mẫu của số liệu ghép nhóm là 40.

b) (TH) Có 7 học sinh có thành tích chạy từ 6,5 đến dưới 7,0 giây.

c) (**TH**)Tần số lớn nhất thuộc nhóm $[7,5;\ 8,0)$.

d) (**VD**) Thành tích chạy $50m$ của các học sinh trung bình là $7,5$ giây.

Lời giải

Đúng; Sai; Đúng; Đúng

Ta có bảng sau:

Thành tích	$[6,0;\ 6,5)$	$[6,5;\ 7,0)$	$[7,0;\ 7,5)$	$[7,5;\ 8,0)$	$[8,0;\ 8,5)$	$[8,5;\ 9,0)$
Giá trị đại diện	6,25	6,75	7,25	7,75	8,25	8,75
Số học sinh	2	6	10	14	6	2

0) **Đúng**

Cỡ mẫu: $n = 2+6+10+14+6+2 = 40$

b) Sai

Theo bảng số liệu ghép nhóm trên có 6 học sinh có thành tích chạy từ 6,5 đến dưới 7,0 giây.

c) Đúng

$[7,5;\ 8,0)$ có tần số lớn nhất là

Theo bảng số liệu ghép nhóm trên thì nhóm
14.

d) Đúng

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

$$\overline{x}=\frac{2.6,25+6.6,75+10.7,25+14.7,75+6.8,25+2.8,75}{40}=7,525$$

Câu 2: Số lượng người đi xem một bộ phim mới theo độ tuổi trong một rạp chiếu phim được ghi lại theo bảng phân phối ghép nhóm sau:

Độ tuổi	[10;20)	[20;30)	[30;40)	[40;50)	[50;60)
Số người	6	12	16	7	3

a) (TH) Giá trị đại diện nhóm $[50;60)$ là 55.

b) (TH) Độ tuổi được dự báo là ít xem phim đó nhất là thuộc nhóm $[50;60)$

c) (NB) Nhóm chứa mốt là nửa khoảng $[40;50)$.

d) (VD) Độ tuổi được dự báo là thích xem phim đó nhiều nhất là 31 tuổi.

Lời giải

Đúng; Đúng; Sai; Sai.

a) Giá trị đại diện nhóm $[50;60)$ là $\frac{1}{2}(50+60)=55.$

b) Nhóm $[50;60)$ có tần số ít nhất suy ra độ tuổi được dự báo là ít xem phim đó nhất là thuộc nhóm $[50;60)$.

c) Ta có nhóm $[30;40)$ có tần số lớn nhất suy ra nhóm chứa mốt là $[30;40)$

Do nhóm chứa mốt là $[30;40)$ nên ta có:

$u_m=30, n_m=16, n_{m-1}=12, n_{m+1}=7, u_{m+1}-u_m=40-30=10$.

Ta có mốt là: $M_0=30+\frac{16-12}{(16-12)+(16-7)}\cdot 10=\frac{430}{13}\approx 33,08$.

d) Vậy độ tuổi được dự báo là thích xem phim đó nhiều nhất là 33 tuổi.

PHẦN 3: Trả lời ngắn.

Câu 1. (VD) Theo Tổ chức Y tế Thế giới (WHO), thiếu máu là tình trạng giảm lượng huyết sắc tố (Hb) dẫn tới sự thiếu cung cấp oxygen cho các mô trong cơ thể. Đối với nam giới trên 15 tuổi, chỉ số Hb (đơn vị tính là g/l) lớn hơn hoặc bằng 130 được xem là không bị thiếu máu, từ 110 đến dưới 130 là thiếu máu mức nhẹ, từ 80 đến dưới 110 là thiếu máu mức vừa, dưới 80 là mức nặng. Đo chỉ số Hb của một số học sinh nam lớp 12 cho kết quả như sau:

132,135,137,131,129,125,140,147,138,137,128,112,
127,129,125,98,139,138,139,141,140,105,136,133,
137,138,108,133,136,141,144,134,136,137,142.

Gọi x là số học sinh có chỉ số Hb từ 130 trở lên (không bị thiếu máu) và y là số học sinh có chỉ số Hb từ 80 đến dưới 110 (thiếu máu mức trung bình). Tính giá trị biểu thức $T=x+2y.$

Lời giải

Ta có mẫu số liệu ghép nhóm:

Chỉ số Hb (g/l)	Từ 130 trở lên	[110;130)	[80;110)
Số học sinh	25	7	3

Dựa vào bảng số liệu ghép nhóm ta có:

$x = 25$ học sinh có chỉ số Hb từ 130 trở lên (không bị thiếu máu),

$y = 3$ học sinh có chỉ số Hb từ 80 đến dưới 110 (thiếu máu mức trung bình).

Vậy $T = x + 2y = 25 + 2.3 = 31$.

ĐÁP SỐ: 31.

Câu 2. (VD) Một công ty may quần áo đồng phục học sinh cho biết cỡ áo theo chiều cao của học sinh được tính như sau:

Chiều cao (cm)	[150;160)	[160;167)	[167;170)	[170;175)	[175;180)
Cỡ áo	S	M	L	XL	XXL

Công ty muốn ước lượng tỉ lệ các cỡ áo khi may cho học sinh lớp 11 đã đo chiều cao của 36 học sinh nam khối 11 của một trường và thu được mẫu số liệu sau (đơn vị là centimét):

160 161 161 162 162 162 163 163 163 164 164 164 164

165 165 165 165 165 166 166 166 166 167 167 168 168

168 168 169 169 170 171 171 172 172 174

Công ty may 500 áo đồng phục cho học sinh lớp 11 thì nên may số lượng áo của cỡ M là bao nhiêu chiếc?

Lời giải

Đếm số giá trị thuộc mỗi nhóm, ta lập được bảng tần số ghép nhóm của mẫu số liệu với các nhóm đã cho ở bảng trên như sau:

Chiều cao (cm)	[150;160)	[160;167)	[167;170)	[170;175)	[175;180)
Số học sinh	0	22	8	6	0

Công ty may 500 áo đồng phục cho học sinh lớp 11 thì nên may số lượng áo theo mỗi cỡ như sau:

- Không nên may áo cỡ S và cỡ XXL;
- Số lượng áo cỡ M nên may là $\frac{22}{36}.500 \approx 306$ (chiếc);
- Số lượng áo cỡ L nên may là $\frac{8}{36}.500 \approx 111$ (chiếc);
- Số lượng áo cỡ XL nên may là $500 - 306 - 111 = 83$ (chiếc).

ĐÁP SỐ: 306 **chiếc.**

2. CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM

Câu 1: (NB) Giả sử mẫu số liệu được cho dưới dạng bảng tần số ghép nhóm:

Nhóm	Nhóm 1	Nhóm 2	…	Nhóm k
Giá trị đại diện	c_1	c_2	…	c_k
Tần số	n_1	n_2	…	n_k

Đặt $n = n_1 + n_2 + \ldots + n_k$.

Số trung bình của mẫu số liệu ghép nhóm, kí hiệu $\overline{x}$, được tính theo công thức nào?

A. $\overline{x}=\frac{n_1c_1+n_2c_2+\ldots+n_kc_k}{n}$. **B.** $\overline{x}=\frac{n_1c_1+n_2c_2+\ldots+n_kc_k}{2n}$.

C. $\overline{x}=\frac{n_1^2c_1+n_2^2c_2+\ldots+n_k^2c_k}{n}$. **D.** $\overline{x}=\frac{n_1c_1+n_2c_2+\ldots+n_kc_k}{\sqrt{n}}$.

Lời giải

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

Nhóm	Nhóm 1	Nhóm 2	⋯	Nhóm k
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 $\overline{x}$, được tính như sau:

$\overline{x}=\frac{n_1c_1+n_2c_2+\ldots+n_kc_k}{n}$ trong đó $n=n_1+n_2+\ldots+n_k$.

Câu 2: (NB) Kết quả khảo sát cân nặng của 25 quả cam ở lô hàng A được cho ở bảng sau:

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

Nhóm chứa mốt là nhóm nào

A. $[150;155)$. **B.** $[155;160)$. **C.** $[165;170)$. **D.** $[170;175)$.

Lời giải

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

Nhóm chứa mốt của mẫu số liệu ghép nhóm là nhóm $[165;170)$.

Câu 3: (NB) Khảo sát thời gian tập thể dục 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

Giá trị đại diện của nhóm $[20;40)$ là

A. 10. **B.** 20. **C.** 30. **D.** 40.

Lời giải

Giá trị đại diện của nhóm $[20;40)$ là $c=\frac{20+40}{2}=30$.

Câu 4: (NB) Một cuộc khảo sát đã tiến hành xác định tuổi (theo năm) của 120 chiếc ô tô. Kết quả điều tra được cho trong bảng sau.

Số tuổi (theo năm)	[0;4)	[4;8)	[8;12)	[12;16)	[20;24)
Số ô tô	23	25	37	26	19

Mẫu số liệu trên có bao nhiêu nhóm

A. 10. **B.** 11. **C.** 7. **D.** 5.

Câu 5: (TH) 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 trung vị của mẫu số liệu là

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

Câu 6: (TH) Một cuộc khảo sát đã tiến hành xác định tuổi (theo năm) của 120 chiếc ô tô. Kết quả điều tra được cho trong bảng sau.

Số tuổi (theo năm)	[0;4)	[4;8)	[8;12)	[12;16)	[20;24)
Số ô tô	23	25	37	26	19

Có bao nhiêu ô tô có độ tuổi dưới 12

A.75. **B.** 37. **C.** 45. **D.** 26.

ĐÚNG - SAI

Câu 1. Người ta tiến hành phỏng vấn 30 người về một bộ phim mới chiếu trên truyền hình. Người điều tra yêu cầu cho điểm bộ phim (thang điểm là 100). Kết quả được trình bày trong bảng phân bố tần số ghép lớp sau đây:

Số điểm	$[50;60)$	$[60;70)$	$[70;80)$	$[80;90)$	$[90;100)$
Số người	2	6	10	8	4

(VD) a) Ước lượng số trung bình của mẫu ghép là $77.$

(NB) b) Giá trị đại diện của nhóm $[90;100)$ là 95.

(NB) c) Nhóm chứa mốt của mẫu số liệu trên là nhóm $[80;90)$.

(VD) d) Mốt của mẫu số liệu là $74,67.$

Lời giải

a) Đúng	**b) Đúng**	**c) Sai**	**d) Sai**

Số điểm	$[50;60)$	$[60;70)$	$[70;80)$	$[80;90)$	$[90;100)$
Giá trị đại diện	55	65	75	85	95
Số người	2	6	10	8	4

Ước lượng số trung bình của mẫu ghép là:

$$\overline{x}=\frac{55.2+65.6+75.10+85.8+95.4}{30}=77.$$

Nhóm chứa mốt của mẫu số liệu trên là nhóm $[70;80)$.

Do đó, $u_m=70;n_{m-1}=6;n_m=10;n_{m+1}=8;u_{m+1}-u_m=80-70=10$.

$$M_o=70+\frac{10-6}{(10-6)+(10-8)}\cdot 10=76,67.$$

Câu 2: Dựa vào bảng tần số mẫu số liệu ghép nhóm sau, hãy tìm tứ phân vị của nó.

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

Các mệnh đề sau đúng hay sai ?

(NB) a) Cỡ mẫu của mẫu số liệu là $n=40$.

(NB) b) Giá trị đại diện của nhóm $[40;50)$ là 45.

(VD) c) Số trung bình của mẫu số liệu ghép nhóm là: $\overline{x}=65$.

(VD) d) Trung vị của mẫu số liệu ghép nhóm là: $M_e=45$.

Lời giải

a) Đúng	b) Đúng	c) Sai	d) Sai

a) Cỡ mẫu của mẫu số liệu là $n=2+10+16+8+2+2=40.$

b) Giá trị đại diện của nhóm $[40;50)$ là $c=\frac{40+50}{2}=45$.

c) Số trung bình của mẫu số liệu ghép nhóm là:

$\overline{x}=\frac{35.2+45.10+55.16+65.8+75.2+85.2}{40}=56$.

d) Nhóm chứa trung vị là $[50;60)$.

Trung vị của mẫu số liệu ghép nhóm là:

$M_e=50+\frac{\frac{40}{2}-(2+10)}{16}.(60-50)=55$.

TRẢ LỜI NGẮN

Câu 1: (VD) Bảng sau cho ta cân nặng của học sinh một lớp 11:

Cân nặng (kg)	$[40;45)$	$[45;50)$	$[50;55)$	$[55;60)$	$[60;65)$
Số học sinh	10	7	16	4	3

Tìm cân nặng trung bình của học sinh lớp 11 đó, làm tròn đến chữ số thập phân thứ nhất.

Trả lời: $50,4(kg)$

Lời giải

Bảng thống kê cân nặng của học sinh theo giá trị đại diện:

Cân nặng (kg)	42,5	47,5	52,5	57,5	62,5
Số học sinh	10	7	16	4	3

Cân nặng trung bình của học sinh lớp 11 đó là:

$\overline{x}=\frac{42,5.10+47,5.7+52,5.16+57,5.4+62,5.3}{40}\approx 50,4(kg).$

Câu 2: (VD) Bảng số liệu ghép nhóm sau cho biết chiều cao (cm) của 40 học sinh lớp 11A.

Khoảng chiều cao (cm)	$[145;150)$	$[150;155)$	$[155;160)$	$[160;165)$	$[165;170)$
Số học sinh	7	13	6	8	6

Tìm mốt của mẫu số liệu ghép nhóm này, làm tròn đến hàng đơn vị.

Trả lời: 152.

Lời giải

Nhóm chứa mốt của mẫu số liệu trên là $[150;155)$

Mốt của mẫu số liệu ghép nhóm là: $M_o = 150 + \frac{13-7}{(13-7)+(13-6)} \cdot 5 \approx 152.$.

3. KHOẢNG BIẾN THIÊN VÀ KHOẢNG TỨ PHÂN VỊ

PHẦN I. TRẮC NGHIỆM 4 PHƯƠNG ÁN

Câu 1.[NB] Một ý nghĩa của khoảng tứ phân vị là

A. Khoảng tứ phân vị của mẫu số liệu ghép nhóm giúp xác định các giá trị không bất thường của mẫu số liệu đó.

B. Khoảng tứ phân vị thường không được sử dụng thay cho khoảng biến thiên.

C. Khoảng tứ phân vị của mẫu số liệu ghép nhóm xấp xỉ khoảng tứ phân vị của mẫu số liệu gốc và là một đại lượng cho biết mức độ phân tán của nửa giữa mẫu số liệu.

D. Khoảng tứ phân vị của mẫu số liệu ghép nhóm xấp xỉ khoảng tứ phân vị của mẫu số liệu gốc và là một đại lượng cho biết mức độ không phân tán của nửa giữa mẫu số liệu.

Câu 2.[NB] Cho bảng khảo sát về khối lượng của 30 củ khoai tây thu hoạch ở một nông trường như sau:

Khối lượng (gam)	[70;80)	[80;90)	[90;100)	[100;110)	[110;120)	Cộng
Số lượng	3	6	12	6	3	30

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là

A. 50. **B.** 10. **C.** 30. **D.** 40.

Câu 3. [NB] Công thức tính khoảng tứ phân vị của mẫu số liệu ghép nhóm là

A. $\Delta_Q = Q_3 - Q_1$. **B.** $\Delta_Q = Q_1 - Q_3$. **C.** $\Delta_Q = Q_3.Q_1$. **D.** $\Delta_Q = Q_3 + Q_1$.

Câu 4. [TH] Bạn Linh thống kê chiều cao (đơn vị: cm) của các bạn học sinh nữ lớp $12A$ và lớp $12B$ ở bảng sau:

Chiều cao (cm)	$[150;155)$	$[155;160)$	$[160;165)$	$[165;170)$	$[170;175)$	$[175;180)$
Số học sinh nữ lớp 12 A	2	7	12	3	0	1
Số học sinh nữ lớp 12 B	0	9	8	2	1	5

Gọi R_1; R_2 lần lượt là khoảng biến thiên của mẫu số liệu ghép nhóm về chiều cao của các bạn học sinh nữ lớp $12A$ và $12B$. Tìm R_1; R_2.

A. $R_1 = 30\,(cm)\ ;\ R_2 = 25\,(cm)$. **B.** $R_1 = 30\,(cm)\ ;\ R_2 = 30\,(cm)$.

$R_1 = 25\,(cm)\ ;\ R_2 = 25\,(cm)$ $R_1 = 12\,(cm)\ ;\ R_2 = 9\,(cm)$

C. . **D.** .

Câu 5. [TH] Mức thưởng Tết cho các nhân viên của 2 tổ tại một công ty được thống kê trong bảng sau:

Mức thưởng Tết (triệu đồng)	$[5;10)$	$[10;15)$	$[15;20)$	$[20;25)$	$[25;30)$
Số nhân viên tổ A	40	25	20	10	5
Số nhân viên tổ B	50	30	20	10	0

Gọi $R_1; R_2$ tương ứng là khoảng biến thiên của mẫu số liệu ghép nhóm về mức thưởng Tết của các nhân viên Tổ A và Tổ B. Chọn phương án đúng?

A. $R_1 = 20$. **B.** $R_2 = 25$. **C.** $R_1 > R_2$. **D.** $R_1 < R_2$.

Câu 6. [TH] Khảo sát thời gian nghe nhạc trong ngày của một số học sinh khối 12 thu được mẫu số liệu ghép nhóm sau:

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

Nhóm chứa tứ phân vị thứ ba là

A. $[20;40)$. **B.** $[40;60)$. **C.** $[60;80)$. **D.** $[80;100)$.

Câu 7. [TH] Kết quả điều tra tổng thu nhập trong năm 2024 của một số hộ gia đình ở thành phố Nha

Trang được ghi lại ở bảng sau:

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

Tứ phân vị Q_1 bằng

A. $\frac{675}{62}$. **B.** $\frac{9775}{31}$. **C.** $\frac{16715}{62}$. **D.** $\frac{16175}{62}$.

Câu 8. [VD] Điểm kiểm tra 15 phút của 36 học sinh lớp 11A được cho bởi bảng tần số ghép

nhóm sau:

Nhóm điểm	**Tần số**	Tần số tích lũy
$[1;3)$	3	3
$[3;5)$	2	5

$[5;7)$	10	15
$[7;9)$	14	29
$[9;11)$	7	36
	$n=36$	

Khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu ghép nhóm trên lần lượt là

A. $10;\ 9,2$. **B.** $10;\ 2,9$. **C.** $10;\ 25,3$. **D.** $6;\ 20,5$.

Câu 9. [VD] Điều tra 42 học sinh của một lớp 11 về số giờ tự học ở nhà, người ta có bảng sau đây:

Lớp (Số giờ tự học)	Tần số	Tần số tích lũy
$[1;2)$	8	8
$[2;3)$	10	18
$[3;4)$	12	30
$[4;5)$	9	39
$[5;6)$	3	42
	$n=42$	

Khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu ghép nhóm trên lần lượt là

A. $5;\ 1,95$. **B.** $2;\ 3,1$. **C.** $2;\ 3,2$. **D.** $3;\ 1,2$.

Câu 10. [VD] Bốn bạn Ánh, Ba, Châu, Dũng cùng là thành viên của một câu lạc bộ rubik. Trong một lần luyện tập rubik với nhau, mỗi bạn đã cùng giải rubik 30 lần liên tiếp và thống kê kết quả lại ở bảng sau:

Thời gian giải rubik (giây)	$[6;8)$	$[8;10)$	$[10;12)$	$[12;14)$	$[14;16)$
Ánh	1	8	5	7	9
Ba	4	8	5	6	7
Châu	5	1	6	5	13
Dũng	2	6	6	8	8

Nếu so sánh theo khoảng tứ phân vị của mẫu số liệu ghép nhóm thì bạn nào có tốc độ giải rubik đồng đều nhất?

A. Ánh. **B.** Ba. **C.** Châu. **D.** Dũng.

PHẦN II: CÂU HỎI ĐÚNG SAI

Câu 1. [NB-NB-NB-NB] Khoảng tứ phân vị của mẫu số liệu ghép nhóm là:

a. Hiệu giữa tứ phân vị thứ ba và tứ phân vị thứ nhất của mẫu số liệu ghép nhóm đó.

b. Tổng giữa tứ phân vị thứ ba và tứ phân vị thứ nhất của mẫu số liệu ghép nhóm đó.

c. Hiệu giữa hai tứ phân vị bất kì của mẫu số liệu ghép nhóm đó.

d. Tổng giữa hai tứ phân vị bất kì của mẫu số liệu ghép nhóm đó.

Câu 2. [NB-TH-TH-VD-VD]Thời gian (phút) truy cập Internet mỗi buổi tối của một số học sinh được cho ở 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	4	12	14	23	3

Trong các khẳng định sau, khẳng định nào đúng, khẳng định nào sai?

a/ Khoảng biến thiên của mẫu số liệu là 15 .

b/ Nhóm chứa tứ phân vị thứ ba là $[15,5;18,5)$.

c/ Tứ phân vị thứ nhất là $Q_1=15$.

d/ Khoảng tứ phân vị của mẫu số liệu ghép nhóm bé hơn 6 .

Câu 3. [NB-VD-VD-VD] Bảng tần số ghép nhóm dưới đây thống kê số giờ ngủ buổi tối của các học sinh lớp 12A1 và 12A2:

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

a) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian ngủ của các bạn nam là 5

b) Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian ngủ của các bạn nam là $2,09$

c) Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian ngủ của các bạn nữ trong khoảng $(2;3)$

d) Học sinh nam có thời gian ngủ đồng đều hơn.

PHẦN III: TRẢ LỜI NGẮN

Câu 1. [NB] Cho bảng số liệu ghép nhóm về chiều cao (đơn vị centimét) của 36 học sinh trong lớp 12A1 như sau:

Nhóm	$[150;155)$	$[155;160)$	$[160;165)$	$[165;170)$	$[170;175)$	$[175;180)$
Tần số	6	11	8	7	3	1

Xác định khoảng biến thiên của mẫu số liệu ghép nhóm?

Trả lời:.....

Câu 2. [TH] Cho bảng số liệu ghép nhóm về chiều cao (đơn vị centimét) của 36 học sinh trong lớp 12A1 như sau:

Nhóm	$[150;155)$	$[155;160)$	$[160;165)$	$[165;170)$	$[170;175)$	$[175;180)$
Tần số	6	11	8	7	3	1

Nhóm chứa tứ phân vị thứ nhất có tần số bằng bao nhiêu?

Trả lời:.....

Câu 3. [VD] Khoảng tứ phân vị của mẫu số liệu ghép nhóm được cho ở bảng sau là bao nhiêu?(làm tròn kết quả đến hai chữ số thập phân).

Nhóm	$[25;30)$	$[30;35)$	$[35;40)$	$[40;45)$
Tần số	2	17	10	25

Trả lời:.....

Câu 4. [VD] Cho bảng số liệu về khối lượng của 30 củ khoai tây thu hoạch từ một thửa ruộng như dưới đây.

Khối lượng (gam)	[70;80)	[80;90)	[90;100)	[100;110)	[110;120)
Tần số	3	6	12	6	3

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là bao nhiêu?

Trả lời:.....

Câu 5. [VD] Điều tra về chiều cao của học sinh khối lớp 10, ta có kết quả sau:

Nhóm	Chiều cao (cm)	Số học sinh
1	[150;152)	5
2	[152;154)	18
3	[154;156)	40
4	[156;158)	26
5	[158;160)	8
6	[160;162)	3
		$N=100$

Tính tổng tứ phân vị thứ nhất Q_1 và tứ phân vị thứ ba Q_3.

Trả lời:.....

BÀI TẬP PHÁT TRIỂN(BÀI TẬP VỀ NHÀ)

PHẦN I. TRẮC NGHIỆM 4 PHƯƠNG ÁN

Câu 1.[NB] Điểm kiểm tra của nhóm học sinh lớp 10 được cho như sau:

Lớp điểm	[3;4]	[5;6]	[7;8]	[9;10]
Số học sinh	3	3	2	2

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là

A. 7 **B.** 3 **C.** 5 **D.** 6

Câu 2.[NB]Cho mẫu số liệu ghép nhóm với bộ ba tứ phân vị lần lượt là $Q_1 = 11,5$ $Q_2 = 14,5$; $Q_3 = 21,3.$; Khi đó khoảng tứ phân vị của mẫu số liệu trên là

A. $\Delta Q = 3,0$. **B.** $\Delta Q = 6,8$. **C.** $\Delta Q = 9,8$. **D.** $\Delta Q = 32,8$.

Câu 3. [NB] Cho mẫu số liệu ghép nhóm về khối lượng (đơn vị: gram) của 30 củ khoai tây như sau:

Khối lượng	$[70;\ 80)$	$[80;\ 90)$	$[90;\ 100)$	$[100;\ 110)$	$[110;120)$
Tần số	3	6	12	6	3

Giá trị đại diện của nhóm $[90;100)$

A. 85. **B.** 95. **C.** 90. **D.** 100.

Câu 4. [TH] Ta có bảng sau về thời gian tập thể dục buổi sáng của bác Bình và bác An:

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

Hỏi hiệu khoảng biến thiên của mẫu số liệu của bác Bình và bác An là bao nhiêu?

A. 11. **B.** 9. **C.** 15. **D.** 10.

Câu 5. [TH]Đo cân nặng của 40 học sinh lớp 12A9 ta được bảng số liệu như sau:

Khối lượng(kg)	[40;45)	[45;50)	[50;55)	[55;60)	[60;65)	[65;70)	[70;75)	[75;80]
Số học sinh	4	13	7	5	6	2	1	2

Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm thuộc khoảng nào sau đây?

A. $[40;45]$. **B.** $[45;50]$. **C.** $[50;55]$. **D.** $[55;60]$.

Câu 6. [TH] Thống kê thời gian sử dụng mạng xã hội trong ngày của các bạn học sinh tổ 1 và tổ 2 lớp 12A thu được bảng sau:

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

Tìm khoảng biến thiên R_1, R_2 cho thời gian sử dụng mạng xã hội của tổ 1 và tổ 2.

A. $R_1 = 60;\ R_2 = 90$.

B. $R_1 = 90;\ R_2 = 60$.

C. $R_1 = 30;\ R_2 = 30$.

D. $R_1 = 10;\ R_2 = 30$.

Câu 7. [VD] Điều tra về khối lượng 27 củ khoai tây (đơn vị: gam) thu hoạch tại nông trường, ta có kết quả sau:

Nhóm	Tần số	Tần số tích lũy
$[74; 80)$	4	4
$[80; 86)$	6	10
$[86; 92)$	3	13
$[92; 98)$	4	17
$[98; 104)$	3	20
$[104; 110)$	7	27
	$n = 27$	

Khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu ghép nhóm trên lần lượt là

A. $36;\ 21,45$.

B. $7;\ 23$.

C. $11;\ 25,3$.

D. $33;\ 20,5$.

Câu 8. [VD] Thống kê chiều cao của một số cây bạch đàn giống 1 tháng tuổi của 4 nông trường được cho bởi bảng sau:

Chiều cao (cm)	$[5;7)$	$[7;9)$	$[9;11)$	$[11;13)$	$[13;15)$
Nông trường A	5	8	16	8	3
Nông trường B	5	10	8	9	6
Nông trường C	13	9	9	3	9
Nông trường D	3	12	8	12	4

Nếu xét theo khoảng tứ phân vị thì cây bạch đàn giống 1 tháng tuổi ở nông trường nào có chiều cao đồng đều nhất?

A. Nông trường A.

B. Nông trường B.

C. Nông trường C.

D. Nông trường D.

Câu 9. [VD] Thống kê điểm kiểm tra môn Toán giữa kì I của bốn lớp 12 của một trường THPT cho bởi bảng sau:

Điểm	$[5;6)$	$[6;7)$	$[7;8)$	$[8;9)$	$[9;10]$
Lớp 12B 1	7	3	15	12	4
Lớp 12B 2	5	9	12	11	3
Lớp 12B 3	10	10	9	6	1
Lớp 12B 4	14	3	15	9	1

Nhà trường muốn đánh giá mức độ "học đều" môn Toán của các lớp. Nếu xét theo khoảng tứ phân vị thì điểm kiểm tra môn Toán giữa kì I của lớp nào đồng đều nhất?

A. Lớp 12B1. **B.** Lớp 12B2. **C.** Lớp 12B3. **D.** Lớp 12B4.

PHẦN II: CÂU HỎI ĐÚNG SAI

Câu 1. [NB-NB-NB-NB]Cho mẫu số liệu ghép nhóm có cùng đơn vị.

a) Khoảng biến thiên của mẫu số liệu ghép nhóm đo mức độ phân tán của mẫu số liệu đó.

b) Khoảng biến thiên càng lớn thì mẫu số liệu càng ít phân tán.

c) Khoảng biến thiên của mẫu số liệu ghép nhóm không bị ảnh hưởng bởi các giá trị bất thường.

d) Khoảng biến thiên của mẫu số liệu ghép nhóm xấp xỉ cho khoảng biến thiên của mẫu số liệu gốc.

Câu 2. [NB-NB-TH-TH] Cho bảng mẫu số liệu ghép nhóm về điểm môn Toán của hai lớp $12A$ và $12B$ được cho như sau:

Khoảng điểm	$[3;4)$	$[4;5)$	$[5;6)$	$[6;7)$	$[7;8)$	$[8;9)$	$[9;10)$
Số học sinh 12A	0	1	3	6	17	7	2
Số học sinh 12B	1	3	13	11	5	3	0

Trong các khẳng định sau , khẳng định nào đúng, khẳng định nào sai?

a) Khoảng biến thiên cho điểm môn Toán của lớp $12A$ là 7.

b) Khoảng biến thiên cho điểm môn Toán của lớp $12B$ là 6.

c) Nhóm chứa tứ phân vị thứ nhất của lớp $12A$ là nhóm $[6;7)$.

d) Nhóm chứa tứ phân vị thứ ba của lớp $12B$ là nhóm $[7;8)$.

Câu 3. [NB-VD-VD-VD] Bảng thống kê thời gian (đơn vị: phút) tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình và bạn An

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

a) Khoảng biến thiên của mẫu số liệu về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn An là 20.

b) Khoảng tứ phân vị của mẫu số liệu về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình là 28.

c) Khoảng tứ phân vị của mẫu số liệu về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn An là 22.

d) Dựa vào khoảng tứ phân vị của hai mẫu số liệu trên thì thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình phân tán hơn bạn An.

PHẦN III: TRẢ LỜI NGẮN

Câu 1. [NB] Khoảng biến thiên của mẫu số liệu ghép nhóm được cho ở bảng sau là bao nhiêu?

Nhóm	$[15;22)$	$[22;29)$	$[29;36)$	$[36;43)$	$[43;50)$
Tần số	1	6	21	21	11

Trả lời:.....

Câu 2. [TH] Cho bảng thống kê thời gian sử dụng điện thoại vào buổi sáng mỗi ngày trong tháng 4/2024 của Tuấn và An ở bảng như sau

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

Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian sử dụng điện thoại vào buổi sáng mỗi ngày của Tuấn và An lần lượt là R_1, R_2. Tính tổng $R_1 + 2R_2$.

Trả lời:.....

Câu 3. [VD] Bảng thống kê cân nặng của 50 quả xoài được lựa chọn ngẫu nhiên sau khi thu hoạch ở nông trường như sau

Cân nặng	$[250;290)$	$[290;330)$	$[330;370)$	$[370;410)$	$[410;450)$

(g)					
Số quả xoài	2	12	19	12	5

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

Trả lời:.....

Câu 4. [VD] An tìm hiểu hàm lượng chất béo (đơn vị: g) có trong 100 g mỗi loại thực phẩm. Sau khi thu thập dữ liệu về 60 loại thực phẩm, An lập được bảng thống kê

Hàm lượng chất béo (g)	$[2;6)$	$[6;10)$	$[10;14)$	$[14;18)$	$[18;22)$	$[22;26)$
Tần số	2	6	10	13	16	13

Tính khoảng tứ phân vị của mẫu số liệu?

Trả lời:.....

ĐÁP ÁN CHI TIẾT

PHẦN I. TRẮC NGHIỆM 4 PHƯƠNG ÁN

Câu 1.[NB] Một ý nghĩa của khoảng tứ phân vị là

A. Khoảng tứ phân vị của mẫu số liệu ghép nhóm giúp xác định các giá trị không bất thường của mẫu số liệu đó.

B. Khoảng tứ phân vị thường không được sử dụng thay cho khoảng biến thiên.

C. Khoảng tứ phân vị của mẫu số liệu ghép nhóm xấp xỉ khoảng tứ phân vị của mẫu số liệu gốc và là một đại lượng cho biết mức độ phân tán của nửa giữa mẫu số liệu.

D. Khoảng tứ phân vị của mẫu số liệu ghép nhóm xấp xỉ khoảng tứ phân vị của mẫu số liệu gốc và là một đại lượng cho biết mức độ không phân tán của nửa giữa mẫu số liệu.

Lời giải

Ý nghĩa của khoảng tứ phân vị:

- Khoảng tứ phân vị của mẫu số liệu ghép nhóm xấp xỉ khoảng tứ phân vị của mẫu số liệu gốc và là một đại lượng cho biết mức độ phân tán của nửa giữa mẫu số liệu.

- Khoảng tứ phân vị của mẫu số liệu ghép nhóm giúp xác định các giá trị bất thường của mẫu đó. - Khoảng tứ phân vị thường được sử dụng thay cho khoảng biến thiên vì nó loại trừ hầu hết giá trị bất thường của mẫu số liệu và nó không bị ảnh hưởng bởi các giá trị bất thường đó.

Câu 2.[NB] Cho bảng khảo sát về khối lượng của 30 củ khoai tây thu hoạch ở một nông trường như sau:

Khối lượng (gam)	[70;80)	[80;90)	[90;100)	[100;110)	[110;120)	Cộng
Số lượng	3	6	12	6	3	30

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là

A. 50. **B.** 10. **C.** 30. **D.** 40.

Lời giải

Ta có khoảng biến thiên của mẫu số liệu trên là

$120-70=50.$

Câu 3. [NB] Công thức tính khoảng tứ phân vị của mẫu số liệu ghép nhóm là

A. $\Delta_Q=Q_3-Q_1$. **B.** $\Delta_Q=Q_1-Q_3$. **C.** $\Delta_Q=Q_3.Q_1$. **D.** $\Delta_Q=Q_3+Q_1$.

Lời giải

Công thức tính khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q=Q_3-Q_1$

Câu 4. [TH] Bạn Linh thống kê chiều cao (đơn vị: cm) của các bạn học sinh nữ lớp $12A$ và lớp $12B$ ở bảng sau:

Chiều cao (cm)	$[150;155)$	$[155;160)$	$[160;165)$	$[165;170)$	$[170;175)$	$[175;180)$
Số học sinh nữ lớp 12 A	2	7	12	3	0	1
Số học sinh nữ lớp 12 B	0	9	8	2	1	5

Gọi R_1; R_2 lần lượt là khoảng biến thiên của mẫu số liệu ghép nhóm về chiều cao của các bạn học sinh nữ lớp $12A$ và $12B$. Tìm R_1; R_2.

A. $R_1=30\,(cm)\,;\,R_2=25\,(cm)$. **B.** $R_1=30\,(cm)\,;\,R_2=30\,(cm)$.

C. $R_1=25\,(cm)\,;\,R_2=25\,(cm)$. **D.** $R_1=12\,(cm)\,;\,R_2=9\,(cm)$.

Lời giải

Khoảng biến thiên của mẫu số liệu ghép nhóm về chiều cao của các bạn học sinh nữ lớp $12A$ là: $R_1=180-150=30$ (cm).

Trong mẫu số liệu ghép nhóm về chiều cao của các bạn học sinh nữ lớp $12B$, khoảng đầu tiên chứa dữ liệu là [155; 160) và khoảng cuối cùng chứa dữ liệu là [175; 180).

Khoảng biến thiên của mẫu số liệu ghép nhóm về chiều cao của các bạn học sinh nữ lớp $12B$ là: $R_2=180-155=25$ (cm).

Câu 5. [TH] Mức thưởng Tết cho các nhân viên của 2 tổ tại một công ty được thống kê trong bảng sau:

Mức thưởng Tết (triệu đồng)	$[5;10)$	$[10;15)$	$[15;20)$	$[20;25)$	$[25;30)$
Số nhân viên tổ A	40	25	20	10	5

Số nhân viên tổ B	50	30	20	10	0

Gọi $R_1;R_2$ tương ứng là khoảng biến thiên của mẫu số liệu ghép nhóm về mức thưởng Tết của các nhân viên Tổ A và Tổ B. Chọn phương án đúng?

A. $R_1=20$. **B.** $R_2=25$. **C.** $R_1>R_2$. **D.** $R_1<R_2$.

Lời giải

Ta có: $R_1=a_{k+1}-a_1=30-5=25$.

$R_2=a_{k+1}-a_1=25-5=20$.

Vậy $R_1>R_2$.

Câu 6. [TH] Khảo sát thời gian nghe nhạc trong ngày của một số học sinh khối 12 thu được mẫu số liệu ghép nhóm sau:

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

Nhóm chứa tứ phân vị thứ ba là

A. $[20;40)$. **B.** $[40;60)$. **C.** $[60;80)$. **D.** $[80;100)$.

Lời giải

Gọi $x_1;x_2;\ldots;x_{42}$ là mẫu số liệu gốc về thời gian nghe nhạc trong ngày của 42 học sinh khối 12 được xếp theo thứ tự tăng dần.

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

Câu 7. [TH] Kết quả điều tra tổng thu nhập trong năm 2024 của một số hộ gia đình ở thành phố Nha

Trang được ghi lại ở bảng sau:

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

Tứ phân vị Q_1 bằng

A. $\frac{675}{62}$. **B.** $\frac{9775}{31}$. **C.** $\frac{16715}{62}$. **D.** $\frac{16175}{62}$.

Lời giải

Số hộ gia đình được khảo sát (cỡ mẫu) là $n = 24+62+34+21+9 = 150$.

Ta có, $\frac{n}{4} = \frac{150}{4} = \frac{75}{2}$ suy ra $24 < \frac{75}{2} < 24+62$ nên nhóm thứ hai $[250;300)$ là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng $\frac{75}{2}$. Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là:

$$Q_1 = 250 + \frac{\frac{150}{4} - 24}{62}(300-250) = \frac{16175}{62}.$$

Câu 8. **[VD]**Điểm kiểm tra 15 phút của 36 học sinh lớp 11A được cho bởi bảng tần số ghép

nhóm sau:

Nhóm điểm	**Tần số**	Tần số tích lũy
$[1;3)$	3	3
$[3;5)$	2	5
$[5;7)$	10	15
$[7;9)$	14	29
$[9;11)$	7	36
	$n = 36$	

Khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu ghép nhóm trên lần lượt là

A. $10;\ 9,2$. **B.** $10;\ 2,9$. **C.** $10;\ 25,3$. **D.** $6;\ 20,5$.

Lời giải

Trong mẫu số liệu ghép nhóm trên, ta có: đầu mút trái của nhóm 1 là $a_1 = 1$, đầu mút phải của nhóm 5 là $a_6 = 11$. Vậy khoảng biến thiên của mẫu số liệu ghép nhóm đó là : $R = a_6 - a_1 = 11-1 = 10$ (điểm)

Số phần tử của mẫu là $n = 36$

Ta có: $\frac{n}{4}=\frac{36}{4}=9$ mà $5<9<15$. Suy ra nhóm 3 là nhóm đầu tiên có tần số tích luỹ lớn hơn hoặc bằng 9. Xét nhóm 3 là nhóm $[5;7)$ có $s=5$; $h=2$; $n_3=10$ và nhóm 2 là nhóm $[3;5)$ có $cf_2=5$.

Áp dụng công thức, ta có tứ phân vị thứ nhất là:

$$Q_1=5+\left(\frac{9-5}{10}\right).2=5,8$$

(điểm)

Ta có: $\frac{3n}{4}=\frac{3.36}{4}=27$ mà $15<27<29$. Suy ra nhóm 4 là nhóm đầu tiên có tần số tích luỹ lớn hơn hoặc bằng 27. Xét nhóm 4 là nhóm $[7;9)$ có $t=7$; $l=2$; $n_4=14$ và nhóm 3 là nhóm $[5;7)$ có $cf_3=15$.

Áp dụng công thức, ta có tứ phân vị thứ ba là:

$$Q_3=7+\left(\frac{27-15}{14}\right).2\approx 8,7$$

(điểm)

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

$$\Delta_Q=Q_3-Q_1\approx 8,7-5,8=2,9$$

(điểm)

Câu 9. **[VD]**Điều tra 42 học sinh của một lớp 11 về số giờ tự học ở nhà, người ta có bảng sau đây:

Lớp (Số giờ tự học)	Tần số	Tần số tích lũy
$[1;2)$	8	8
$[2;3)$	10	18
$[3;4)$	12	30
$[4;5)$	9	39
$[5;6)$	3	42
	$n=42$	

Khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu ghép nhóm trên lần lượt là

A. $5;1,95$. **B.** $2;3,1$. **C.** $2;3,2$. **D.** $3;1,2$.

Lời giải

Trong mẫu số liệu ghép nhóm trên, ta có: đầu mút trái của nhóm 1 là $a_1 = 1$, đầu mút phải của nhóm 5 là $a_6 = 6$. Vậy khoảng biến thiên của mẫu số liệu ghép nhóm đó là : $R = a_6 - a_1 = 6 - 1 = 5$ (giờ)

Số phần tử của mẫu là $n = 42$

Ta có: $\frac{n}{4} = \frac{42}{4} = 10,5$ mà $8 < 10,5 < 18$. Suy ra nhóm 2 là nhóm đầu tiên có tần số tích luỹ lớn hơn hoặc bằng $10,5$. Xét nhóm 2 là nhóm $[2;3)$ có $s = 2$; $h = 1$; $n_2 = 10$ và nhóm 1 là nhóm $[1;2)$ có $cf_1 = 8$.

Áp dụng công thức, ta có tứ phân vị thứ nhất là:

$$Q_1 = 2 + \left(\frac{10,5 - 8}{10}\right).1 = 2,25$$ (giờ)

Ta có: $\frac{3n}{4} = \frac{3.42}{4} = 31,5$ mà $30 < 31,5 < 39$. Suy ra nhóm 4 là nhóm đầu tiên có tần số tích luỹ lớn hơn hoặc bằng $31,5$. Xét nhóm 4 là nhóm $[4;5)$ có $t = 4$; $l = 1$; $n_4 = 9$ và nhóm 3 là nhóm $[3;4)$ có $cf_3 = 30$.

Áp dụng công thức, ta có tứ phân vị thứ ba là:

$$Q_3 = 4 + \left(\frac{31,5 - 30}{9}\right).1 \approx 4,2$$ (giờ)

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

$$\Delta_Q = Q_3 - Q_1 \approx 4,2 - 2,25 = 1,95$$ (giờ)

Câu 10. [VD] Bốn bạn Ánh, Ba, Châu, Dũng cùng là thành viên của một câu lạc bộ rubik. Trong một lần luyện tập rubik với nhau, mỗi bạn đã cùng giải rubik 30 lần liên tiếp và thống kê kết quả lại ở bảng sau:

Thời gian giải rubik (giây)	$[6;8)$	$[8;10)$	$[10;12)$	$[12;14)$	$[14;16)$
Ánh	1	8	5	7	9
Ba	4	8	5	6	7
Châu	5	1	6	5	13
Dũng	2	6	6	8	8

Nếu so sánh theo khoảng tứ phân vị của mẫu số liệu ghép nhóm thì bạn nào có tốc độ giải rubik đồng đều nhất?

A. Ánh. **B.** Ba. **C.** Châu. **D.** Dũng.

Lời giải

Bạn Ánh:

$Q_1 = 8 + \dfrac{\frac{30}{4} - 1}{8} \cdot 2 = \dfrac{77}{8}$, $Q_3 = 14 + \dfrac{\frac{30 \cdot 3}{4} - (1+8+5+7)}{9} \cdot 2 = \dfrac{43}{3}$,

$\Delta_Q = Q_3 - Q_1 = \dfrac{113}{24}$.

Bạn Ba:

$Q_1 = 8 + \dfrac{\frac{30}{4} - 4}{8} \cdot 2 = \dfrac{71}{8}$, $Q_3 = 12 + \dfrac{\frac{30 \cdot 3}{4} - (4+8+5)}{6} \cdot 2 = \dfrac{83}{6}$,

$\Delta_Q = Q_3 - Q_1 = \dfrac{119}{24}$.

Bạn Châu:

$Q_1 = 10 + \dfrac{\frac{30}{4} - (5+1)}{6} \cdot 2 = \dfrac{21}{2}$, $Q_3 = 14 + \dfrac{\frac{30 \cdot 3}{4} - (5+1+6+5)}{13} \cdot 2 = \dfrac{193}{13}$,

$\Delta_Q = Q_3 - Q_1 = \dfrac{113}{26}$.

Bạn Dũng:

$Q_1 = 8 + \dfrac{\frac{30}{4} - 2}{6} \cdot 2 = \dfrac{59}{6}$, $Q_3 = 14 + \dfrac{\frac{30 \cdot 3}{4} - (2+6+6+8)}{8} \cdot 2 = \dfrac{113}{8}$,

$\Delta_Q = Q_3 - Q_1 = \dfrac{103}{24}$.

Ta thấy khoảng tứ phân vị ở mẫu số liệu của bạn Dũng nhỏ nhất nên nếu so sánh theo khoảng tứ phân vị của mẫu số liệu ghép nhóm thì bạn Dũng có tốc độ giải rubik đồng đều nhất.

PHẦN II: CÂU HỎI ĐÚNG SAI

Câu 1. [NB-NB-NB-NB] Khoảng tứ phân vị của mẫu số liệu ghép nhóm là:

A. Hiệu giữa tứ phân vị thứ ba và tứ phân vị thứ nhất của mẫu số liệu ghép nhóm đó.

B. Tổng giữa tứ phân vị thứ ba và tứ phân vị thứ nhất của mẫu số liệu ghép nhóm đó.

C. Hiệu giữa hai tứ phân vị bất kì của mẫu số liệu ghép nhóm đó.

D. Tổng giữa hai tứ phân vị bất kì của mẫu số liệu ghép nhóm đó.

Lời giải

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

Câu 2. **[NB-TH-TH-VD-VD]**Thời gian (phút) truy cập Internet mỗi buổi tối của một số học sinh được cho ở 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	4	12	14	23	3

Trong các khẳng định sau, khẳng định nào đúng, khẳng định nào sai?

a/ Khoảng biến thiên của mẫu số liệu là 15 .

b/ Nhóm chứa tứ phân vị thứ ba là $[15,5;18,5)$.

c/ Tứ phân vị thứ nhất là $Q_1=15$.

d/ Khoảng tứ phân vị của mẫu số liệu ghép nhóm bé hơn 6 .

Lời giải

a/ Khoảng biến thiên của mẫu số liệu là 15 .

$R=a_6-a_1=24,5-9,5=15$.

Mệnh đề đúng .

b/ Nhóm chứa tứ phân vị thứ ba là $[15,5;18,5)$.

Cỡ mẫu $n=4+12+14+23+3=56$.

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

Mệnh đề sai .

c/ Tứ phân vị thứ nhất là $Q_1=15$.

$$Q_1=12,5+\frac{\frac{56}{4}-4}{12}.3=15.$$

Mệnh đề đúng .

d/ Khoảng tứ phân vị của mẫu số liệu ghép nhóm bé hơn 6 .

Tứ phân vị thứ ba Q_3 là $\frac{x_{42}+x_{43}}{2}$ nên nhóm chứa tứ phân vị thứ ba là $[18,5;21,5)$.

$$Q_3 = 18,5 + \frac{\frac{3.56}{4} - 30}{23}.3 = \frac{923}{46}.$$

$$\Delta_Q = Q_3 - Q_1 = \frac{233}{46} < 6$$

Vậy khoảng tứ phân vị là .

Mệnh đề đúng .

Câu 3. [NB-VD-VD-VD] Bảng tần số ghép nhóm dưới đây thống kê số giờ ngủ buổi tối của các học sinh lớp 12A1 và 12A2:

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

a) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian ngủ của các bạn nam là 5

b) Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian ngủ của các bạn nam là $2,09$

c) Khoảng tứ phân vị của mẫu số liệu ghép nhóm về thời gian ngủ của các bạn nữ trong khoảng $(2;3)$

d) Học sinh nam có thời gian ngủ đồng đều hơn.

Lời giải

a) Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian ngủ của các bạn nam là

9 – 4 = 5

Mệnh đề đúng.

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

Cỡ mẫu $n = 6 + 10 + 13 + 9 + 7 = 45$

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

Tứ phân vị thứ nhất của mẫu số liệu là $\frac{x_{11}+x_{12}}{2}$ thuộc nhóm $[5;6)$ nên nhóm chứa tứ phân vị thứ nhất là $[5;6)$. Ta có: $Q_1 = 5 + \frac{6-5}{10}\left(\frac{45}{4} - 6\right) \approx 5,53$

Tứ phân vị thứ ba của mẫu số liệu là $\frac{x_{34}+x_{35}}{2}$ thuộc nhóm $[7;8)$ nên nhóm chứa tứ phân vị thứ nhất là $[7;8)$. Ta có: $Q_3 = 7 + \frac{8-7}{9}\left(\frac{3.45}{4} - 29\right) \approx 7,53$

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

Mệnh đề sai.

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

Cỡ mẫu $n = 4 + 8 + 10 + 11 + 8 = 41$

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

Tứ phân vị thứ nhất của mẫu số liệu là $\frac{x_{10}+x_{11}}{2}$ thuộc nhóm $[5;6)$ nên nhóm chứa tứ phân vị thứ nhất là $[5;6)$. Ta có: $Q_1 = 5 + \frac{6-5}{8}\left(\frac{41}{4} - 4\right) \approx 5,78$

Tứ phân vị thứ ba của mẫu số liệu là $\frac{x_{31}+x_{32}}{2}$ thuộc nhóm $[7;8)$ nên nhóm chứa tứ phân vị thứ nhất là $[7;8)$. Ta có: $Q_3 = 7 + \frac{8-7}{11}\left(\frac{3.41}{4} - 22\right) \approx 7,80$

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

Mệnh đề đúng.

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

Mệnh đề sai.

PHẦN III: TRẢ LỜI NGẮN

Câu 1. [NB] Cho bảng số liệu ghép nhóm về chiều cao (đơn vị centimét) của 36 học sinh trong lớp 12A1 như sau:

Nhóm	$[150;155)$	$[155;160)$	$[160;165)$	$[165;170)$	$[170;175)$	$[175;180)$
Tần số	6	11	8	7	3	1

Xác định khoảng biến thiên của mẫu số liệu ghép nhóm?

Trả lời:.....

Lời giải

Trả lời: 30.

Ta có khoảng biến thiên của mẫu số liệu ghép nhóm trên là $R = a_{k+1} - a_1 = 180 - 150 = 30$.

Câu 2. [TH] Cho bảng số liệu ghép nhóm về chiều cao (đơn vị centimét) của 36 học sinh trong lớp 12A1 như sau:

Nhóm	$[150;155)$	$[155;160)$	$[160;165)$	$[165;170)$	$[170;175)$	$[175;180)$
Tần số	6	11	8	7	3	1

Nhóm chứa tứ phân vị thứ nhất có tần số bằng bao nhiêu?

Trả lời:.....

Lời giải

Trả lời: 11.

Do $\frac{n}{4} = \frac{36}{4} = 9$ nên nhóm chứa tứ phân vị thứ nhất là nhóm $[155;160)$, nên nó có tần số bằng 11.

Câu 3. [VD] Khoảng tứ phân vị của mẫu số liệu ghép nhóm được cho ở bảng sau là bao nhiêu?(làm tròn kết quả đến hai chữ số thập phân).

Nhóm	$[25;30)$	$[30;35)$	$[35;40)$	$[40;45)$
Tần số	2	17	10	25

Trả lời:.....

Lời giải

Trả lời: **8,92.**

Cỡ mẫu $n=2+17+10+25=54$.

Tứ phân vị thứ nhất Q_1 là x_{14} nên nhóm chứa tứ phân vị thứ nhất là $[30;35)$.

$$Q_1=30+\frac{\frac{54}{4}-2}{17}.5=\frac{1135}{34}.$$

Tứ phân vị thứ ba Q_3 là x_{41} nên nhóm chứa tứ phân vị thứ ba là $[40;45)$.

$$Q_3=40+\frac{\frac{3.54}{4}-29}{25}.5=\frac{423}{10}.$$

$$\Delta_Q=Q_3-Q_1=\frac{758}{85}\approx 8,92.$$

Vậy khoảng tứ phân vị là

Câu 4. [VD] Cho bảng số liệu về khối lượng của 30 củ khoai tây thu hoạch từ một thửa ruộng như dưới đây.

Khối lượng (gam)	[70;80)	[80;90)	[90;100)	[100;110)	[110;120)
Tần số	3	6	12	6	3

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là bao nhiêu?

Trả lời:.....

Lời giải

Trả lời: 15.

Cỡ mẫu $n=3+6+12+6+3=30$

Gọi $x_1;x_2;\ldots;x_{30}$ là khối lượng của 30 củ khoai tây được sắp xếp theo thứ tự tăng dần.

Tứ phân vị thứ nhất của mẫu số liệu là x_8 thuộc nhóm [80;90) nên nhóm chứa tứ phân vị thứ nhất là [80;90). Ta có:

$$Q_1=80+\frac{90-80}{6}\left(\frac{30}{4}-3\right)\approx 87,5$$

Tứ phân vị thứ ba của mẫu số liệu là x_{23} thuộc nhóm [100;110) nên nhóm chứa tứ phân vị thứ nhất là [100;110). Ta có:

$$Q_3=100+\frac{110-100}{6}\left(\frac{3.30}{4}-21\right)\approx 102,5$$

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta Q = Q_3 - Q_1 = 102,5 - 87,5 = 15$

Câu 5. [VD] Điều tra về chiều cao của học sinh khối lớp 10, ta có kết quả sau:

Nhóm	Chiều cao (cm)	Số học sinh
1	[150;152)	5
2	[152;154)	18
3	[154;156)	40
4	[156;158)	26
5	[158;160)	8
6	[160;162)	3
		$N = 100$

Tính tổng tứ phân vị thứ nhất Q_1 và tứ phân vị thứ ba Q_3.

Trả lời:.....

Lời giải

Đáp án: $\frac{40433}{130}$

Nhóm chứa tứ phân vị thứ nhất là $[154;156)$, nhóm chứa tứ phân vị thứ ba là $[156;158)$.

Ta có

$$Q_1 = 154 + \frac{\frac{1.100}{4} - (5+18)}{40}.(156-154) = \frac{1541}{10},$$

$$Q_3 = 156 + \frac{\frac{3.100}{4} - (5+18+40)}{26}.(158-156) = \frac{2040}{13}.$$

Vậy $Q_1 + Q_3 = \frac{1541}{10} + \frac{2043}{13} = \frac{40433}{130}$.

BÀI TẬP PHÁT TRIỂN(BÀI TẬP VỀ NHÀ)

PHẦN I. TRẮC NGHIỆM 4 PHƯƠNG ÁN

Câu 1.[NB] Điểm kiểm tra của nhóm học sinh lớp 10 được cho như sau:

Lớp điểm	[3;4]	[5;6]	[7;8]	[9;10]
Số học sinh	3	3	2	2

Khoảng biến thiên của mẫu số liệu ghép nhóm trên là

A. 7 **B.** 3 **C.** 5 **D.** 6

Lời giải

Ta có khoảng biến thiên của mẫu số liệu trên là

$10-3=7.$

Câu 2.[NB]Cho mẫu số liệu ghép nhóm với bộ ba tứ phân vị lần lượt là $Q_1=11,5$ $Q_2=14,5$; $Q_3=21,3.$; Khi đó khoảng tứ phân vị của mẫu số liệu trên là

A. $\Delta Q=3,0$. **B.** $\Delta Q=6,8$. **C.** $\Delta Q=9,8$. **D.** $\Delta Q=32,8$.

Lời giải

Khoảng tứ phân vị của mẫu số liệu là: $\Delta Q=Q_3-Q_1=21,3-11,5=9,8$.

Câu 3. [NB] Cho mẫu số liệu ghép nhóm về khối lượng (đơn vị: gram) của 30 củ khoai tây như sau:

Khối lượng	$[70;\ 80)$	$[80;\ 90)$	$[90;\ 100)$	$[100;\ 110)$	$[110;120)$
Tần số	3	6	12	6	3

Giá trị đại diện của nhóm $[90;100)$

A. 85. **B.** 95. **C.** 90. **D.** 100.

Lời giải

Giá trị đại diện của nhóm $[90;100)$ là: $\frac{90+100}{2}=95$.

Câu 4. [TH] Ta có bảng sau về thời gian tập thể dục buổi sáng của bác Bình và bác An:

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

Hỏi hiệu khoảng biến thiên của mẫu số liệu của bác Bình và bác An là bao nhiêu?

A. 11. **B.** 9. **C.** 15. **D.** 10.

Lời giải

Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bác Bình là:

40 – 15 = 25 (phút).

Trong mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bác An, khoảng đầu tiên chứa dữ liệu là [20; 25) và khoảng cuối cùng chứa dữ liệu là [25; 30). Do đó khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian tập thể dục buổi sáng của bác An là: 30 – 20 = 10 (phút).

Vậy hiệu khoảng biến thiên của bác Bình và bác An là: $25-10=15$.

Câu 5. [TH] Đo cân nặng của 40 học sinh lớp 12A9 ta được bảng số liệu như sau:

Khối lượng(kg)	[40;45)	[45;50)	[50;55)	[55;60)	[60;65)	[65;70)	[70;75)	[75;80]
Số học sinh	4	13	7	5	6	2	1	2

Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm thuộc khoảng nào sau đây?

A. $[40;45]$. **B.** $[45;50]$. **C.** $[50;55]$. **D.** $[55;60]$.

Lời giải

Gọi $x_1;x_2;\ldots;x_{40}$ là mẫu số liệu gốc về cân nặng của 40 học sinh lớp 12A9 được xếp theo thứ tự tăng dần.

Tứ phân vị thứ nhất của mẫu số liệu $x_1;x_2;\ldots;x_{40}$ là $x_{10}\in[45;50)$

Câu 6. [TH] Thống kê thời gian sử dụng mạng xã hội trong ngày của các bạn học sinh tổ 1 và tổ 2 lớp 12A thu được bảng sau:

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

Tìm khoảng biến thiên R_1,R_2 cho thời gian sử dụng mạng xã hội của tổ 1 và tổ 2.

A. $R_1=60;\ R_2=90$. **B.** $R_1=90;\ R_2=60$.

C. $R_1=30;\ R_2=30$. **D.** $R_1=10;\ R_2=30$.

Lời giải

Khoảng biến thiên cho thời gian sử dụng mạng xã hội của tổ 1 là $R_1=90-0=90$

Khoảng biến thiên cho thời gian sử dụng mạng xã hội của tổ 2 là $R_2=60-0=60$

Câu 7. [VD] Điều tra về khối lượng 27 củ khoai tây (đơn vị: gam) thu hoạch tại nông trường, ta có kết quả sau:

Nhóm	Tần số	Tần số tích lũy
$[74;80)$	4	4
$[80;86)$	6	10

$[86;92)$	3	13
$[92;98)$	4	17
$[98;104)$	3	20
$[104;110)$	7	27
	$n=27$	

Khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu ghép nhóm trên lần lượt là

A. $36;\ 21,45$. **B.** $7;\ 23$. **C.** $11;\ 25,3$. **D.** $33;\ 20,5$.

Lời giải

Trong mẫu số liệu ghép nhóm trên, ta có: đầu mút trái của nhóm 1 là $a_1=74$, đầu mút phải của nhóm 6 là $a_7=110$. Vậy khoảng biến thiên của mẫu số liệu ghép nhóm đó là : $R=a_7-a_1=110-74=36$ (gam)

Số phần tử của mẫu là $n=27$

Ta có: $\frac{n}{4}=\frac{27}{4}=6,75$ mà $4<6,75<10$. Suy ra nhóm 2 là nhóm đầu tiên có tần số tích luỹ lớn hơn hoặc bằng $6,75$. Xét nhóm 2 là nhóm $[80;86)$ có $s=80$; $h=6$; $n_2=6$ và nhóm 1 là nhóm $[74;80)$ có $cf_1=4$.

Áp dụng công thức, ta có tứ phân vị thứ nhất là:

$$Q_1=80+\left(\frac{6,75-4}{6}\right).6=82,75$$ (gam)

Ta có: $\frac{3n}{4}=\frac{3.27}{4}=20,25$ mà $20<20,25<27$. Suy ra nhóm 6 là nhóm đầu tiên có tần số tích luỹ lớn hơn hoặc bằng $20,25$. Xét nhóm 6 là nhóm $[104;109)$ có $t=104$; $l=6$; $n_6=7$ và nhóm 5 là nhóm $[98;104)$ có $cf_5=20$.

Áp dụng công thức, ta có tứ phân vị thứ ba là:

$$Q_3=104+\left(\frac{20,25-20}{7}\right).6=\frac{1459}{14}\approx 104,2$$ (gam)Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm đã cho là:

$$\Delta_Q=Q_3-Q_1\approx 104,2-82,75=21,45$$ (gam)

Câu 8. [VD] Thống kê chiều cao của một số cây bạch đàn giống 1 tháng tuổi của 4 nông trường được cho bởi bảng sau:

Chiều cao (cm)	$[5;7)$	$[7;9)$	$[9;11)$	$[11;13)$	$[13;15)$
Nông trường A	5	8	16	8	3
Nông trường B	5	10	8	9	6
Nông trường C	13	9	9	3	9
Nông trường D	3	12	8	12	4

Nếu xét theo khoảng tứ phân vị thì cây bạch đàn giống 1 tháng tuổi ở nông trường nào có chiều cao đồng đều nhất?

A. Nông trường A.

B. Nông trường B.

C. Nông trường C.

D. Nông trường D.

Lời giải

Nông trường A:

$n = 5+8+16+8+3 = 40$.

$Q_1 = 7 + \dfrac{\frac{40}{4}-5}{8}\cdot 2 = \dfrac{33}{4}$ $Q_3 = 11 + \dfrac{\frac{40\cdot 3}{4}-(5+8+16)}{8}\cdot 2 = \dfrac{45}{4}$,

$\Delta_Q = Q_3 - Q_1 = 3$.

Nông trường B:

$n = 5+10+8+9+6 = 38$.

$Q_1 = 7 + \dfrac{\frac{38}{4}-5}{10}\cdot 2 = \dfrac{79}{10}$ $Q_3 = 11 + \dfrac{\frac{38\cdot 3}{4}-(5+10+8)}{9}\cdot 2 = \dfrac{110}{9}$,

$\Delta_Q = Q_3 - Q_1 = \dfrac{389}{90}$.

Nông trường C:

$n = 13+9+9+3+9 = 43$.

$Q_1 = 5 + \dfrac{\dfrac{43}{4}}{13} \cdot 2 = \dfrac{173}{26}$ $Q_3 = 11 + \dfrac{\dfrac{43 \cdot 3}{4} - (13 + 9 + 9)}{3} \cdot 2 = \dfrac{71}{6}$

,

$\Delta_Q = Q_3 - Q_1 = \dfrac{202}{39}$

.

Nông trường D:

$n = 3 + 12 + 8 + 12 + 4 = 39$

.

$Q_1 = 7 + \dfrac{\dfrac{39}{4} - 3}{12} \cdot 2 = \dfrac{65}{8}$ $Q_3 = 11 + \dfrac{\dfrac{39 \cdot 3}{4} - (3 + 12 + 8)}{12} \cdot 2 = \dfrac{289}{24}$

,

$\Delta_Q = Q_3 - Q_1 = \dfrac{47}{12}$

.

Ta thấy khoảng tứ phân vị của nông trường A là nhỏ nhất nên nếu xét theo khoảng tứ phân vị thì cây bạch đàn giống 1 tháng tuổi ở nông trường A có chiều cao đồng đều nhất.

Câu 9. [VD] Thống kê điểm kiểm tra môn Toán giữa kì I của bốn lớp 12 của một trường THPT cho bởi bảng sau:

Điểm	$[5;6)$	$[6;7)$	$[7;8)$	$[8;9)$	$[9;10]$
Lớp 12B1	7	3	15	12	4
Lớp 12B2	5	9	12	11	3
Lớp 12B3	10	10	9	6	1
Lớp 12B4	14	3	15	9	1

Nhà trường muốn đánh giá mức độ "học đều" môn Toán của các lớp. Nếu xét theo khoảng tứ phân vị thì điểm kiểm tra môn Toán giữa kì I của lớp nào đồng đều nhất?

A. Lớp 12B1. **B.** Lớp 12B2. **C.** Lớp 12B3. **D.** Lớp 12B4.

Lời giải

Lớp 12B1:

$n = 7 + 3 + 15 + 12 + 4 = 41$

$Q_1 = 7 + \dfrac{\frac{41}{4} - (7+3)}{15} \cdot 1 = \dfrac{421}{60}$, $Q_3 = 8 + \dfrac{\frac{41 \cdot 3}{4} - (7+3+15)}{12} \cdot 1 = \dfrac{407}{48}$.

$\Delta_Q = Q_3 - Q_1 = \dfrac{117}{80}$.

Lớp 12B2:

$n = 5+9+12+11+3 = 40$

$Q_1 = 6 + \dfrac{\frac{40}{4} - 5}{9} \cdot 1 = \dfrac{59}{9}$, $Q_3 = 8 + \dfrac{\frac{40 \cdot 3}{4} - (5+9+12)}{11} \cdot 1 = \dfrac{92}{11}$.

$\Delta_Q = Q_3 - Q_1 = \dfrac{179}{99}$.

Lớp 12B3:

$n = 10+10+9+6+1 = 36$

$Q_1 = 5 + \dfrac{\frac{36}{4}}{10} \cdot 1 = \dfrac{59}{10}$, $Q_3 = 7 + \dfrac{\frac{36 \cdot 3}{4} - (10+10)}{9} \cdot 1 = \dfrac{70}{9}$.

$\Delta_Q = Q_3 - Q_1 = \dfrac{169}{90}$.

Lớp 12B4:

$n = 14+3+15+9+1 = 42$

$Q_1 = 5 + \dfrac{\frac{42}{4}}{14} \cdot 1 = \dfrac{23}{4}$, $Q_3 = 7 + \dfrac{\frac{42 \cdot 3}{4} - (14+3)}{15} \cdot 1 = \dfrac{239}{30}$.

$\Delta_Q = Q_3 - Q_1 = \dfrac{133}{60}$.

Ta thấy khoảng tứ phân vị của lớp 12B1 nhỏ nhất nên nếu xét theo khoảng tứ phân vị thì điểm kiểm tra môn Toán giữa kì I của lớp 12B1 đồng đều nhất.

PHẦN II: CÂU HỎI ĐÚNG SAI

Câu 1. [NB-NB-NB-NB]Cho mẫu số liệu ghép nhóm có cùng đơn vị.

a) Khoảng biến thiên của mẫu số liệu ghép nhóm đo mức độ phân tán của mẫu số liệu đó.

b) Khoảng biến thiên càng lớn thì mẫu số liệu càng ít phân tán.

c) Khoảng biến thiên của mẫu số liệu ghép nhóm không bị ảnh hưởng bởi các giá trị bất thường.

d) Khoảng biến thiên của mẫu số liệu ghép nhóm xấp xỉ cho khoảng biến thiên của mẫu số liệu gốc.

Lời giải

a) Đúng	b) Sai	c) Sai	d) Đúng

Câu a) Đúng.

Câu b) Sai – Do khoảng biến thiên càng lớn thì mẫu số liệu ghép nhóm càng phân tán.

Câu c) Sai – Do khoảng biến thiên chỉ sử dụng hai giá trị a_1 và a_{m+1} của mẫu số liệu nên dễ bị ảnh hưởng bởi các giá trị bất thường.

Câu d) Đúng.

Câu 2. [NB-NB-TH-TH] Cho bảng mẫu số liệu ghép nhóm về điểm môn Toán của hai lớp $12A$ và $12B$ được cho như sau:

Khoảng điểm	$[3;4)$	$[4;5)$	$[5;6)$	$[6;7)$	$[7;8)$	$[8;9)$	$[9;10)$
Số học sinh 12A	0	1	3	6	17	7	2
Số học sinh 12B	1	3	13	11	5	3	0

Trong các khẳng định sau , khẳng định nào đúng, khẳng định nào sai?

a) Khoảng biến thiên cho điểm môn Toán của lớp $12A$ là 7.

b) Khoảng biến thiên cho điểm môn Toán của lớp $12B$ là 6.

c) Nhóm chứa tứ phân vị thứ nhất của lớp $12A$ là nhóm $[6;7)$.

d) Nhóm chứa tứ phân vị thứ ba của lớp $12B$ là nhóm $[7;8)$.

Lời giải

a) Ta có khoảng biến thiên của điểm môn Toán của lớp $12A$ là $R_1 = 10 - 4 = 6$.
Mệnh đề sai.

b) Khoảng biến thiên cho điểm môn Toán của lớp $12B$ là $R_2 = 9 - 3 = 6$.
Mệnh đề đúng.

c) Ta có $n=1+3+13+11+5+3=36$, . Gọi $x_1,\ldots,x_{36}$ là điểm của 36 học sinh lớp $12A$ được sắp xếp theo thứ tự tăng dần. Tứ phân vị thứ ba có số liệu gốc là x_9 nên nhóm chứa phân vị thứ nhất là nhóm $[6;7)$.

Mệnh đề đúng.

d) Ta có $n=1+3+13+11+5+3=36$, . Gọi $x_1,\ldots,x_{36}$ là điểm của 36 học sinh lớp $12B$ được sắp xếp theo thứ tự tăng dần. Tứ phân vị thứ ba có số liệu gốc là x_{27} nên nhóm chứa phân vị thứ ba là nhóm $[6;7)$.

Mệnh đề sai.

Câu 3. [NB-VD-VD-VD] Bảng thống kê thời gian (đơn vị: phút) tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình và bạn An

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

a) Khoảng biến thiên của mẫu số liệu về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn An là 20.

b) Khoảng tứ phân vị của mẫu số liệu về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình là 28.

c) Khoảng tứ phân vị của mẫu số liệu về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn An là 22.

d) Dựa vào khoảng tứ phân vị của hai mẫu số liệu trên thì thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình phân tán hơn bạn An.

Lời giải

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

Mệnh đề đúng.

b) Cỡ mẫu là: 28. Gọi $x_1;x_2;\ldots;x_{28}$ là mẫu số liệu gốc về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình đã được sắp xếp theo thứ tự không giảm. Nên tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_7+x_8)$ nên nhóm chứa tứ phân vị thứ nhất là nhóm $[20;25)$ và ta có $Q_1=20+\frac{\left(\frac{1.28}{4}-5\right)}{10}.5=21$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{21}+x_{22})$ nên nhóm chứa tứ phân vị thứ ba là nhóm $[25;30)$ và ta có $Q_3 = 25 + \frac{\left(\frac{3.28}{4}-15\right)}{10}.5 = 28$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = Q_3 - Q_1 = 28 - 21 = 7$

Mệnh đề Sai.

c) Cỡ mẫu là: 28. Gọi $x_1; x_2; ...; x_{28}$ thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn An và giả sử dãy số liệu gốc này đã được sắp xếp theo thứ tự không giảm. Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_7+x_8)$ nên nhóm chứa tứ phân vị thứ nhất là nhóm $[20;25)$ và ta có

$Q_1 = 20 + \frac{\left(\frac{1.28}{4}-5\right)}{5}.5 = 22$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{21}+x_{22})$ nên nhóm chứa tứ phân vị thứ ba là nhóm $[25;30)$ và ta có $Q_3 = 25 + \frac{\left(\frac{3.28}{4}-10\right)}{15}.5 = 26$

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

Mệnh đề Sai.

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

Mệnh đề đúng.

PHẦN III: TRẢ LỜI NGẮN

Câu 1. [NB] Khoảng biến thiên của mẫu số liệu ghép nhóm được cho ở bảng sau là bao nhiêu?

Nhóm	$[15;22)$	$[22;29)$	$[29;36)$	$[36;43)$	$[43;50)$
Tần số	1	6	21	21	11

Trả lời:.....

Lời giải

Trả lời: 35.

Khoảng biến thiên của mẫu số liệu ghép nhóm ở bảng trên là $R = a_6 - a_1 = 50 - 15 = 35$

Câu 2. [TH] Cho bảng thống kê thời gian sử dụng điện thoại vào buổi sáng mỗi ngày trong tháng 4/2024 của Tuấn và An ở bảng như sau

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

Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian sử dụng điện thoại vào buổi sáng mỗi ngày của Tuấn và An lần lượt là $R_1, R_2.$ Tính tổng $R_1 + 2R_2.$

Trả lời:.....

Lời giải

Trả lời: 45.

*Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian sử dụng điện thoại vào buổi sáng mỗi ngày của Tuấn là: $40-15=25$ phút.

*Khoảng biến thiên của mẫu số liệu ghép nhóm về thời gian sử dụng điện thoại vào buổi sáng mỗi ngày của An là: $30-20=10$ phút.

* Khi đó tổng $R_1 + 2R_2 = 25 + 2.10 = 45.$

Câu 3. [VD] Bảng thống kê cân nặng của 50 quả xoài được lựa chọn ngẫu nhiên sau khi thu hoạch ở nông trường như sau

Cân nặng (g)	[250;290)	[290;330)	[330;370)	[370;410)	[410;450)
Số quả xoài	2	12	19	12	5

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

Trả lời:.....

Lời giải

***Trả lời*: 60**

Cỡ mẫu là: 50. Gọi $x_1; x_2; ...; x_{50}$ là mẫu số liệu gốc gồm cân nặng của 50 quả xoài được sắp xếp theo thứ tự không giảm. Tứ phân vị thứ nhất của mẫu số liệu gốc là x_{13} nên nhóm chứa tứ phân vị thứ nhất là nhóm $[290;330)$ và ta có

$$Q_1 = 290 + \frac{\left(\frac{1.50}{4} - 2\right)}{12}.(330-290) = 325$$

Tứ phân vị thứ ba của mẫu số liệu gốc là x_{38} nên nhóm chứa tứ phân vị thứ ba là nhóm $[370;410)$ và ta có

$$Q_3 = 370 + \frac{\left(\frac{3.50}{4} - 33\right)}{12}.40 = 385$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là $\Delta_Q = Q_3 - Q_1 = 385 - 325 = 60$.

Câu 4. [VD] An tìm hiểu hàm lượng chất béo (đơn vị: g) có trong 100 g mỗi loại thực phẩm. Sau khi thu thập dữ liệu về 60 loại thực phẩm, An lập được bảng thống kê

Hàm lượng chất béo (g)	$[2;6)$	$[6;10)$	$[10;14)$	$[14;18)$	$[18;22)$	$[22;26)$
Tần số	2	6	10	13	16	13

Tính khoảng tứ phân vị của mẫu số liệu?

Trả lời:.....

Lời giải

Trả lời: $8,7$

Hàm lượng chất béo (g)	$[2;6)$	$[6;10)$	$[10;14)$	$[14;18)$	$[18;22)$	$[22;26)$
Tần số	2	6	10	13	16	13
Tần số tích lũy	2	8	18	31	47	60

Trung vị thứ nhất và thứ ba:

$$Q_1 = 10 + \frac{15-8}{10}.4 = 12,8$$

$$Q_3 = 18 + \frac{45-31}{16}.4 = 21,5$$

Khoảng tứ phân vị của mẫu số liệu là $\Delta_Q = Q_3 - Q_1 = 21,5 - 12,8 = 8,7$

4. PHƯƠNG SAI, ĐỘ LỆCH CHUẨN CỦA MẪU SỐ LIỆU GHÉP NHÓM

Xét ví dụ:

Thầy giáo thống kê lại điểm trung bình cuối năm của các học sinh lớp và 11B ở bảng sau:
(1) Nếu so sánh theo khoảng biến thiên thì học sinh lớp nào có điểm trung bình ít phân tán hơn?
(2) Nếu so sánh theo độ lệch chuẩn thì học sinh lớp nào có điểm trung bình ít phân tán hơn?

Lời giải

(1) Nếu so sánh theo khoảng biến thiên thì học sinh lớp nào có điểm trung bình ít phân tán hơn?

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

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

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

(2) Nếu so sánh theo độ lệch chuẩn thì học sinh lớp nào có điểm trung bình ít phân tán hơn?

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

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

» ***Xét mẫu số liệu của lớp 11A:***

Cỡ mẫu là $n_1=1+11+22+6=40$.

Số trung bình của mẫu số liệu ghép nhóm là

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

Phương sai của mẫu số liệu ghép nhóm là

$$S_1^2=\frac{1}{40}\left(1\cdot 5,5^2+11\cdot 7,5^2+22\cdot 8,5^2+6\cdot 9,5^2\right)-8,3^2=0,61.$$

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

» ***Xét mẫu số liệu của lớp 11B:***

Cỡ mẫu là $n_2=6+8+14+12=40$.

Số trung bình của mẫu số liệu ghép nhóm là

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

Phương sai của mẫu số liệu ghép nhóm là

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

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

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

Chú ý

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

» Với các mẫu số liệu ghép nhóm có cùng số trung bình (hoặc xấp xỉ nhau), ta thường sử dụng phương sai và độ lệch chuẩn để so sánh mức độ phân tán của các mẫu số liệu đó.

B Các dạng bài tập

Dạng 1. Lập mẫu số liệu ghép nhóm và tính giá trị trung bình

Phương pháp

» Xét mẫu số liệu ghép nhóm được cho ở bảng sau:
Số trung bình cộng được tính như sau:
■*Bước 1:* tính giá trị đại diện mỗi nhóm .
■*Bước 2:* tính .
■*Bước 3:* .

Ví dụ 1.1.

Chiều cao của cây giống được cho bởi bảng sau:
Tính số trung bình (làm tròn đến chữ số thứ hai sau dấu phẩy) của bảng nói trên.

Lời giải

Bảng tần số ghép lớp của bảng nói trên là:

Chiều cao (cm)	$[40;44)$	$[45;49)$	$[50;54)$	$[55;59)$	$[60;64)$	$[65;69)$
Giá trị đại diện	42	47	52	57	62	67
Số cây giống	2	5	3	4	3	3

Áp dụng công thức tính số trung bình cho bảng tần số ghép lớp ta có:

$$\overline{x}=\frac{42.2+47.5+52.3+57.4+62.3+67.3}{20}=54,5\,(cm)$$

Ví dụ 1.2.

Các bạn học sinh lớp 12A trả lời 40 câu hỏi trong một bài kiểm tra. Kết quả được thống kê ở bảng sau:
Tính giá trị trung bình số câu trả lời đúng.

Lời giải

Giá trị đại diệm của mỗi nhóm như sau:

Số câu trả lời đúng	$[16;21)$	$[21;26)$	$[26;31)$	$[31;36)$	$[36;41)$
Giá trị đại diện	18,5	23,5	28,5	33,5	38,5
Số học sinh	4	6	8	18	4

Tổng số học sinh là: 40.

Vậy giá trị trung bình là

$$\overline{x}=\frac{m_1.x_1+m_2.x_2+...+m_k.x_k}{n}=\frac{4.18,5+6.23,5+8.28,5+18.33,5+4.38,5}{40}=30.$$

Ví dụ 1.3.

Trong tuần lễ bảo vệ môi trường, các em học sinh đã tiến hành thu nhặt vỏ chai nhựa để tái chế. Nhà trường thống kê kết quả thu nhặt vỏ chai nhựa của học sinh ở bảng sau:
Tính giá trị trung bình số câu trả lời đúng.

Lời giải

Ta hiệu chỉnh bảng số liệu và tính giá trị đại diện, kết quả như bảng sau:

Số vỏ chai nhựa	$[0,5;5,5)$	$[5,5;10,5)$	$[10,5;15,5)$	$[15,5;20,5)$	$[20,5;25,5)$
Giá trị đại diện	3	8	13	18	23
Số học sinh	53	82	48	39	18

Tổng số học sinh là: 240.
Vậy giá trị trung bình là

$$\overline{x}=\frac{m_1.x_1+m_2.x_2+...+m_k.x_k}{n}=\frac{53.3+82.8+48.13+39.18+18.23}{240}=10,375.$$

Dạng 2. Tính phương sai của mẫu số liệu ghép nhóm

Phương pháp

» Xét mẫu số liệu ghép nhóm được cho ở bảng sau:
Phương sai được tính như sau:
■ ***Bước 1:*** tính giá trị đại diện mỗi nhóm .
■ ***Bước 2:*** tính .
■ ***Bước 3:*** .
■ ***Bước 4:***
Hoặc : .

Ví dụ 2.1.

Cân nặng (kg) của một số quả mít trong một khu vườn được thống kê ở bảng sau:
Hãy tính phương sai của mẫu số liệu ghép nhóm trên.

Lời giải

Ta có giá trị đại diện được thể hiện trong bảng sau:

Cân nặng (kg)	$[4;6)$	$[6;8)$	$[8;10)$	$[10;12)$	$[12;14)$
Giá trị đại diện	5	7	9	11	13
Số quả mít	6	12	19	9	4

Cỡ mẫu: $n = 50$.

Số trung bình $\overline{x} = \dfrac{m_1.x_1 + m_2.x_2 + ... + m_k.x_k}{n} = \dfrac{6.5 + 12.7 + 19.9 + 9.11 + 4.13}{50} = 8,72$.

Phương sai: $s^2 = \dfrac{1}{n}\left(m_1.x_1^2 + m_2.x_2^2 + ... + m_k.x_k^2\right) - \left(\overline{x}\right)^2$

$= \dfrac{1}{50}\left(6.5^2 + 12.7^2 + 19.9^2 + 9.11^2 + 4.13^2\right) - \left(8,72\right)^2 = 4,8016$.

Ví dụ 2.2.

Thống kê điểm trung bình của hai lớp 12A và 12B được cho ở bảng sau
Nếu so sánh phương sai thì học sinh lớp nào có điểm trung bình ít phân tán hơn?

Lời giải

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

Điểm trung bình	$[5;6)$	$[6;7)$	$[7;8)$	$[8;9)$	$[9;10)$
Giá trị đại diện	5,5	6,5	7,5	8,5	9,5
Số học sinh lớp 12A	1	0	11	22	6
Số học sinh lớp 12B	0	6	8	14	12

» **Xét mẫu số liệu lớp 12A:**

Cỡ mẫu: $n_1 = 40$.

Số trung bình: $\overline{x} = \frac{m_1.x_1 + m_2.x_2 + ... + m_k.x_k}{n_1} = \frac{1.5,5 + 0.6,5 + 11.7,5 + 22.8,5 + 6.9,5}{40} = 8,3$.

Phương sai: $s_1^2 = \frac{1}{n_1}\left(m_1.x_1^2 + m_2.x_2^2 + ... + m_k.x_k^2\right) - \left(\overline{x}\right)^2$

$= \frac{1}{40}\left(1.5,5^2 + 0.6,5^2 + 11.7,5^2 + 22.8,5^2 + 6.9,5^2\right) - \left(8,3\right)^2 = 0,61$.

» **Xét mẫu số liệu lớp 12B:**

Cỡ mẫu: $n_2 = 40$.

Số trung bình: $\overline{y} = \frac{m_1.y_1 + m_2.y_2 + ... + m_k.y_k}{n_2} = \frac{0.5,5 + 6.6,5 + 8.7,5 + 14.8,5 + 12.9,5}{40} = 8,3$.

Phương sai: $s_2^2 = \frac{1}{n_2}\left(m_1.y_1^2 + m_2.y_2^2 + ... + m_k.y_k^2\right) - \overline{y}^2$

$= \frac{1}{40}\left(0.5,5^2 + 6.6,5^2 + 8.7,5^2 + 14.8,5^2 + 12.9,5^2\right) - \left(8,3\right)^2 = 1,06$.

$\longrightarrow$ Vì $s_1^2 < s_2^2$ nên điểm trung bình của học sinh lớp 12B phân tán ít hơn điểm trung bình của học sinh lớp 12A.

Ví dụ 2.3.

Tuổi thọ của một số linh kiện điện tử (đơn vị: năm) được sản xuất bởi hai phân xưởng được cho như sau:

Tính phương sai của mỗi mẫu số liệu ghép nhóm và nhận xét về độ phân tán của tuổi thọ các linh kiện điện tử được sản xuất bởi mỗi phân xưởng.

Lời giải

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

Tuổi thọ (năm)	$[1,5;2)$	$[2;2,5)$	$[2,5;3)$	$[3;3,5)$	$[3,5;4)$
Giá trị đại diện	1,75	2,25	2,75	3,25	3,75
Số linh kiện của phân xưởng 1	4	9	13	8	6
Số linh kiện của phân xưởng 2	2	8	20	7	3

» **Xét mẫu số liệu Phân xưởng 1:**

Tổng số linh kiện: $4 + 9 + 13 + 8 + 6 = 40$

Giá trị trung bình $\overline{x}_1 = \frac{1,75.4+2,25.9+2,75.13+3,25.8+3,75.6}{4+9+13+8+6} = \frac{223}{80}$

Phương sai: $s_1^2 = \frac{1}{40}\left(1,75^2.4+2,25^2.9+2,75^2.13+3,25^2.8+3,75^2.6\right)-\left(\frac{223}{80}\right)^2 = \frac{2271}{6400} \approx 0,354$

» **Xét mẫu số liệu Phân xưởng 2:**

Tổng số linh kiện: $2+8+20+7+3=40$

Giá trị trung bình $\overline{x}_2 = \frac{1,75.2+2,25.8+2,75.20+3,25.7+3,75.3}{2+8+20+7+3} = \frac{221}{80}$

Phương sai: $s_2^2 = \frac{1}{40}\left(1,75^2.2+2,25^2.8+2,75^2.20+3,25^2.7+3,75^2.3\right)-\left(\frac{221}{80}\right)^2 = \frac{1399}{6400} \approx 0,218$

$\longrightarrow$ Vì $s_1^2 > s_2^2$ nên số linh kiện phân xưởng 2 phân tán nhiều hơn số linh kiện phân xưởng 1.

 Dạng 3. Tính độ lệch chuẩn của mẫu số liệu ghép nhóm & ý nghĩa

Phương pháp

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

Phương sai được tính như sau:

■***Bước 1:*** tính giá trị đại diện mỗi nhóm .

■***Bước 2:*** tính .

■***Bước 3:*** .

■***Bước 4:*** Phương sai

Hoặc : .

■***Bước 5:*** Độ lệch chuẩn

⁂ **Ý nghĩa:** Khi hai mẫu số liệu ghép nhóm có cùng đơn vị đo và có số trung bình cộng xấp xỉ nhau, mẫu số liệu nào có độ lệch chuẩn nhỏ hơn thì mức độ phân tán so với số trung bình cộng của các số liệu trong mẫu đó sẽ thấp hơn.

Ví dụ 3.1.

Cho mẫu số liệu ghép nhóm về độ tuổi của cư dân trong một khu phố.

Tính phương sai và độ lệch chuẩn của mẫu số liệu đó. (*Kết quả làm tròn đến hàng phần mười*)

✍ ***Lời giải***

Số trung bình cộng là $\overline{x}=\frac{n_1x_1+n_2x_2+\mathrm{K}+n_kx_k}{N}$

$=\frac{25\cdot 25+20\cdot 35+20\cdot 45+15\cdot 55+14\cdot 65+6\cdot 75}{100}=44,1$ (tuổi).

Phương sai là $S^2=\frac{n_1\left(x_1-\overline{x}\right)^2+n_2\left(x_2-\overline{x}\right)^2+\mathrm{K}+n_k\left(x_k-\overline{x}\right)^2}{N}$

$S^2=\frac{25\left(25-\overline{x}\right)^2+20\left(35-\overline{x}\right)^2+20\left(45-\overline{x}\right)^2+15\left(55-\overline{x}\right)^2+14\left(65-\overline{x}\right)^2+6\left(75-\overline{x}\right)^2}{100}=244,2$.

Độ lệch chuẩn là $S=\sqrt{S^2}\approx 15,6$ (tuổi).

Ví dụ 3.2.

Cho hai mẫu số liệu ghép nhóm thống kê mức lương của hai công ty , (đơn vị là triệu đồng)

(1) Tính phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm lần lượt biểu diễn mức lương của hai công ty , . (*Kết quả làm tròn đến hàng phần mười*)

(2) Công ty nào có mức lương đồng đều hơn?

Lời giải

(1) Tính phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm lần lượt biểu diễn mức lương của hai công ty A, B. (*Kết quả làm tròn đến hàng phần mười*)

» **Xét Mức lương tại công ty A**

Số trung bình của công ty A là $\overline{x}_A=\frac{n_1x_1+n_2x_2+\mathrm{K}+n_kx_k}{N}=\frac{62}{3}\approx 20,7$ (triệu đồng).

Phương sai của công ty A là $S_A{}^2=\frac{n_1\left(x_1-\overline{x}\right)^2+n_2\left(x_2-\overline{x}\right)^2+\mathrm{K}+n_k\left(x_k-\overline{x}\right)^2}{N}=\frac{1769}{36}$.

Độ lệch chuẩn của công ty A là $S_A=\sqrt{S_A{}^2}\approx 7$ (triệu đồng).

» **Xét Mức lương tại công ty B**

Số trung bình của công ty B là $\overline{x}_B=\frac{n_1x_1+n_2x_2+\mathrm{K}+n_kx_k}{N}=\frac{229}{12}\approx 19,1$ (triệu đồng).

Phương sai của công ty B là $S_A{}^2=\frac{n_1\left(x_1-\overline{x}\right)^2+n_2\left(x_2-\overline{x}\right)^2+\mathrm{K}+n_k\left(x_k-\overline{x}\right)^2}{N}=\frac{8339}{144}$.

Độ lệch chuẩn của công ty B là $S_B=\sqrt{S_B{}^2}\approx 7,6$ (triệu đồng).

(2) Công ty nào có mức lương đồng đều hơn?

Vì độ lệch chuẩn của công ty A nhỏ hơn độ lệch chuẩn của công ty B và lương bình quân của cả hai công ty xấp xỉ nhau nên công ty A có mức lương đồng đều hơn.

Dạng 4. Sử dụng phương sai, độ lệch chuẩn đo độ rủi ro

Phương pháp

» Xét mẫu số liệu ghép nhóm được cho ở bảng sau:
Phương sai được tính như sau:
■***Bước 1:*** tính giá trị đại diện mỗi nhóm .
■***Bước 2:*** tính .
■***Bước 3:*** .
■***Bước 4:*** Phương sai
Hoặc : .
■***Bước 5:*** Độ lệch chuẩn
⁂ **Ý nghĩa:** Khi hai mẫu số liệu ghép nhóm có cùng đơn vị đo và có số trung bình cộng xấp xỉ nhau, mẫu số liệu nào có ***phương sai*** và ***độ lệch chuẩn*** cao hơn thì mức độ ***rủi ro lớn hơn***.

Ví dụ 4.1.

Giá đóng cửa của một cổ phiếu là giá của cổ phiếu đó cuối một phiên giao dịch. Bảng sau thống kê giá đóng của (đơn vị là nghìn đồng) của hai mã cổ phiếu và trong ngày giao dịch liên tiếp.
(1) Tính phương sai và độ lệch chuẩn của hai mã cổ phiếu , . (*Kết quả làm tròn đến hàng phần mười*)
(2) So sánh độ rủi ro của cổ phiếu và ?

Lời giải

(1) Tính phương sai và độ lệch chuẩn của hai mã cổ phiếu A, B.

Số trung bình của cổ phiếu A là $\overline{x}_A = \dfrac{n_1x_1 + n_2x_2 + \mathrm{K} + n_kx_k}{N} = 125,2$.

Phương sai của cổ phiếu A là $S_A^2 = \dfrac{n_1\left(x_1 - \overline{x}\right)^2 + n_2\left(x_2 - \overline{x}\right)^2 + \mathrm{K} + n_k\left(x_k - \overline{x}\right)^2}{N} = 7,4$.

Độ lệch chuẩn của cổ phiếu A là $S_A = \sqrt{S_A^{2\,2}} \approx 2,7$.

Số trung bình của cổ phiếu B là $\overline{x}_B = \dfrac{n_1x_1 + n_2x_2 + \mathrm{K} + n_kx_k}{N} = 124,8$.

Phương sai của cổ phiếu B là $S_B^2 = \frac{n_1(x_1-\overline{x})^2 + n_2(x_2-\overline{x})^2 + \mathrm{K} + n_k(x_k-\overline{x})^2}{N} = 8,1$.

Độ lệch chuẩn của cổ phiếu B là $S_B = \sqrt{S_B^2} \approx 2,9$.

(2) So sánh độ rủi ro của cổ phiếu A và B?

Vì độ lệch chuẩn về giá của cổ phiếu A nhỏ hơn độ lệch chuẩn về giá của cổ phiếu B nên cổ phiếu A có độ rủi ro thấp hơn cổ phiếu B.

Ví dụ 4.2.

Chỉ số P/B (viết tắt của từ Price to book ratio) là một chỉ số tài chính quan trọng, được sử dụng để so sánh giữa giá của cổ phiếu với giá trị ghi sổ của cổ phiếu đó. Khi chỉ số P/B cao, nghĩa là thị trường đang có nhiều kỳ vọng đối với cổ phiếu này, doanh nghiệp có tiềm năng phát triển tốt trong tương lai. Bảng sau thống kê chỉ số P/B của hai mã cổ phiếu và trong năm giao dịch liên tiếp.

(1) Tính phương sai và độ lệch chuẩn của hai mã cổ phiếu , . (*Kết quả làm tròn đến hàng phần mười*)

(2) So sánh độ rủi ro của cổ phiếu và ? Biết rằng các chỉ số tài chính còn lại của hai cổ phiếu đó là xấp xỉ nhau.

Lời giải

(1) Tính phương sai và độ lệch chuẩn của hai mã cổ phiếu A, B.

Số trung bình của cổ phiếu A là $\overline{x}_A = \frac{n_1x_1 + n_2x_2 + \mathrm{K} + n_kx_k}{N} = 6,4$.

Phương sai của cổ phiếu A là $S_A^2 = \frac{n_1(x_1-\overline{x})^2 + n_2(x_2-\overline{x})^2 + \mathrm{K} + n_k(x_k-\overline{x})^2}{N} = 7,2$.

Độ lệch chuẩn của cổ phiếu A là $S_A = \sqrt{S_A^2} \approx 2,7$.

Số trung bình của cổ phiếu B là $\overline{x}_B = \frac{n_1x_1 + n_2x_2 + \mathrm{K} + n_kx_k}{N} = 6,4$.

Phương sai của cổ phiếu B là $S_B^2 = \frac{n_1(x_1-\overline{x})^2 + n_2(x_2-\overline{x})^2 + \mathrm{K} + n_k(x_k-\overline{x})^2}{N} = 9,6$.

Độ lệch chuẩn của cổ phiếu B là $S_B = \sqrt{S_B^2} \approx 3,1$.

(2) So sánh độ rủi ro của cổ phiếu A và B?

Vì độ lệch chuẩn về chỉ số P/B của cổ phiếu A nhỏ hơn độ lệch chuẩn của cổ phiếu B nên cổ phiếu A có độ rủi ro thấp hơn cổ phiếu B.

BÀI TẬP LUYỆN TẬP

A. Câu hỏi – Trả lời trắc nghiệm

» Câu 1. 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

$$\overline{x}=\frac{19,25.13+19,75.45+20,25.24+20,75.12+21,25.6}{100}=20,015.$$

» Câu 2. 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:

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

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

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

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

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

+ Phương sai của mẫu số liệu ghép nhóm là

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

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

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

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:

$$\overline{x}=\frac{22,5.6+27,5.6+32,5.4+37,5.1+42,5.1}{18}=\frac{85}{3}.$$

+ Phương sai của mẫu số liệu ghép nhóm là

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

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

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

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:

$$\overline{x}=\frac{9.4+11.6+13.8+15.4+17.3}{25}=12,68.$$

+ Phương sai của mẫu số liệu ghép nhóm là

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

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

» Câu 6. 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:

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

+ Phương sai của mẫu số liệu ghép nhóm là

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

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

» Câu 7. 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

$$\overline{x}_A=\frac{1}{25}.\left(2.7,5+5.12,5+8.17,5+6.22,5+4.27,5\right)=18,5$$

$$\overline{x}_B=\frac{1}{25}.\left(8.7,5+4.12,5+2.17,5+5.22,5+6.27,5\right)=16,9$$

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

$$s_A=\sqrt{\frac{1}{25}.\left(2.7,5^2+5.12,5^2+8.17,5^2+6.22,5^2+4.27,5^2\right)-18,5^2}=5,8$$

$$s_B=\sqrt{\frac{1}{25}.\left(8.7,5^2+4.12,5^2+2.17,5^2+5.22,5^2+6.27,5^2\right)-16,9^2}=8,04.$$

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 8. 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
Số trung bình của mẫu số liệu ghép nhóm là:

$$\overline{x}_1=\frac{8.121+9.123+12.125+10.127+11.129}{50}=125,28.$$

Phương sai của mẫu số liệu ghép nhóm là

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

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

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

Phương sai của mẫu số liệu ghép nhóm là

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

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 cồ phiếu B.

» Câu 9. 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

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

» **Câu 10.** 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

Số trung bình: $\overline{x}=\frac{42,5.4+47,5.14+52,5.8+57,5.10+62,5.6+67,5.2}{44}=\frac{585}{11}\approx 53,2$

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

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

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

$$\overline{x}=\frac{1}{20}.(2,85.3+3,15.6+3,45.5+3,75.4+4,05.2)=3,39.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2=\frac{1}{20}\left(3.2,85^2+6.3,15^2+5.3,45^2+4.3,45^2+2.4,05^2\right)-3,39^2=0,1314.$$

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

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

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: $\overline{x}=\frac{1}{18}.(22,5.6+27,5.6+32,5.4+37,5.1+42,5.1)=\frac{85}{3}.$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2=\frac{1}{18}\left(6.22,5^2+6.27,5^2+4.32,5^2+1.37,5^2+1.42,5^2\right)-\left(\frac{85}{3}\right)^2=31,25.$$

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

» Câu 13. 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: $\overline{x}=\frac{1}{35}.(3.4+7.8+11.13+15.6+19.4)\approx 10,77$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2=\frac{1}{35}\left(4.3^2+8.7^2+13.11^2+6.15^2+4.19^2\right)-10,77^2\approx 21,01$$

» Câu 14. 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:

$$\overline{x}=\frac{9.15+10.45+9.75+15.105+7.135}{50}=75,6$$ (phút).

Phương sai của mẫu số kiệu ghép nhóm là

$$S^2=\frac{1}{50}\left(9.15^2+10.45^2+9.75^2+15.105^2+7.135^2\right)-75,6^2=1601,64$$

» Câu 15. 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

$$\overline{x}=\frac{1}{56}.(3.11+12.14+15.17+24.20+2.23)\approx 17,54$$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2=\frac{1}{56}\left(3.11^2+12.14^2+15.17^2+24.20^2+2.23^2\right)-17,54^2\approx 8,56$$

$$S=\sqrt{8,56}\approx 2,93$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là

» Câu 16. 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: $\overline{x}=\frac{1}{85}.(10.3+14.7+31.11+2.15+5.19+23.23)\approx 13,21$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2=\frac{1}{85}.(10.3^2+14.7^2+31.11^2+2.15^2+5.19^2+23.23^2)-13,21^2\approx 48,43$$

$$S=\sqrt{48,43}\approx 6,96$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là

» Câu 17. 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

$$\overline{x}=\frac{1}{20}.(3.89+6.107+3.125+5.143+3.161)=124,1$$

Phương sai của mẫu số liệu ghép nhóm là

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

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

Độ lệch chuẩn của mẫu số liệu ghép nhóm là

» Câu 18. 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	8	9	12	10	11

của cổ phiếu A					
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

$$\overline{x}_A = \frac{1}{50}.(8.121+9.123+12.125+10.127+11.129) = 125,28$$

Phương sai của mẫu số liệu ghép nhóm là

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

Độ lệch chuẩn của mẫu số liệu ghép nhóm là

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

Xét mẫu số liệu của cổ phiếu B

Số trung bình của mẫu số liệu là

$$\overline{x}_B = \frac{1}{50}.(16.121+4.123+3.125+6.127+21.129) = 125,48$$

Phương sai của mẫu số liệu ghép nhóm là

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

Độ lệch chuẩn của mẫu số liệu ghép nhóm là

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

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 19. 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

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

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 $s=\sqrt{\overline{x^2}-\left(\overline{x}\right)^2}\approx\sqrt{5,2}\approx 2,3$.

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

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

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 mấu số liệu là $s=\sqrt{\overline{x^2}-\left(\overline{x}\right)^2}\approx\sqrt{7,7}\approx 2,8$.

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

» Câu 20. 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:

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

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

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

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

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

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 21. 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

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

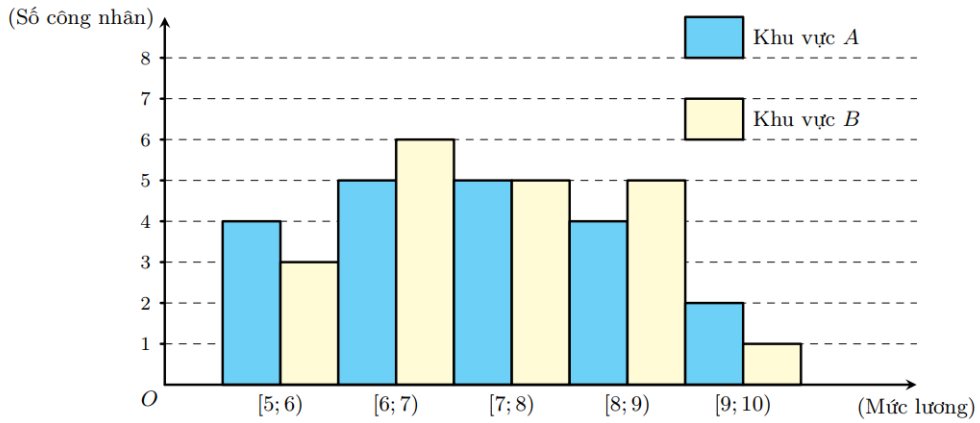
Phương sai

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

.

» Câu 22. 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

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

Phương sai của mẫu số liệu ghép nhóm là

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

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

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

Phương sai của mẫu số liệu ghép nhóm là

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

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

Tổng: khoảng $2,3647$.

» Câu 23. Bộ phận kiểm tra chất lượng sản phẩm dùng máy để đo (chính xác đến $0,001\text{mm}$) 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

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

Phương sai của mẫu số liệu là

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

Độ lệch chuẩn của mẫu số liệu là $S^2=\sqrt{\frac{179}{225}}=\frac{\sqrt{179}}{15}\approx 0,89$.

» Câu 24. 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 và 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

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

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

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

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

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

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

B. Câu hỏi – Trả lời Đúng/sai

» **Câu 25.** 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:

	Mệnh đề	Đúng	Sai
(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 $\overline{x}=\frac{20.4+30.6+40.10+50.7+60.3}{30}$.		
(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* $\overline{x}=\frac{20.4+30.6+40.10+50.7+60.3}{30}$.

Công thức tính độ tuổi bình quân của khách hàng mua bảo hiểm trong bảng thống kê là

$$\overline{x}=\frac{25.4+35.6+45.10+55.7+65.3}{30}$$

» **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

$$\overline{x}=\frac{25.4+35.6+45.10+55.7+65.3}{30}\approx 44,67.$$

» **Chọn** ĐÚNG.

» **Câu 26.** Xét mẫu dữ liệu ghép nhóm được cho ở bảng sau với $n=n_1+n_2+\ldots+n_k$.

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:

	Mệnh đề	Đúng	Sai

(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 $\overline{x}=\frac{1}{n}(c_1+c_2+\text{L}+c_k)$.		
(c)	Phương sai của mẫu số liệu ghép nhóm là $s^2=\frac{1}{n}\left[n_1(c_1-\overline{x})^2+n_2(c_2-\overline{x})^2+\text{L}+n_k(c_k-\overline{x})^2\right]$.		
(d)	Phương sai của mẫu số liệu ghép nhóm là $s^2=\frac{1}{n}\left(n_1c_1^2+n_2c_2^2+\ldots+n_kc_k^2\right)-\overline{x}$.		

✍ 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* $\overline{x}=\frac{1}{n}(c_1+c_2+\text{L}+c_k)$.

Giá trị trung bình của mẫu số liệu ghép nhóm là $\overline{x}=\frac{1}{n}(n_1c_1+n_2c_2+\text{L}+n_kc_k)$.

» **Chọn** SAI.

(c) *phương sai của mẫu số liệu ghép nhóm là* $s^2=\frac{1}{n}\left[n_1(c_1-\overline{x})^2+n_2(c_2-\overline{x})^2+\text{L}+n_k(c_k-\overline{x})^2\right]$.

» **Chọn** ĐÚNG.

(d) *phương sai của mẫu số liệu ghép nhóm là* $s^2=\frac{1}{n}\left(n_1c_1^2+n_2c_2^2+\ldots+n_kc_k^2\right)-\overline{x}$.

Phương sai của mẫu số liệu ghép nhóm là

$$s^2=\frac{1}{n}\left(n_1c_1^2+n_2c_2^2+\ldots+n_kc_k^2\right)-\overline{x}^2.$$

» **Chọn** SAI.

» **Câu 27.** 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:

	Mệnh đề	Đúng	Sai
(a)	Giá trị đại diện của lớp $[36;41)$ là 38,5.		
(b)	Công thức tính số trung bình là $\overline{x}=\frac{18,5.4+23,5.6+28,5.8+33,5.18+38,5.4}{40}$.		
(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* $\overline{x}=\frac{18,5.4+23,5.6+28,5.8+33,5.18+38,5.4}{40}$.

» **Chọn** ĐÚNG.

(c) *số trung bình là 30.*

số trung bình là $\overline{x}=\frac{18,5.4+23,5.6+28,5.8+33,5.18+38,5.4}{40}=30$.

» **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 28.** 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:

	Mệnh đề	**Đúng**	**Sai**
(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).*

$$\overline{x}=\frac{25.3+35.9+45.6+55.4+65.2}{24}=\frac{1010}{24}\approx 42,08$$

» **Chọn** ĐÚNG.

(d) *Phương sai của mẫu số liệu ghép nhóm là* $129,99$ *(làm tròn đến hàng phần trăm).*

$$S^2=\frac{1}{24}\left(3.25^2+9.35^2+6.45^2+4.55^2+2.65^2\right)-\left(\frac{1010}{24}\right)^2\approx 128,99$$

» **Chọn** SAI.

» **Câu 29.** 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:

	Mệnh đề	Đúng	Sai
(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 $n_1 = 2+10+6+4+3 = 25$

Số trung bình: $\overline{x_1} = \dfrac{5,5.2+15,5.10+25,5.6+35,5.4+45,5.3}{25} = 23,9$

Phương sai:

$$S_1^2 = \frac{1}{25}\left(2.5,5^2+10.15,5^2+6.25,5^2+4.35,5^2+3.45,5^2\right)-23,9^2 = 133,44$$

» **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:

Cỡ mẫu là $n_2 = 3+8+10+2+4 = 27$

Số trung bình: $\overline{x_2} = \dfrac{5,5.3+15,5.8+25,5.10+35,5.2+45,5.4}{27} = \dfrac{648,5}{27} \approx 24,02$

$$S_2^2 = \frac{1}{27}\left(3.5,5^2+8.15,5^2+10.25,5^2+2.35,5^2+4.45,5^2\right)-24,02^2 \approx 138,47$$

$$S_2 \approx \sqrt{138,47} \approx 11,77$$

» **Chọn** SAI.

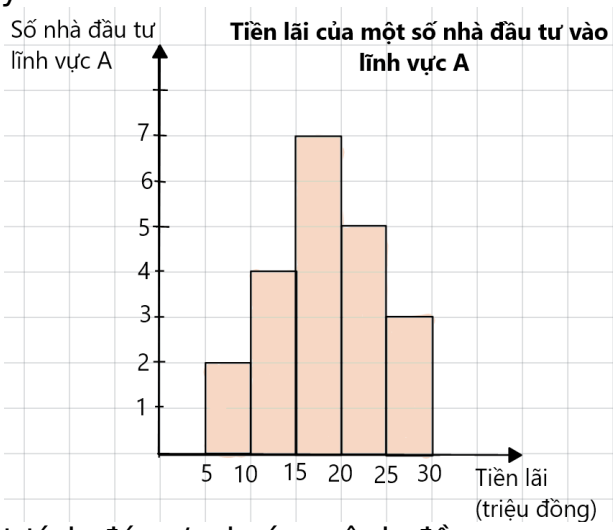
(d) *Nếu so sánh theo độ lệch chuẩn thì thời gian để học sinh hoàn thành một câu hỏi thi của lớp 11A ít phân tán hơn lớp 11B.*

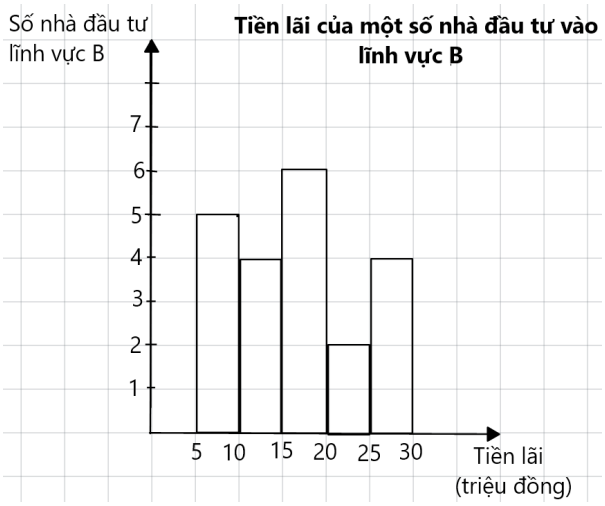
Ta có: $S_1 < S_2$

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 30. 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:

	Mệnh đề	Đúng	Sai
(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 $n_1 = 2+4+7+5+3 = 21$

Số trung bình: $\overline{x_1} = \dfrac{7,5.2+12,5.4+17,5.7+22,5.5+27,5.3}{21} = \dfrac{255}{14}$

Phương sai:

$$S_1^2=\frac{1}{21}\left(2.7,5^2+4.12,5^2+7.17,5^2+5.22,5^2+3.27,5^2\right)-\left(\frac{255}{14}\right)^2=\frac{5000}{147}$$

$$S_1=\sqrt{\frac{5000}{147}}\approx 5,83$$

» **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 $n_2=5+4+6+2+4=21$

Số trung bình: $\overline{x_2}=\frac{7,5.5+12,5.4+17,5.6+22,5.2+27,5.4}{21}=\frac{695}{42}$

$$S_2^2=\frac{1}{21}\left(5.7,5^2+4.12,5^2+6.17,5^2+2.22,5^2+4.27,5^2\right)-\left(\frac{695}{42}\right)^2=\frac{21650}{441}$$

$$S_2=\sqrt{\frac{21650}{441}}\approx 7,01$$

» **Chọn** ĐÚNG.

(c) *Về trung bình, đầu tư vào lĩnh vực B đem lại tiền lãi cao hơn lĩnh vực A.*

Số trung bình: $\overline{x_1}=\frac{7,5.2+12,5.4+17,5.7+22,5.5+27,5.3}{21}=\frac{255}{14}\approx 18,21$

Số trung bình: $\overline{x_2}=\frac{7,5.5+12,5.4+17,5.6+22,5.2+27,5.4}{21}=\frac{695}{42}\approx 16,55$

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 31.** 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:

<table>
<tr><th></th><th>Mệnh đề</th><th>Đúng</th><th>Sai</th></tr>
<tr><td>(a)</td><td>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.</td><td></td><td></td></tr>
<tr><td>(b)</td><td>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
<table>
<tr><td>Chỉ số Cholesterol toàn phần do bác sĩ A điều trị</td><td>$[3,7;4,14)$</td><td>$[4,14;4,58)$</td><td>$[4,58;5,02)$</td><td>$[5,02;5,46)$</td><td>$[5,46;5,9)$</td></tr>
<tr><td>Số bệnh nhân</td><td>4</td><td>3</td><td>5</td><td>3</td><td>3</td></tr>
</table></td><td></td><td></td></tr>
<tr><td>(c)</td><td>Giá trị độ lệch chuẩn của mẫu số liệu ghép nhóm đầu tiên là</td><td></td><td></td></tr>
</table>

	$[3,7;4,14)$ 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 $x_1;x_2;....;x_{18}$ 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.

Trung vị $Q_2=\frac{1}{2}\left(x_9+x_{10}\right)=\frac{1}{2}\left(4,6+4,7\right)=4,65$.

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 $\Delta Q=Q_3-Q_1=5,2-4,2=1$.

Số trung bình của mẫu số liệu trên do bác sĩ A điều trị bằng

$$\overline{x}=\frac{2.3,8+4,0+4,1+4,2+4,3+4,4+4,5+4,6+4,7+4,8+5,0+5,1+5,2+5,3+2.5,6+5,8}{18}=\frac{212}{45}$$

Phương sai của mẫu số liệu trên do bác sĩ A điều trị bằng

$$S^2=\frac{x_1^2+x_2^2+x_3^2+....+x_{18}^2}{18}-\overline{x}^2=\frac{3023}{8100}.$$

Độ lệch chuẩn của mẫu số liệu trên do bác sĩ A điều trị bằng $s=\sqrt{S^2}=0,61$.

» **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

Số trung bình của mẫu số liệu trên do bác sĩ A điều trị bằng

$$\overline{x_A}=\frac{4.3,92+4.4,36+4.4,8+3.5,24+3.5,68}{18}=\frac{709}{150}$$

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 $s_A=\sqrt{S_A^2}=0,609$.

» **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

$$\overline{x_B}=\frac{3.3,92+4.4,36+3.4,8+4.5,24+4.5,68}{18}=\frac{1091}{225}$$

Phương sai của mẫu số liệu trên do bác sĩ B điều trị bằng

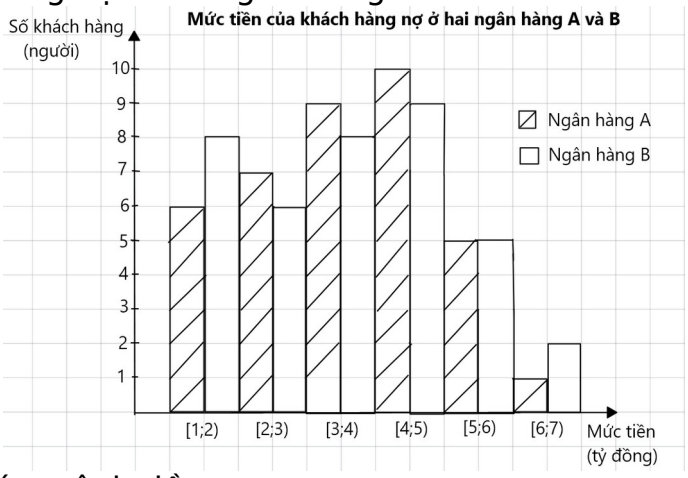
$$S_B^2=\frac{3.3,92^2+4.4,36^2+3.4,8^2+4.5,24^2+4.5,68^2}{18}-\left(\frac{1091}{225}\right)^2\approx 0,3848.$$

Độ lệch chuẩn của mẫu số liệu trên do bác sĩ B điều trị bằng $s_B=\sqrt{S_B^2}=0,62$.

Vì $s_A<s_B$ 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 32.** 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:

	Mệnh đề	**Đúng**	**Sai**
(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

$$\overline{x}_A = \frac{6.1,5+7.2,5+9.3,5+10.4,5+5.5,5+1.6,5}{38} = \frac{137}{38}$$

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 $s_A = \sqrt{S_A^2} = \frac{\sqrt{661}}{19}$.

» **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

$$\overline{x}_B = \frac{8.1,5+6.2,5+8.3,5+9.4,5+5.5,5+2.6,5}{38} = \frac{68}{19}$$

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 $s_B = \sqrt{S_B^2} = \sqrt{\frac{3221}{1444}}$.

» **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 $s_A < s_B$ nên rủi ro của ngân hàng A thấp hơn rủi ro của ngân hàng B khi cho khách hàng vay nợ.

» **Chọn** SAI.

C. Câu hỏi – Trả lời ngắn

» Câu 33. 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 $S = 43x + 47y + 1000z = 43.3 + 47.3 + 1000.3 = 3270$.

» Câu 34. 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: $\overline{x} = \dfrac{2.6 + 7.8 + 7.10 + 3.12 + 1.14}{20} = 9,4$

» Câu 35. 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:

$$\overline{x}=\frac{8\cdot 1,5+10\cdot 2,5+12\cdot 3,5+9\cdot 4,5+3\cdot 5,5}{42}=\frac{68}{21}\approx 3,238.$$

Phương sai là:

$$S^2=\frac{1}{42}\left[8\cdot(1,5)^2+10\cdot(2,5)^2+12\cdot(3,5)^2+9\cdot(4,5)^2+3\cdot(5,5)^2\right]-\left(\frac{68}{21}\right)^2=\frac{2525}{1764}\approx 1,431.$$

» Câu 36. 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:

$$\overline{x}=\frac{3\cdot 25+5\cdot 35+5\cdot 45+8\cdot 55+7\cdot 65+5\cdot 75+3\cdot 85+4\cdot 95}{40}=59,5.$$

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 37. 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:

$$\overline{x}=\frac{5\cdot 10+5\cdot 30+6\cdot 50+2\cdot 70+2\cdot 90}{20}=41.$$

Phương sai là:

$$S^2=\frac{1}{20}\left[5\cdot(10)^2+5\cdot(30)^2+6\cdot(50)^2+2\cdot(70)^2+2\cdot(90)^2\right]-(41)^2=619.$$

Độ lệch chuẩn: $S=\sqrt{619}\approx 24,88$.

» Câu 38. 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:

$$\overline{x}=\frac{29\cdot 100+11\cdot 300+3\cdot 500+4\cdot 700+3\cdot 900}{50}=264.$$

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 39. 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:

$$\overline{x}=\frac{3\cdot 44+7\cdot 48+4\cdot 52+3\cdot 56+3\cdot 60}{20}=51,2.$$

Phương sai của mẫu số liệu ghép nhóm là:

$$S^2=\frac{1}{20}\left(3\cdot 44^2+7\cdot 48^2+4\cdot 52^2+3\cdot 56^2+3\cdot 60^2\right)-(51,2)^2\approx 26,6.$$

» Câu 40. 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: $n_A = 25+38+20+10+7 = 100;\ n_B = 22+27+19+18+14 = 100$

Đường kính trung bình của thân cây xoan đào trồng tại địa điểm A là:

$$\overline{x}_A = \frac{25\cdot 31 + 38\cdot 33 + 20\cdot 35 + 10\cdot 37 + 7\cdot 39}{100} = 33,72$$

Phương sai của mẫu số liệu ghép nhóm vè đường kính của thân cây xoan đào trồng tại địa điểm A là:

$$S_A^2 = \frac{1}{100}\left(25\cdot 31^2 + 38\cdot 33^2 + 20\cdot 35^2 + 10\cdot 37^2 + 7\cdot 39^2\right) - \left(33,72\right)^2 \approx 5,40$$

Đường kính trung bình của thân cây xoan đào trồng tại địa điểm B là:

$$\overline{x}_B = \frac{22\cdot 31 + 27\cdot 33 + 19\cdot 35 + 18\cdot 37 + 14\cdot 39}{100} = 34,5$$

Phương sai của mẫu số liệu ghép nhóm về đường kính của thân cây xoan đào trồng tại địa điểm B là:

$$S_B^2 = \frac{1}{100}\left(22\cdot 31^2 + 27\cdot 33^2 + 19\cdot 35^2 + 18\cdot 37^2 + 14\cdot 39^2\right) - \left(34,5\right)^2 = 7,31$$

Vậy $\left|S_A^2 - S_B^2\right| = \left|5,40 - 7,31\right| = 1,91$

» Câu 41. 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 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:

$$\overline{x}_X = \frac{8\cdot 6,5 + 10\cdot 7,5 + 13\cdot 8,5 + 10\cdot 9,5 + 9\cdot 10,5}{50} = 8,54$$

Phương sai của mẫu số liệu ghép nhóm của trường X là:

$$S_X^2 = \frac{1}{50}\left[8\cdot\left(6,5\right)^2 + 10\cdot\left(7,5\right)^2 + 13\cdot\left(8,5\right)^2 + 10\cdot\left(9,5\right)^2 + 9\cdot\left(10,5\right)^2\right] - \left(8,54\right)^2 = 1,76$$

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:

$$\overline{x}_Y = \frac{4\cdot 6,5 + 12\cdot 7,5 + 17\cdot 8,5 + 14\cdot 9,5 + 3\cdot 10,5}{50} = 8,5$$

Phương sai của mẫu số liệu ghép nhóm của trường Y là:

$$S_Y^2 = \frac{1}{50}\left[4\cdot(6,5)^2 + 12\cdot(7,5)^2 + 17\cdot(8,5)^2 + 14\cdot(9,5)^2 + 3\cdot(10,5)^2\right] - (8,5)^2 = 1,08$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm của trường Y là:

$$S_Y = \sqrt{S_Y^2} = \sqrt{1,08} \approx 1,04$$

Vậy $S_X - S_Y = 1,33 - 1,04 = 0,29$

» Câu 42. 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

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

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: $S^2 = \frac{1}{40}\left(3\cdot 6,34^2 + 7\cdot 6,58^2 + 5\cdot 6,82^2 + 20\cdot 7,06^2 + 5\cdot 7,30^2\right) - (6,92)^2$

$$= \frac{2,9824}{40} \approx 0,07$$

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

» Câu 43. 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 $D = S_A - S_B$ nếu $D > 0$ thì đầu tư vào lĩnh

vực A "rủi ro" hơn, ngược lại nếu $D<0$ thì đầu tư vào lĩnh vực B "rủi ro" hơn.

Tính $D=S_A-S_B$ (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

$$\overline{x}_A=\frac{20\cdot 7,5+5.12,5+10.17,5+5.22,5+20\cdot 27,5}{60}=17,5$$ (triệu đồng)

Phương sai của mẫu số liệu ghép nhóm là

$$S_A^2=\frac{1}{60}\left(20\cdot 7,5^2+5.12,5^2+10.17,5^2+5.22,5^2+20\cdot 27,5^2\right)-(17,5)^2=70,83$$

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

$$\overline{x}_B=\frac{5\cdot 7,5+10.12,5+30.17,5+10.22,5+5\cdot 27,5}{60}=17,5$$ (triệu đồng)

Phương sai của mẫu số liệu ghép nhóm là

$$S_B^2=\frac{1}{60}\left(5\cdot 7,5^2+10.12,5^2+30.17,5^2+10.22,5^2+5\cdot 27,5^2\right)-(17,5)^2=25$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_B=\sqrt{S_B^2}=\sqrt{25}=5$.

$\Rightarrow D=S_A-S_B=8,42-5=3,42$

» Câu 44. 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 $D=S_A-S_B$ nếu $D>0$ thì đầu tư vào lĩnh vực A "rủi ro" hơn, ngược lại nếu $D<0$ thì đầu tư vào lĩnh vực B "rủi ro" hơn.

Tính $D=S_A-S_B$ (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

$$\overline{x}_A=\frac{8.121+9\cdot 123+12\cdot 125+10\cdot 127+11\cdot 129}{50}=125,28.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_A^2 = \frac{1}{50}\left(8.121^2 + 9.123^2 + 12.125^2 + 10.127^2 + 11.129^2\right) - (125,28)^2 \approx 7,5$$

.

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

Phương sai của mẫu số liệu ghép nhóm là

$$S_B^2 = \frac{1}{50}\left(16 \cdot 121^2 + 4.123^2 + 3.125^2 + 6.127^2 + 21.129^2\right) - (125,48)^2 \approx 12,4$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_B = \sqrt{S_B^2} = \sqrt{12,4} \approx 3.5$.

$\Rightarrow D = S_A - S_B = 2,7 - 3,5 = -0,8$

III. BÀI TẬP CỦNG CỐ KIẾN THỨC:

Câu 1: Cho mẫu số liệu ghép nhóm về thống kê chiều cao của 35 cây bạch đàn trong rừng, ta có bảng số liệu sau:

Khoảng chiều cao (m)	$[6,5; 7,0)$	$[7,0; 7,5)$	$[7,5; 8,0)$	$[8,0; 8,5)$
Số cây	6	15	11	3

Tính chiều cao trung bình của 35 cây bạch đàn trên.

A. $7,407(m)$. **B.** $4,707(m)$. **C.** $7,704(m)$. **D.** $7,5(m)$.

Câu 2: Tìm hiểu thời gian làm bài tập về nhà trong tuần trước của một số học sinh thu được kết quả như sau:

Thời gian	$[18;20)$	$[20;22)$	$[22;24)$
Số học sinh	16	35	8

Tính thời gian làm bài tập về nhà trung bình trong tuần trước của các bạn học sinh này.

A. 20,73 giờ. **B.** 21,73 giờ. **C.** 20,72 giờ. **D.** 21,72 giờ.

Câu 3: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một của 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 4: Thống kê nhiệt độ tại một địa điểm trong 40 ngày, ta có bảng số liệu sau:

Nhiệt độ $(^\circ C)$	$[19;22)$	$[22;25)$	$[25;28)$	$[28;31)$
Số ngày	7	15	12	6

Nhiệt độ trung bình là?

A. $25,456$. **B.** $24,775$. **C.** $23,021$. **D.** $22,036$.

Câu 5: Thống kê cân nặng của học sinh lớp 11D cho trong bảng dưới đây:

Cân	$[40,5;45,5)$	$[45,5;50,5)$	$[50,5;55,5)$	$[55,5;60,5)$	$[60,5;65,5)$	$[65,5;70,5)$

nặng						
Số học sinh	10	7	16	4	2	3

Tính cân nặng trung bình của học sinh lớp 11D?

A. $50,1$. **B.** $52,83$. **C.** $50,81$. **D.** $51,81$.

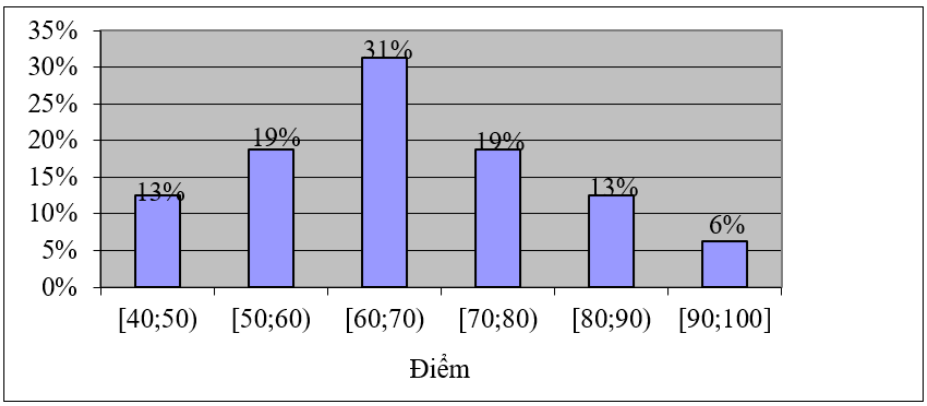
Câu 6: Khảo sát thời gian tập thể dục của một số học sinh khối 11 thu được mẫu số liệu ghép nhóm sau:

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

Mốt của mẫu số liệu trên là

A. 42. **B.** 52. **C.** 53. **D.** 54.

Câu 7: Kết quả điểm thi của 32 học sinh trong kì thi Tiếng Anh (thang điểm 100) được biểu diễn ở biểu đồ dưới đây:



Hãy ước lượng mốt của mẫu số liệu trên?

A. 55. **B.** 65. **C.** 72. **D.** 68.

Câu 8: Khảo sát thời gian tự học bài ở nhà của một số em học sinh lớp 11 thu được mẫu ghép nhóm số lượng như sau:

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

Mốt của số liệu ghép nhóm trên là

A. $102,86$. **B.** $102,85$. **C.** $102,8$. **D.** $102,9$.

Câu 9: Thời gian 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	2	10	6	4	3

Tìm mốt của mẫu số liệu ghép nhóm này.

A. $17,42$. **B.** $14,56$. **C.** $17,16$. **D.** $12,67$.

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

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

Tìm mốt của mẫu số liệu ghép nhóm này.

A. $18,37$. **B.** $19,37$. **C.** $19,73$. **D.** $18,73$.

Câu 11: Thời gian đọc sách mỗi ngày của một số học sinh được cho trong bảng sau

Thời gian (phút)	$[10;15)$	$[15;20)$	$[20;25)$	$[25;30)$	$[30;35)$
Số học sinh	3	10	12	15	20

Xác định trung vị của mẫu số liệu ghép nhóm.

A. 25. **B.** 26. **C.** $25,56$. **D.** $26,67$.

Câu 12: Cho mẫu số liệu ghép nhóm về chiều cao của 25 cây dừa giống như sau:

Chiều cao (cm)	$[0;10)$	$[10;20)$	$[20;30)$	$[30;40)$	$[40;50)$
Số cây	4	6	7	5	3

Trung vị của mẫu số liệu ghép nhóm này là

A. $M_e = \frac{175}{7}$. **B.** $M_e = \frac{165}{5}$. **C.** $M_e = \frac{165}{7}$. **D.** $M_e = \frac{165}{3}$.

Câu 13: Cho mẫu số liệu ghép nhóm về thống kê điểm số của 100 học sinh tham dự kỳ thi học sinh giỏi toán, ta có bảng số liệu sau:

Điểm	$[8;10)$	$[10;12)$	$[12;14)$	$[14;16)$	$[16;18)$	$[18;20)$
Số học sinh	6	21	30	25	14	4

Tìm trung vị của mẫu số liệu ghép nhóm trên.

A. $12,18$. **B.** $12,81$. **C.** $13,35$. **D.** $13,53$.

Câu 14: Cho mẫu số liệu ghép nhóm về thống kê huyết áp của 20 người, ta có bảng số liệu sau:

Huyết áp (mmHg)	$[70;80)$	$[80;90)$	$[90;100)$	$[100;110)$	$[110;120)$	$[120;130)$
Số người	4	2	3	6	3	2

Tìm trung vị của mẫu số liệu ghép nhóm trên.

A. $100,67$. **B.** $101,67$. **C.** $101,76$. **D.** $100,76$.

Câu 15: Khảo sát thời gian tập thể dục của một số học sinh khối 11 thu được mẫu số liệu ghép nhóm sau:

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

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

A. $[40;60)$. **B.** $[20;40)$. **C.** $[60;80)$. **D.** $[80;100)$.

Câu 16: Khảo sát thời gian tập thể dục của một số học sinh khối 11 thu được mẫu số liệu ghép nhóm sau:

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

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

A. $[40;60)$. **B.** $[20;40)$. **C.** $[60;80)$. **D.** $[80;100)$.

Câu 17: Tìm hiểu thời gian hoàn thành một bài tập của một số học sinh thu được kết quả sau:

Thời gian (phút)	$[0;4)$	$[4;8)$	$[8;12)$	$[12;16)$	$[16;20)$
Số học sinh	2	4	7	4	3

Tứ phân vị thứ ba của mẫu số liệu ghép nhóm này là

A. $Q_3 = 13$. **B.** $Q_3 = 14$. **C.** $Q_3 = 15$. **D.** $Q_3 = 12$.

Câu 18: Cho mẫu số liệu ghép nhóm về chiều cao của 25 cây dừa giống như sau:

Chiều cao (cm)	$[0;10)$	$[10;20)$	$[20;30)$	$[30;40)$	$[40;50)$
Số cây	4	6	7	5	3

Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm này là

A. $Q_1 = 13,5$. **B.** $Q_1 = 13,9$. **C.** $Q_1 = 15,75$. **D.** $Q_1 = 13,75$.

Câu 19: Tìm hiểu thời gian hoàn thành một bài tập của một số học sinh thu được kết quả sau:

Thời gian (phút)	$[0;4)$	$[4;8)$	$[8;12)$	$[12;16)$	$[16;20)$
Số học sinh	2	4	7	4	3

Tứ phân vị thứ ba của mẫu số liệu ghép nhóm này là

A. $Q_3=13$. **B.** $Q_3=14$. **C.** $Q_3=15$. **D.** $Q_3=12$.

Câu 20: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một của hàng được ghi lại ở bảng sau:

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

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

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

Câu 21: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một của hàng được ghi lại ở bảng sau:

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

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

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

Câu 22: Tìm hiểu thời gian hoàn thành một bài tập của một số học sinh thu được kết quả sau:

Thời gian (phút)	$[0;4)$	$[4;8)$	$[8;12)$	$[12;16)$	$[16;20)$
Số học sinh	2	4	7	4	3

Tứ phân vị thứ ba của mẫu số liệu ghép nhóm này là

A. $Q_3=13$. **B.** $Q_3=14$. **C.** $Q_3=15$. **D.** $Q_3=12$.

Câu 23: Cô Hà thống kê lại đường kính thân gỗ của một số cây xoan đào 6 năm tuổi được trồng ở một lâm trường ở bảng sau.

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

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

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

Câu 24: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khoẻ. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

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

Khoảng biến thiên của mẫu số liệu ghép nhóm là

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

Câu 25: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khoẻ. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

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

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

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

Câu 26: Bạn Chi rất thích nhảy hiện đại. Thời gian tập nhảy mỗi ngày trong thời gian gần đây của bạn Chi được thống kê lại ở bảng sau:

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

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

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

Câu 27: Dũng là học sinh rất giỏi chơi rubik, bạn có thể giải nhiều loại khối rubik khác nhau. Trong một lần tập luyện giải khối rubik 3×3, bạn Dũng đã tự thống kê lại thời gian giải rubik trong 25 lần giải liên tiếp ở bảng sau:

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

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

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

Câu 28: Thời gian (phút) truy cập Internet mỗi buổi tối của một số học sinh được cho trong bảng sau:

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

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

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

Câu 29: Khối lượng các túi đường được đóng gói (đơn vị là kg) được thống kê ở bảng sau.

Khối lượng (kg)	$[1,5;1,7)$	$[1,7;1,9)$	$[1,9;2,1)$	$[2,1;2,3)$	$[2,3;2,5)$
Số túi đường	3	5	23	5	4

Phương sai của mẫu số liệu ghép nhóm trên gần nhất với kết quả nào sau đây.

A. $0,04$. **B.** $0,07$. **C.** $0,08$. **D.** 0.09.

Câu 30: Cân nặng của các học sinh lớp $10A$ trường Trung học phổ thông M như sau.

Cân nặng (kg)	$[30;36)$	$[36;42)$	$[42;48)$	$[48;54)$	$[54;60)$	$[60;66)$
Số học sinh lớp $10A$	1	2	5	15	9	6

Phương sai của mẫu số liệu ghép nhóm trên gần nhất với kết quả nào sau đây.

A. $50,4$. **B.** $7,1$. **C.** $51,8$. **D.** $7,2$.

Câu 31: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khoẻ. 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 32: 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. 33,09. **B.** 5,59. **C.** 5,75. **D.** 31,25.

Câu 33: 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

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

A. $5,98$. **B.** 6,23. **C.** 2,44. **D.** 2,50.

Câu 34: Cho bảng phân bố tần số ghép lớp về độ dài của 60 lá dương xỉ trưởng thành như sau sau

Độ dài (cm)	$[10;20)$	$[20;30)$	$[30;40)$	$[40;50]$
Tần số	8	18	24	10

Tính độ lệch chuẩn bảng phân bố tần số ghép lớp đã cho

A. $9,1$. **B.** $9,2$. **C.** $9,3$. **D.** 84.

Câu 35: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một của 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

Độ lệch chuẩn của mẫu số liệu trên gần nhất với giá trị nào dưới đây?

A. $2,02$. **B.** $2,03$. **C.** $4,04$. **D.** $2,01$.

Câu 36: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khoẻ. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

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

Độ lệch chuẩn của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

A. 3,41. **B.** 11,62. **C.** 0,017. **D.** 0,36.

Câu 37: 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

Độ lệch chuẩn của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

A. $5,59$. **B.** 5,95. **C.** 5,75. **D.** 31,25.

Câu 38: 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.** 6,23. **C.** 2,44. **D.** 2,50.

HƯỚNG DẪN GIẢI CHI TIẾT

Câu 1: Cho mẫu số liệu ghép nhóm về thống kê chiều cao của 35 cây bạch đàn trong rừng, ta có bảng số liệu sau:

Khoảng chiều cao (m)	$[6,5;\ 7,0)$	$[7,0;\ 7,5)$	$[7,5;\ 8,0)$	$[8,0;\ 8,5)$
Số cây	6	15	11	3

Tính chiều cao trung bình của 35 cây bạch đàn trên.

A. $7,407(m)$. **B.** $4,707(m)$. **C.** $7,704(m)$. **D.** $7,5(m)$.

Lời giải

Ta có giá trị đại diện các nhóm được cho dưới bảng sau:

Khoảng chiều cao	$[6,5;7,0)$	$[7,0;7,5)$	$[7,5;8,0)$	$[8,0;8,5)$
Giá trị đại diện	6,75	7,25	7,75	8,25
Tần số	6	15	11	3

Từ đó suy ra chiều cao trung bình của 35 cây bạch đàn là:

$$\overline{x}=\frac{6,75.6+7,25.15+7,75.11+8,25.3}{35}=7,047$$

Câu 2: Tìm hiểu thời gian làm bài tập về nhà trong tuần trước của một số học sinh thu được kết quả như sau:

Thời gian	$[18;20)$	$[20;22)$	$[22;24)$
Số học sinh	16	35	8

Tính thời gian làm bài tập về nhà trung bình trong tuần trước của các bạn học sinh này.

A. 20,73 giờ. **B.** 21,73 giờ. **C.** 20,72 giờ. **D.** 21,72 giờ.

Lời giải

Trong mỗi khoảng thời gian, giá trị đại diện là trung bình cộng của giá trị hai đầu mút nên ta có bảng sau:

Thời gian	19	21	23
Số học sinh	16	35	8

Thời gian làm bài tập về nhà trung bình trong tuần trước của các bạn học sinh này là:

$$\frac{19.16+21.35+23.8}{16+35+8}=\frac{1223}{59}\approx 20,73.$$

Câu 3: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một của hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

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

Số trung bình của mẫu số liệu trên thuộc khoảng nào trong các khoảng dưới đây?

A. $[7;\ 9)$. **B.** $[9;\ 11)$. **C.** $[11;\ 13)$. **D.** $[13;\ 15)$.

Lời giải

Bảng tần số ghép nhóm theo giá trị đại diện là

Doanh thu	$[5;7)$	$[7;9)$	$[9;11)$	$[11;13)$	$[13;15)$
Giá trị đại diện	6	8	10	12	14
Số ngày	2	7	7	3	1

Số trung bình: $\overline{x} = \dfrac{2.6 + 7.8 + 7.10 + 3.12 + 1.14}{20} = 9,4$

Câu 4: Thống kê nhiệt độ tại một địa điểm trong 40 ngày, ta có bảng số liệu sau:

Nhiệt độ $(^\circ C)$	$[19;22)$	$[22;25)$	$[25;28)$	$[28;31)$
Số ngày	7	15	12	6

Nhiệt độ trung bình là?

A. $25,456$. **B.** $24,775$. **C.** $23,021$. **D.** $22,036$.

Lời giải

Ta có bảng giá trị đại diện là:

Nhiệt độ $(^\circ C)$	$[19;22)$	$[22;25)$	$[25;28)$	$[28;31)$
Số ngày	7	15	12	6
Giá trị đại diện	20,5	23,5	26,5	29,5

Nhiệt độ trung bình là: $\dfrac{7.20,5 + 15.23,5 + 12.26,5 + 6.29,5}{40} = 24,775$

Câu 5: Thống kê cân nặng của học sinh lớp 11D cho trong bảng dưới đây:

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

Tính cân nặng trung bình của học sinh lớp 11D?

A. $50,1$. **B.** $52,83$. **C.** $50,81$. **D.** $51,81$.

Lời giải

Trong mỗi khoảng cân nặng, giá trị đại diện là trung bình cộng của giá trị hai đầu mút nên ta có bảng sau:

Cân nặng (kg)	43	48	53	58	63	68
Số học sinh	10	7	16	4	2	3

Tổng số học sinh là $n = 42.$

Cân nặng trung bình của học sinh lớp 11D là:

$\overline{x} = \dfrac{10.43 + 7.48 + 16.53 + 4.58 + 2.63 + 3.68}{42} \approx 51,81$ kg.

Câu 6: Khảo sát thời gian tập thể dục 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

Mốt của mẫu số liệu trên là

A. 42. **B.** 52. **C.** 53. **D.** 54.

Lời giải

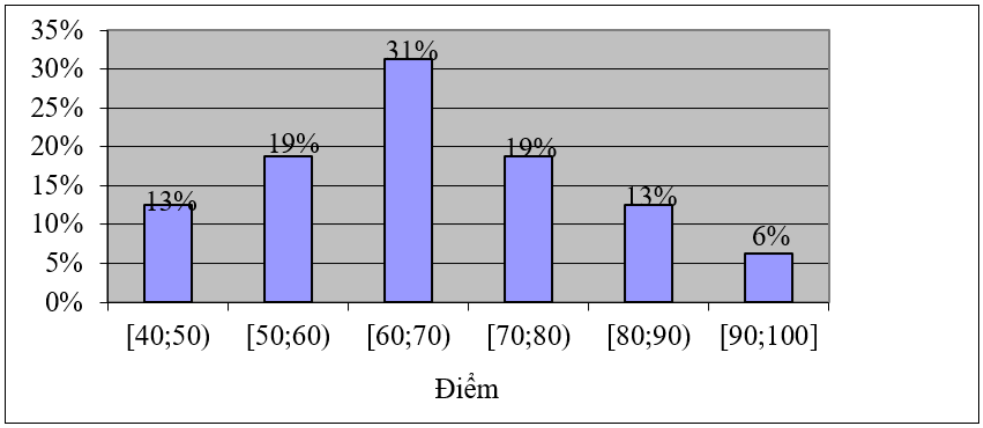
Mốt M_0 chứa trong nhóm $[40;60)$

Do đó: $u_m = 40; u_{m+1} = 60 \Rightarrow u_{m+1} - u_m = 60 - 40 = 20$

$n_{m-1} = 9; n_m = 12; n_{m+1} = 10$

$$M_0 = 40 + \frac{12-9}{(12-9) + (12-10)}(60-20) = 52$$

Câu 7: Kết quả điểm thi của 32 học sinh trong kì thi Tiếng Anh (thang điểm 100) được biểu diễn ở biểu đồ dưới đây:



Hãy ước lượng mốt của mẫu số liệu trên?

A. 55. **B.** 65. **C.** 72. **D.** 68.

Lời giải

Ta có bảng tần số ghép nhóm sau:

Lớp điểm	$[40;50)$	$[50;60)$	$[60;70)$	$[70;80)$	$[80;90)$	$[90;100]$
Tần số	4	6	10	6	4	2

Nhóm chứa mốt của mẫu số liệu là: $[60;70)$

Mốt của mẫu số liệu là: $60 + \frac{10-6}{10-6+10-6}.10 = 65$

Câu 8: Khảo sát thời gian tự học bài ở nhà của một số em học sinh lớp 11 thu được mẫu ghép nhóm số lượng như sau:

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

Mốt của số liệu ghép nhóm trên là

A. $102,86$. **B.** $102,85$. **C.** $102,8$. **D.** $102,9$.

Lời giải

$$M_0 = 90 + \frac{15-9}{15-9+15-7}.30 \approx 102,86$$

.

Câu 9: Thời gian 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	2	10	6	4	3

Tìm mốt của mẫu số liệu ghép nhóm này.

A. $17,42$. **B.** $14,56$. **C.** $17,16$. **D.** $12,67$.

Lời giải

Tần số lớn nhất là 10 nên nhóm chứa mốt là nhóm $[10,5;\ 20,5]$. Ta có:

$u_m = 10,5, u_{m+1} = 20,5, n_m = 10, n_{m+1} = 6; n_{m-1} = 2, u_{m+1} - u_m = 10.$

Do đó: $M_o = 10.5 + \dfrac{10-2}{(10-2)+(10-6)} \times 10 = 17,16$

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

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

Tìm mốt của mẫu số liệu ghép nhóm này.

A. $18,37$. **B.** $19,37$. **C.** $19,73$. **D.** $18,73$.

Lời giải

Tần số lớn nhất là 24 nên nhóm chứa mốt là nhóm $[18,5;\ 21,5]$. Ta có,

$u_m = 18,5, u_{m+1} = 21,5, n_m = 24, n_{m+1} = 15; n_{m-1} = 2, u_{m+1} - u_m = 3.$

Do đó: $M_o = 18,5 + \dfrac{24-15}{(24-15)+(24-2)} \times 3 = 18,5 + \dfrac{27}{31} \approx 19,37$

Câu 11: Thời gian đọc sách mỗi ngày của một số học sinh được cho trong bảng sau

Thời gian (phút)	[10;15)	[15;20)	[20;25)	[25;30)	[30;35)
Số học sinh	3	10	12	15	20

Xác định trung vị của mẫu số liệu ghép nhóm.

A. 25. **B.** 26. **C.** $25,56$. **D.** $26,67$.

Lời giải

Cỡ mẫu là $n = 3+10+12+15+20 = 60$.

Ta có: $\dfrac{n}{2} = 30$.

Mà $cf_3 = 25 < 30 < cf_4 = 40$. Suy ra nhóm 4 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 30.

Xét nhóm 4 là nhóm $[25;30)$ có $r = 25; d = 5$, $n_4 = 15$; và nhóm 3 có $cf_3 = 25$.

Vậy trung vị của mẫu số liệu là:

$M_e = 25 + \dfrac{30-25}{15}.5 \approx 26,67$.

Câu 12: Cho mẫu số liệu ghép nhóm về chiều cao của 25 cây dừa giống như sau:

Chiều cao (cm)	$[0;10)$	$[10;20)$	$[20;30)$	$[30;40)$	$[40;50)$
Số cây	4	6	7	5	3

Trung vị của mẫu số liệu ghép nhóm này là

A. $M_e=\frac{175}{7}$. **B.** $M_e=\frac{165}{5}$. **C.** $M_e=\frac{165}{7}$. **D.** $M_e=\frac{165}{3}$.

Lời giải

Cỡ mẫu: $n=4+6+7+5+3=25$.

Giả sử $x_1,x_2,\ldots,x_{25}$ là chiều cao của 25 cây dừa giống được sắp xếp theo thứ tự không giảm. Khi đó, trung vị là x_{13}. Do x_{13} thuộc nhóm $[20;30)$ nên nhóm này chứa trung vị. Do đó:

$p=3$, $a_3=20$, $m_3=7$, $m_1+m_2=10$, $a_4-a_3=10$. Do đó:

$$M_e=20+\frac{\frac{25}{2}-10}{7}.10=\frac{165}{7}.$$

Câu 13: Cho mẫu số liệu ghép nhóm về thống kê điểm số của 100 học sinh tham dự kỳ thi học sinh giỏi toán, ta có bảng số liệu sau:

Điểm	$[8;10)$	$[10;12)$	$[12;14)$	$[14;16)$	$[16;18)$	$[18;20)$
Số học sinh	6	21	30	25	14	4

Tìm trung vị của mẫu số liệu ghép nhóm trên.

A. $12,18$. **B.** $12,81$. **C.** $13,35$. **D.** $13,53$.

Lời giải

Cỡ mẫu là $n=6+21+30+25+14+4=100$.

Gọi $x_1;x_2;\ldots;x_{100}$ là số điểm của 100 học sinh và được sắp xếp theo thứ tự tăng dần. Khi đó, trung vị là $\frac{x_{50}+x_{51}}{2}$, do hai giá trị x_{50},x_{51} thuộc nhóm $[12;14)$ nên nhóm này chứa trung vị.

Ta có $p=3;a_3=12;m_3=30;m_1+m_2=25;a_4-a_3=2$ khi đó trung vị của mẫu số liệu là

$$M_e=12+\frac{\frac{100}{2}-27}{30}.2=13,53$$

Câu 14: Cho mẫu số liệu ghép nhóm về thống kê huyết áp của 20 người, ta có bảng số liệu sau:

Huyết áp (mmHg)	$[70;80)$	$[80;90)$	$[90;100)$	$[100;110)$	$[110;120)$	$[120;130)$
Số người	4	2	3	6	3	2

Tìm trung vị của mẫu số liệu ghép nhóm trên.

A. $100,67$. **B.** $101,67$. **C.** $101,76$. **D.** $100,76$.

Lời giải

Cỡ mẫu là $n=4+2+3+6+3+2=20$.

Gọi $x_1;x_2;....;x_{20}$ là giá trị huyết áp của 20 người và được sắp xếp theo thứ tự tăng dần. Khi đó, trung vị là $\frac{x_{10}+x_{11}}{2}$, do hai giá trị x_{10},x_{11} thuộc nhóm $[100;110)$ nên nhóm này chứa trung vị.

Ta có $p=4;a_4=100;m_4=6;m_1+m_2+m_3=9;a_5-a_4=10$ khi đó trung vị của mẫu số liệu là $M_e=100+\frac{\frac{20}{2}-9}{6}.10=101,67$

Câu 15: Khảo sát thời gian tập thể dục của một số học sinh khối 11 thu được mẫu số liệu ghép nhóm sau:

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

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

A. $[40;60)$. **B.** $[20;40)$. **C.** $[60;80)$. **D.** $[80;100)$.

Lời giải

Ta có: $n=42$

Nên tứ phân vị thứ nhất của mẫu số liệu trên là $Q_1=x_{11}$

Mà $x_{11}\in[20;40)$

Vậy nhóm chứa tứ phân vị thứ nhất của mẫu số liệu trên là nhóm $[20;40)$

Câu 16: Khảo sát thời gian tập thể dục của một số học sinh khối 11 thu được mẫu số liệu ghép nhóm sau:

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

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

A. $[40;60)$. **B.** $[20;40)$. **C.** $[60;80)$. **D.** $[80;100)$.

Lời giải

Ta có: $n=42$

Nên tứ phân vị thứ nhất của mẫu số liệu trên là $Q_3=x_{33}$

Mà $x_{33}\in[60;80)$

Vậy nhóm chứa tứ phân vị thứ nhất của mẫu số liệu trên là nhóm $[60;80)$

Câu 17: Tìm hiểu thời gian hoàn thành một bài tập của một số học sinh thu được kết quả sau:

Thời gian (phút)	[0;4)	[4;8)	[8;12)	[12;16)	[16;20)
Số học sinh	2	4	7	4	3

Tứ phân vị thứ ba của mẫu số liệu ghép nhóm này là

A. $Q_3=13$. **B.** $Q_3=14$. **C.** $Q_3=15$. **D.** $Q_3=12$.

Lời giải

Cỡ mẫu: $n=2+4+7+4+3=20$.

Tứ phân vị thứ ba Q_3 là $\frac{x_{15}+x_{16}}{2}$. Do x_{15}, x_{16} đều thuộc nhóm $[12;16)$ nên nhóm này chứa Q_3.

Do đó: $p=4$, $a_4=12$, $m_4=4$, $m_1+m_2+m_3=2+4+7=13$, $a_5-a_4=4$. Ta có:

$$Q_3=12+\frac{\frac{3.20}{4}-13}{4}.4=14.$$

Câu 18: Cho mẫu số liệu ghép nhóm về chiều cao của 25 cây dừa giống như sau:

Chiều cao (cm)	$[0;10)$	$[10;20)$	$[20;30)$	$[30;40)$	$[40;50)$
Số cây	4	6	7	5	3

Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm này là

A. $Q_1=13,5$. **B.** $Q_1=13,9$. **C.** $Q_1=15,75$. **D.** $Q_1=13,75$.

Lời giải

Cỡ mẫu: $n=4+6+7+5+3=25$.

Tứ phân vị thứ nhất Q_1 là $\frac{x_6+x_7}{2}$. Do x_6, x_7 đều thuộc nhóm $[10;20)$ nên nhóm này chứa Q_1.

Do đó: $p=2$, $a_2=10$, $m_2=6$, $m_1=4$, $a_3-a_2=10$. Ta có:

$$Q_1=10+\frac{\frac{25}{4}-4}{6}.10=13,75.$$

Câu 19: Tìm hiểu thời gian hoàn thành một bài tập của một số học sinh thu được kết quả sau:

Thời gian (phút)	$[0;4)$	$[4;8)$	$[8;12)$	$[12;16)$	$[16;20)$
Số học sinh	2	4	7	4	3

Tứ phân vị thứ ba của mẫu số liệu ghép nhóm này là

A. $Q_3=13$. **B.** $Q_3=14$. **C.** $Q_3=15$. **D.** $Q_3=12$.

Lời giải

Cỡ mẫu: $n=2+4+7+4+3=20$.

Tứ phân vị thứ ba Q_3 là $\frac{x_{15}+x_{16}}{2}$. Do x_{15}, x_{16} đều thuộc nhóm $[12;16)$ nên nhóm này chứa Q_3.

Do đó: $p=4$, $a_4=12$, $m_4=4$, $m_1+m_2+m_3=2+4+7=13$, $a_5-a_4=4$. Ta có:

$$Q_3 = 12 + \frac{\frac{3.20}{4} - 13}{4}.4 = 14$$.

Câu 20: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một của hàng được ghi lại ở bảng sau:

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

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

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

Lời giải

Gọi $x_1, x_2, ..., x_{20}$ là doanh thu bán hàng trong 20 ngày xếp theo thứ tự không giảm.

Khi đó: $x_1, x_2 \in [5;7)$, $x_3, ..., x_9 \in [7;9)$,, $x_9, ..., x_{16} \in [9;11)$ $x_{17}, ..., x_{19} \in [11;13)$, $x_{20} \in [13;15)$

Do đó, tứ phân vị thứ nhất của mẫu số liệu thuộc nhóm $[7;9)$

$n = 20, n_m = 7, C = 2, u_m = 7, u_{m+1} = 9$

$$Q_1 = 7 + \frac{\frac{1.20}{4} - 2}{7}(9-7) \approx 7,86 \approx 8$$

Câu 21: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một của hàng được ghi lại ở bảng sau:

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

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

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

Lời giải

Gọi $x_1, x_2, ..., x_{20}$ là doanh thu bán hàng trong 20 ngày xếp theo thứ tự không giảm.

Khi đó: $x_1, x_2 \in [5;7)$, $x_3, ..., x_9 \in [7;9)$,, $x_9, ..., x_{16} \in [9;11)$ $x_{17}, ..., x_{19} \in [11;13)$, $x_{20} \in [13;15)$

Do đó, tứ phân vị thứ ba của mẫu số liệu thuộc nhóm $[9;11)$

$n = 20, n_m = 7, C = 9, u_m = 9, u_{m+1} = 11$

$$Q_3 = 9 + \frac{\frac{3.20}{4} - 9}{7}(11-9) \approx 10;71 \approx 11$$

Câu 22: Tìm hiểu thời gian hoàn thành một bài tập của một số học sinh thu được kết quả sau:

Thời gian (phút)	$[0;4)$	$[4;8)$	$[8;12)$	$[12;16)$	$[16;20)$
Số học sinh	2	4	7	4	3

Tứ phân vị thứ ba của mẫu số liệu ghép nhóm này là

A. $Q_3=13$. **B.** $Q_3=14$. **C.** $Q_3=15$. **D.** $Q_3=12$.

Lời giải

Cỡ mẫu: $n=2+4+7+4+3=20$.

Tứ phân vị thứ ba Q_3 là $\frac{x_{15}+x_{16}}{2}$. Do x_{15}, x_{16} đều thuộc nhóm $[12;16)$ nên nhóm này chứa Q_3. Do đó: $p=4$, $a_4=12$, $m_4=4$, $m_1+m_2+m_3=2+4+7=13$, $a_5-a_4=4$.

Ta có: $Q_3=12+\frac{\frac{3.20}{4}-13}{4}.4=14$.

Câu 23: Cô Hà thống kê lại đường kính thân gỗ của một số cây xoan đào 6 năm tuổi được trồng ở một lâm trường ở bảng sau.

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

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

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

Lời giải

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

Câu 24: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khoẻ. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

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

Khoảng biến thiên của mẫu số liệu ghép nhóm là

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

Lời giải

Khoảng biến thiên của mẫu số liệu ghép nhóm là: $4,2-2,7=1,5(km)$

Câu 25: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khoẻ. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

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

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

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

Lời giải

Cỡ mẫu

$n=20$

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

Ta có: $x_1;\ldots;x_3\in[2,7;3,0);x_4;\ldots;x_9\in[3,0;3,3);x_{10};\ldots;x_{14}\in[3,3;3,6);$ $;x_{15};\ldots;x_{18}\in[3,6;3,9);x_{19};x_{20}\in[3,9;4,2).$

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

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

$$Q_1=3,0+\frac{\frac{20}{4}-3}{6}(3,3-3,0)=3,1$$

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

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

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

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

$\Delta_Q=Q_3-Q_1=0,575$

Câu 26: Bạn Chi rất thích nhảy hiện đại. Thời gian tập nhảy mỗi ngày trong thời gian gần đây của bạn Chi được thống kê lại ở bảng sau:

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

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

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

Lời giải

Cỡ mẫu $n=18$

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

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

Tứ phân vị thứ nhất của mẫu số liệu gốc là $x_5\in[20;25)$. Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là: $Q_1=20+\frac{\frac{18}{4}}{6}(25-20)=23,75$

Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{14}\in[30;35)$. Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là: $Q_3=30+\frac{\frac{3.18}{4}-(6+6)}{4}(35-30)=31,875$

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

Câu 27: Dũng là học sinh rất giỏi chơi rubik, bạn có thể giải nhiều loại khối rubik khác nhau. Trong một lần tập luyện giải khối rubik 3×3, bạn Dũng đã tự thống kê lại thời gian giải rubik trong 25 lần giải liên tiếp ở bảng sau:

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

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

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

Lời giải

Cỡ mẫu $n=25$

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

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

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

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_6+x_7)\in[10;12)$. Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là: $Q_1=10+\frac{\frac{25}{4}-4}{6}(12-10)=10,75$

Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{19}\in[14;16)$. Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là: $Q_3=14+\frac{\frac{3.25}{4}-(4+6+8)}{4}(16-14)=14,375$

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

Câu 28: Thời gian (phút) truy cập Internet mỗi buổi tối của một số học sinh được cho trong bảng sau:

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

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là

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

Lời giải

Cỡ mẫu là $n=56$.

Tứ phân vị thứ nhất Q_1 là $\frac{x_{14}+x_{15}}{2}$. Do x_{14},x_{15} đều thuộc nhóm $[12,5;15,5)$ nên nhóm này chứa Q_1. Do đó, $p=2;a_2=12,5;m_2=12;m_1=3,a_3-a_2=3$ và ta có

$$Q_1=12,5+\frac{\frac{56}{4}-3}{12}\cdot 3=15,25$$

Với tứ phân vị thứ ba Q_3 là $\frac{x_{42}+x_{43}}{2}$. Do x_{42},x_{43} đều thuộc nhóm $[18,5;21,5)$ nên nhóm này chứa Q_3. Do đó, $p=4;a_4=18,5;m_4=24;m_1+m_2+m_3=3+12+15=30;a_5-a_4=3$ và ta có $Q_3=18,5+\frac{\frac{3.56}{4}-30}{24}\cdot 3=20.$

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

Câu 29: Khối lượng các túi đường được đóng gói (đơn vị là kg) được thống kê ở bảng sau.

Khối lượng (kg)	$[1,5;1,7)$	$[1,7;1,9)$	$[1,9;2,1)$	$[2,1;2,3)$	$[2,3;2,5)$
Số túi đường	3	5	23	5	4

Phương sai của mẫu số liệu ghép nhóm trên gần nhất với kết quả nào sau đây.

A. $0,04$. **B.** $0,07$. **C.** $0,08$. **D.** 0.09.

Lời giải

Khối lượng trung bình của 40 túi đường là.

$$\bar{x}=\frac{1}{40}(3.1,6+5.1,8+23.2+5.2,2+4.2,4)\approx 2\ kg$$

Phương sai của mẫu số liệu ghép nhóm trên là.

$$s^2=\frac{1}{40}\left(3.(1,6-2)^2+5.(1,8-2)^2+23.(2-2)^2+5.(2,2-2)^2+4.(2,4-2)^2\right)\approx 0,04$$

Câu 30: Cân nặng của các học sinh lớp $10A$ trường Trung học phổ thông M như sau.

Cân nặng (kg)	$[30;36)$	$[36;42)$	$[42;48)$	$[48;54)$	$[54;60)$	$[60;66)$
Số học sinh lớp $10A$	1	2	5	15	9	6

Phương sai của mẫu số liệu ghép nhóm trên gần nhất với kết quả nào sau đây.

A. $50,4$. **B.** $7,1$. **C.** $51,8$. **D.** $7,2$.

Lời giải

Cân nặng trung bình của học sinh lớp $10A$ là.

$$\overline{x_A}=\frac{1}{38}(1.33+2.39+5.45+15.51+9.57+6.63)=52,4\ kg$$

Độ lệch chuẩn về nhóm cân nặng của học sinh lớp $10A$

$$s^2_A=\frac{1}{38}\left[\begin{array}{l}1.(33-52,4)^2+2.(39-52,4)^2+5.(45-52,4)^2+15.(51-52,4)^2\\+9.(57-52,4)^2+6.(63-52,4)^2\end{array}\right]$$

$\approx 50,4$

Câu 31: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khoẻ. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

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

Phương sai của mẫu số liệu ghép nhóm là

A. 3,39. **B.** 11,62. **C.** 0,1314. **D.** 0,36.

Lời giải

Giá trị đại	2,85	3,15	3,45	3,75	4,05

diện					
Số ngày	3	6	5	4	2

Số trung bình:

$$\overline{x}=\frac{3.2,85+6.3,15+5.3,45+4.3,75+2.4,05}{20}=3,39$$

Phương sai:

$$S^2=\frac{3.2,85^2+6.3,15^2+5.3,45^2+4.3,75^2+2.4,05^2}{20}-3,39^2=0,1314$$

Câu 32: 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. 33,09. **B.** 5,59. **C.** 5,75. **D.** 31,25.

Lời giải

Số trung bình: $\overline{x}=\frac{6.22,5+6.27,5+4.32,5+37,5+42,5}{18}\approx 28,33$

Phương sai: $S^2=\frac{6.22,5^2+6.27,5^2+4.32,5^2+37,5^2+42,5^2}{18}-28,33^2=31,25$

Câu 33: 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

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

A. 5,98. **B.** 6,23. **C.** 2,44. **D.** 2,50.

Lời giải

Giá trị đại diện	9	11	13	15	17
Số lần	4	_6	8	4	3

Số trung bình: $\overline{x}=\frac{4.9+6.11+8.13+4.15+3.17}{25}=12,68$

Phương sai:

$$s^2=\frac{4.(9-12,68)^2+6.(11-12,68)^2+8.(13-12,68)^2+4.(15-12,68)^2+3.(17-12,68)^2}{25}\approx 5,98$$

Câu 34: Cho bảng phân bố tần số ghép lớp về độ dài của 60 lá dương xỉ trưởng thành như sau sau

Độ dài (cm)	[10;20)	[20;30)	[30;40)	[40;50]
Tần số	8	18	24	10

Tính độ lệch chuẩn bảng phân bố tần số ghép lớp đã cho

A. $9,1$. **B.** $9,2$. **C.** $9,3$. **D.** 84.

Lời giải

Độ dài (cm)	$[10;20)$	$[20;30)$	$[30;40)$	$[40;50]$
Giá trị đại diện	15	25	35	45
Tần số	8	18	24	10

Trước hết ta có $\overline{x}=\frac{15.8+25.18+35.24+45.10}{60}=31$.

Khi đó phương sai $s_x^2=\frac{8.(15-31)^2+18\cdot(25-31)^2+24.(35-31)^2+10.(45-31)^2}{60}=84$.

$s_x=\sqrt{s_x^2}=\sqrt{84}\approx 9,2$

Câu 35: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một của 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

Độ lệch chuẩn của mẫu số liệu trên gần nhất với giá trị nào dưới đây?

A. $2,02$. **B.** $2,03$. **C.** $4,04$. **D.** $2,01$.

Lời giải

Bảng tần số ghép nhóm theo giá trị đại diện là

Doanh thu	[5;7)	[7;9)	[9;11)	[11;13)	[13;15)
Giá trị đại diện	6	8	10	12	14
Số ngày	2	7	7	3	1

Số trung bình: $\overline{x}=\frac{2.6+7.8+7.10+3.12+1.14}{20}=9,4$.

Phương sai của mẫu số liệu ghép nhóm là:

$s^2=\frac{2(6-9,4)^2+7(8-9,4)^2+7(10-9,4)^2+3(12-9,4)^2+1.(14-9,4)^2}{20}\approx 4,04$

$s=\sqrt{s^2}=\sqrt{4,04}\approx 2,01$

Câu 36: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khoẻ. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

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

Độ lệch chuẩn của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

A. 3,41. **B.** 11,62. **C.** 0,017. **D.** 0,36.

Lời giải

Giá trị đại diện	2,85	3,15	3,45	3,75	4,05
Số ngày	3	6	5	4	2

Số trung bình:

$$\overline{x}=\frac{3.2,85+6.3,15+5.3,45+4.3,75+2.4,05}{20}=3,39$$

Phương sai:

$$S^2=\frac{3.2,85^2+6.3,15^2+5.3,45^2+4.3,75^2+2.4,05^2}{20}-3,39^2=0,1314$$

Độ lệch chuẩn:

$$\sigma=\sqrt{0,1314}\approx 0,36$$

Câu 37: 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

Độ lệch chuẩn của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

A. $5,59$. **B.** 5,95. **C.** 5,75. **D.** 31,25.

Lời giải

Chọn D

Số trung bình: $\overline{x}=\frac{6.22,5+6.27,5+4.32,5+37,5+42,5}{18}\approx 28,33$

$$s^2=\frac{6.22,5^2+6.27,5^2+4.32,5^2+37,5^2+42,5^2}{18}-28,33^2=31,25$$

$$s=\sqrt{s^2}=\sqrt{31,25}\approx 5,59$$

Câu 38: 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.** 6,23. **C.** 2,44. **D.** 2,50.

Lời giải

Giá trị đại diện	9	11	13	15	17
Số lần	4	_6	8	4	3

Số trung bình: $\overline{x}=\frac{4.9+6.11+8.13+4.15+3.17}{25}=12,68$

$$s^2=\frac{4.(9-12,68)^2+6.(11-12,68)^2+8.(13-12,68)^2+4.(15-12,68)^2+3.(17-12,68)^2}{25}\approx 5,98$$

$$s=\sqrt{s^2}\approx 2,44$$