

SỐ LIỆU GHEP NHÓM

A. TÓM TẮT LÝ THUYẾT

I. Số liệu ghép nhóm.

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

Bảng 1: Bảng tần số ghép nhóm

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

Chú ý:

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

❖ Một số quy tắc ghép nhóm của mẫu số liệu

Mỗi mẫu số liệu có thể được ghép nhóm theo nhiều cách khác nhau nhưng thường tuân theo một số quy tắc sau:

Sử dụng từ $k = 5$ đến $k = 20$ nhóm. Cỡ mẫu càng lớn thì cần càng nhiều nhóm số liệu. Các nhóm có cùng độ dài bằng L thoả mãn $R < k.L$, trong đó R là khoảng biến thiên, k là số nhóm.

Giá trị nhỏ nhất của mẫu thuộc vào nhóm $[u_1; u_2)$ và càng gần u_1 càng tốt. Giá trị lớn nhất của mẫu thuộc nhóm $[u_k; u_{k+1})$ và càng gần u_{k+1} càng tốt.

II. Số trung bình.

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 $\bar{x}$, được tính như sau:

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

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

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

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

III. Mốt

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

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

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

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

❖ Ý nghĩa của mốt của mẫu số liệu ghép nhóm

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

IV. Lưu ý

- Các đầu mút của các nhóm có thể không là giá trị của mẫu số liệu.

- Ta hay gặp các bảng số liệu ghép nhóm là số nguyên, chẳng hạn như bảng thống kê số lỗi chính tả trong bài kiểm tra giữa học kì 1 môn Ngữ Văn của học sinh khối 11 như sau:

Số lỗi	[1;2]	[3;4]	[5;6]	[7;8]	[9;10]
Số bài	122	75	14	5	2

Bảng số liệu này không có dạng như trên

Để thuận lợi cho việc tính các số đặc trưng cho bảng số liệu này, người ta hiệu chỉnh về dạng như bảng sau bằng cách thêm và bớt 0,5 đơn vị vào đầu mút bên phải và bên trái của mỗi nhóm số liệu như sau:

Số lỗi	[0,5;2,5]	[2,5;4,5]	[4,5;6,5]	[6,5;8,5]	[8,5;10,5]
Số bài	122	75	14	5	2

B. CÁC DẠNG TOÁN.

DẠNG 1:

Ví dụ 1. Tính giá trị đại diện và độ dài của mỗi nhóm trong mẫu số liệu ở bảng sau:

Khoảng tuổi	[20;30)	[30;40)	[40;50)	[50;60)	[60;70)
Số khách hàng nữ	3	?	?	?	?

Lời giải

Khoảng tuổi	[20;30)	[30;40)	[40;50)	[50;60)	[60;70)
Giá trị đại diện	25	35	45	55	65
Độ dài của nhóm	10	10	10	10	10

DẠNG 2:

Ví dụ 2. Cân nặng của 28 học sinh nam lớp 11 được cho như sau:

55,4 62,6 54,2 56,8 58,8 59,4 60,7 58 59,5 63,6 61,8 52,3 63,4 57,9
 49,7 45,1 56,2 63,2 46,1 49,6 59,1 55,3 55,8 45,5 46,8 54 49,2 52,6

Hãy chia mẫu dữ liệu trên thành 5 nhóm, lập bảng tần số ghép nhóm và xác định giá trị đại diện cho mỗi nhóm.

Lời giải

Khoảng biến thiên của mẫu số liệu trên là $R = 63,6 - 45,1 = 18,5$.

Độ dài mỗi nhóm $L > \frac{R}{k} = \frac{18,5}{5} = 3,7$.

Ta chọn $L = 4$ và chia dữ liệu thành các nhóm $[45;49), [49;53), [53;57), [57;61), [61;65)$.

Khi đó ta có bảng tần số ghép nhóm sau:

Cân nặng	[45;49)	[49;53)	[53;57)	[57;61)	[61;65)
Giá trị đại diện	47	51	55	59	63
Số học sinh	4	5	7	7	5

Ví dụ 3. Một cửa hàng đã thống kê số ba lô bán được mỗi ngày trong tháng 9 với kết quả cho như sau:

12 29 12 19 15 21 19 29 28 12 15 25 16 20 29
 21 12 24 14 10 12 10 23 27 28 18 16 10 20 21

Hãy chia mẫu số liệu trên thành 5 nhóm, lập bảng tần số ghép nhóm, hiệu chỉnh bảng tần số ghép nhóm và xác định giá trị đại diện cho mỗi nhóm.

Lời giải

Khoảng biến thiên của mẫu số liệu trên là $R = 29 - 10 = 19$.

Độ dài mỗi nhóm $L > \frac{R}{k} = \frac{19}{5} = 3,8$.

Ta chọn $L = 4$ và chia dữ liệu thành các nhóm :

$[10;14), [14;18), [18;22), [22;26), [26;30)$

Khi đó ta có bảng tần số ghép nhóm sau:

Số ba lô	[10;14)	[14;18)	[18;22)	[22;26)	[26;30)
Giá trị đại diện	12	16	20	24	28
Số ngày	8	5	8	3	6

DẠNG 3:

Ví dụ 4. Các bạn học sinh lớp 11A1 trả lời 40 câu hỏi trong một bài kiểm tra. Kết quả được thống kê ở bảng sau:

Số câu trả lời đúng	[16;21)	[21;26)	[26;31)	[31;36)	[36;41)
Số học sinh	4	6	8	18	4

a) Tính giá trị đại diện c_p $1 \leq i \leq 5$, của từng nhóm số liệu.

b) Tính $n_1c_1 + n_2c_2 + n_3c_3 + n_4c_4 + n_5c_5$

c) Tính $\bar{x} = \frac{n_1c_1 + n_2c_2 + n_3c_3 + n_4c_4 + n_5c_5}{40}$.

Lời giải

a) $c_1 = 18,5; c_2 = 23,5; c_3 = 28,5; c_4 = 33,5; c_5 = 38,5$

b) $n_1c_1 + n_2c_2 + n_3c_3 + n_4c_4 + n_5c_5 = 1200$

c) $\bar{x} = \frac{n_1c_1 + n_2c_2 + n_3c_3 + n_4c_4 + n_5c_5}{40} = \frac{1200}{40} = 30$

Ví dụ 5. Kết quả khảo sát cân nặng của 25 quả cam ở mỗi lô hàng A và B được cho ở bảng sau:

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

a) Hãy ước lượng cân nặng trung bình của mỗi quả cam ở lô hàng A và lô hàng B .

b) Nếu so sánh theo số trung bình thì cam ở lô hàng nào nặng hơn?

Lời giải

a) Ta có bảng thống kê số lượng cam theo giá trị đại diện:

Cân nặng (g)	[150;155)	[155;160)	[160;165)	[165;170)	[170;175)
Cân nặng đại diện(g)	152,5	157,5	162,5	167,5	172,5
Số cam ở lô hàng A	2	6	12	4	1
Số cam ở lô hàng B	1	3	7	10	4

Cân nặng trung bình của mỗi quả cam ở lô hàng A xấp xỉ bằng

$$(2.152,5 + 6.157,5 + 12.162,5 + 4.167,5 + 1.172,5) : 25 = 161,7 \text{ (g)}.$$

Cân nặng trung bình của mỗi quả cam ở lô hàng B xấp xỉ bằng

$$(1.152,5 + 3.157,5 + 7.162,5 + 10.167,5 + 4.172,5): 25 = 165,1(\text{ g}).$$

b) Nếu so sánh theo số trung bình thì cam ở lô hàng B nặng hơn cam ở lô hàng A .

Ví dụ 6. Cân nặng của 28 học sinh nam lớp 11 được cho như sau:

55,4 62,6 54,2 56,8 58,8 59,4 60,7 58 59,5 63,6 61,8 52,3 63,4 57,9
49,7 45,1 56,2 63,2 46,1 49,6 59,1 55,3 55,8 45,5 46,8 54 49,2 52,6

a) Hãy chia mẫu dữ liệu trên thành 5 nhóm, lập bảng tần số ghép nhóm và xác định giá trị đại diện cho mỗi nhóm.

b) Hãy ước lượng cân nặng trung bình của học sinh lớp 11

Lời giải

a) Khoảng biến thiên của mẫu số liệu trên là $R = 63,6 - 45,1 = 18,5$.

$$\text{Độ dài mỗi nhóm } L > \frac{R}{k} = \frac{18,5}{5} = 3,7.$$

Ta chọn $L = 4$ và chia dữ liệu thành các nhóm

$[45;49), [49;53), [53;57), [57;61), [61;65)$.

Khi đó ta có bảng tần số ghép nhóm sau:

Cân nặng	[45;49)	[49;53)	[53;57)	[57;61)	[61;65)
Giá trị đại diện	47	51	55	59	63
Số học sinh	4	5	7	7	5

b) Cân nặng trung bình của học sinh trong lớp 11 xấp xỉ là:

$$(47.4 + 51.5 + 55.7 + 59.7 + 63.5): 28 = 55,6(\text{ kg})$$

DANG 4:

Ví dụ 7. Một công ty xây dựng khảo sát khách hàng xem họ có nhu cầu mua nhà ở mức giá nào. Kết quả khảo sát được ghi lại ở bảng sau:

Mức giá (triệu đồng/ m^2)	[10;14)	[14;18)	[18;22)	[22;26)	[26;30)
Số khách hàng	54	78	120	45	12

a) Tìm một của mẫu số liệu ghép nhóm trên.

b) Công ty nên xây nhà ở mức giá nào để nhiều người có nhu cầu mua nhất?

Lời giải

a) Nhóm chứa mốt của mẫu số liệu trên là nhóm $[18; 22)$.

Do đó $u_m = 18, n_{m-1} = 78, n_m = 120, n_{m+1} = 45, u_{m+1} - u_m = 22 - 18 = 4$.

Mốt của mẫu số liệu ghép nhóm là

$$M_o = 18 + \frac{120 - 78}{(120 - 78) + (120 - 45)} \cdot 4 = \frac{758}{39} \approx 19,4$$

b) Dựa vào kết quả trên ta có thể dự đoán rằng nếu công ty xây nhà ở mức giá 19,4 triệu đồng/ m^2 thì sẽ có nhiều người có nhu cầu mua nhất.

Ví dụ 8. Hãy sử dụng dữ liệu ở đề tư vấn cho đại lý bảo hiểm xác định khách hàng nam và nữ ở tuổi nào hay mua bảo hiểm nhất.

Số khách hàng mua bảo hiểm ở từng độ tuổi được thống kê như sau:

Độ tuổi	[20;30)	[30;40)	[40;50)	[50;60)	[60;70)
Số khách hàng nam	4	6	10	7	3
Số khách hàng nữ	3	9	6	4	2

Lời giải

Nhóm chứa mốt của mẫu số liệu khách hàng nam là $[40; 50)$.

Do đó $u_m = 40, n_{m-1} = 6; n_{m+1} = 7; u_{m+1} - u_m = 50 - 40 = 10$

Mốt của mẫu số liệu nhóm khách hàng nam là:

$$M_o = 40 + \frac{10 - 6}{(10 - 6) + (10 - 7)} \cdot 10 = 45,7$$

Dựa vào kết quả trên ta có thể dự đoán được khách hàng nam 46 tuổi có nhu cầu mua bảo hiểm cao nhất

Nhóm chứa mốt của mẫu số liệu khách hàng nữ là $[30; 40)$

Do đó $u_m = 30, n_{m-1} = 3; n_{m+1} = 6; u_{m+1} - u_m = 40 - 30 = 10$

Mốt của mẫu số liệu nhóm khách hàng nam là:

$$M_o = 30 + \frac{9 - 3}{(9 - 3) + (9 - 6)} \cdot 10 = 36,7$$

Dựa vào kết quả trên ta có thể dự đoán được khách hàng nữ 37 tuổi có nhu cầu mua bảo hiểm cao nhất

DANG 5:

Ví dụ 9. Số cuộc gọi điện thoại một người thực hiện mỗi ngày trong 30 ngày được lựa chọn ngẫu nhiên được thống kê trong bảng sau:

Số cuộc gọi	[3;5]	[6;8]	[9;11]	[12;14]	[15;17]
Số ngày	5	13	7	3	2

a) Tìm một của mẫu số liệu ghép nhóm trên.

b) Hãy dự đoán xem khả năng người đó thực hiện bao nhiêu cuộc gọi mỗi ngày là cao nhất.

Lời giải

Hiệu chỉnh bảng lại như sau:

Số cuộc gọi	[2,5;5,5)	[5,5;8,5)	[8,5;11,5)	[11,5;14,5)	[14,5;17,5)
Số ngày	5	13	7	3	2

a) Nhóm chứa một của mẫu số liệu trên là nhóm $[5,5;8,5)$.

Do đó $u_m = 5,5; n_{m-1} = 5; n_m = 13; n_{m+1} = 7; u_{m+1} - u_m = 8,5 - 5,5 = 3$.

Mốt của mẫu số liệu ghép nhóm là

$$M_o = 5,5 + \frac{13-5}{(13-5)+(13-7)} \cdot 3 = \frac{101}{14} \approx 7,2$$

b) Dựa vào kết quả trên ta có thể dự đoán rằng khả năng người đó thực hiện 7 cuộc gọi mỗi ngày là cao nhất.

DẠNG 6:

Ví dụ 10. Anh Văn ghi lại cự li 30 lần ném lao của mình ở bảng sau (đơn vị: mét):

72,1	72,9	70,2	70,9	72,2	71,5	72,5	69,3	72,3	69,7
72,3	71,5	71,2	69,8	72,3	71,1	69,5	72,2	71,9	73,1
71,6	71,3	72,2	71,8	70,8	72,2	72,2	72,9	72,7	70,7

a) Tính cự li trung bình của mỗi lần ném.

b) Tổng hợp lại kết quả ném của anh Văn vào bảng tần số ghép nhóm theo mẫu sau:

Cự li (m)	[69,2;70)	[70;70,8)	[70,8;71,6)	[71,6;72,4)	[72,4;73,2)
Số lần	?	?	?	?	?

c) Hãy ước lượng cự li trung bình mỗi lần ném từ bảng tần số ghép nhóm trên.

d) Khả năng anh Văn ném được khoảng bao nhiêu mét là cao nhất?

Lời giải

a) Cự li trung bình của mỗi lần ném là 71,6 (m)

b) Bảng tần số ghép nhóm

Cự li (m)	[69,2;70)	[70;70,8)	[70,8;71,6)	[71,6;72,4)	[72,4;73,2)
Số lần	4	2	9	10	5

c)

Cự li (m)	[69,2;70)	[70;70,8)	[70,8;71,6)	[71,6;72,4)	[72,4;73,2)
Giá trị đại diện	69,6	70,4	71,2	72	72,8
Số lần	4	2	9	10	5

Cự li trung bình mỗi lần ném xấp xỉ bằng

$$\frac{69,6 \cdot 4 + 70,4 \cdot 2 + 71,2 \cdot 9 + 72 \cdot 10 + 72,8 \cdot 5}{30} = 71,5 \text{ (m)}$$

d) Nhóm chứa một của mẫu số liệu trên là $[71,6; 72,4)$

Do đó: $u_m = 71,6; n_{m-1} = 9; n_{m+1} = 5; u_{m+1} - u_m = 72,4 - 71,6 = 0,8$

Mốt của mẫu số liệu ghép nhóm là:

$$M_0 = 71,6 + \frac{10 - 9}{(10 - 9) + (10 - 5)} \cdot 0,8 = 71,7 \text{ (m)}$$

Vậy khả năng anh Văn ném được 71,7 m là cao nhất

C. BÀI TẬP TỰ LUYỆN.

Bài 1: Mẫu số liệu sau cho biết phân bố theo độ tuổi của dân số Việt Nam năm 2019.

Độ tuổi	Dưới 15	Từ 15 đến dưới 65 tuổi	Từ 65 tuổi trở lên
Số người	23371882	65420451	7416651

a) Mẫu số liệu đã cho có là mẫu số liệu ghép nhóm hay không?

b) Nêu các nhóm và tần số tương ứng. Dân số Việt Nam năm 2019 là bao nhiêu?

Bài 2: 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

Đọc và giải thích mẫu số liệu này.

Bài 3: Cân nặng (kg) của 35 người trưởng thành tại một khu dân cư được cho như sau:

43 51 47 62 48 40 50 62 53 56 40 48 56 53 50 42 55
52 48 46 45 54 52 50 47 44 54 55 60 63 58 55 60 58 53.

Chuyển mẫu số liệu trên thành dạng ghép nhóm, các nhóm có độ dài bằng nhau, trong đó có nhóm [40; 45).

Bài 4: 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(m)	[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.

a) Lập bảng tần số ghép nhóm của mẫu số liệu với các nhóm đã cho ở bảng trên.

b) Công ty may 500 áo đồng phục cho học sinh lớp 11 thì nên may số lượng áo theo mỗi cỡ là bao nhiêu chiếc?

Bài 5: Cho bảng tần số ghép nhóm số liệu thống kê chiều cao của 40 mẫu cây ở một vườn thực vật (đơn vị: centimét).

Chiều cao(m)	[30;40)	[40;50)	[50;60)	[60;70)	[70;80)	[80;90)
Số cây						

a) Xác định số trung bình cộng của mẫu số liệu ghép nhóm trên.

b) Mốt của mẫu số liệu trên là bao nhiêu?

Bài 6: Mẫu số liệu sau ghi lại cân nặng của 30 bạn học sinh (đơn vị: kilogam)

17 40 39 40,5 42 51 41,5 39 41 30
40 42 40,5 39,5 41 40,5 37 39,5 40 41
38,5 39,5 40 41 39 40,5 40 38,5 39,5 41,5

a) Lập bảng tần số ghép nhóm cho mẫu số liệu trên có tám nhóm ứng với tám nửa khoảng

$[15; 20), [20; 25), [25; 30), [30; 35), [35; 40), [40; 45), [45; 50), [50; 55)$.

b) Xác định số trung bình cộng của mẫu số liệu ghép nhóm trên.

c) Mốt của mẫu số liệu ghép nhóm trên là bao nhiêu

Bài 7: 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

a) Xác định giá trị trung bình của mẫu số liệu.

b) Tìm mốt của mẫu số liệu.

Bài 8: Người ta đếm số xe ô tô đi qua một trạm thu phí mỗi phút trong khoảng thời gian từ 9 giờ đến 9 giờ 30 phút sáng. Kết quả được ghi lại ở bảng sau:

15	16	13	21	17	23	15	21	6	11	12	23	19	25	11
25	7	29	10	28	29	24	6	11	23	11	21	9	27	15

a) Tính số xe trung bình đi qua trạm thu phí trong mỗi phút.

b) Tổng hợp lại số liệu trên vào bảng tần số ghép nhóm theo mẫu sau:

Số xe	[6;10]	[11;15]	[16;20]	[21;25]	[26;30]
Số lần	?	?	?	?	?

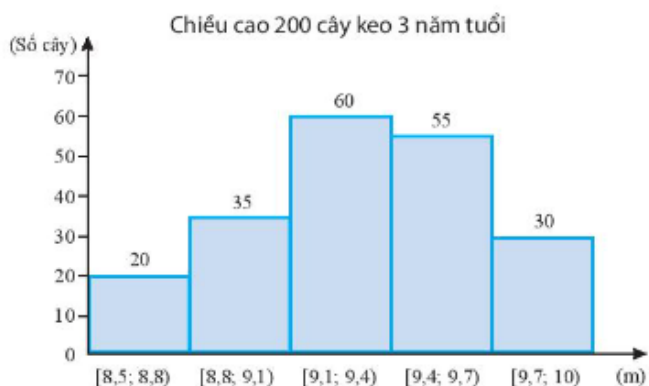
c) Hãy ước lượng trung bình số xe đi qua trạm thu phí trong mỗi phút từ bảng tần số ghép nhóm trên.

Bài 9: Một thư viện thống kê số lượng sách được mượn mỗi ngày trong ba tháng ở bảng sau:

Số sách	[16;20]	[21;25]	[26;30]	[31;35]	[36;40]	[41;45]	[46;50]
Số ngày	3	6	15	27	22	14	5

Hãy ước lượng số trung bình và mốt của mẫu số liệu ghép nhóm trên.

Bài 10: Kết quả đo chiều cao của 200 cây keo 3 năm tuổi ở một nông trường được biểu diễn ở biểu đồ dưới đây.



Hãy ước lượng số trung bình và mốt của mẫu số liệu ghép nhóm trên.

D. BÀI TẬP TRẮC NGHIỆM

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

(Học sinh dựa vào bảng trên trả lời câu 1, câu 2, câu 3, câu 4, câu 5)

Câu 1: Giá trị đại diện của nhóm [8;12) là

- A) 8. B) 12. C) 10. D) 11.

Câu 2: Mẫu số liệu trên có bao nhiêu nhóm

- A) 10. B) 11. C) 7. D) 5.

Câu 3: Có bao nhiêu ô tô có độ tuổi từ 12 đến dưới 16

- A) 23. B) 25. C) 37. D) 26.

Câu 4: Nhóm có tần số 19 là:

- A) [4;8). B) [8;12). C) [12;16). D) [16;20).

Câu 5: Có bao nhiêu ô tô có độ tuổi dưới 12

- A) 75. B) 37. C) 45. D) 26.

Khảo sát cân nặng của 30 bạn học sinh (đơn vị: kilogam), ta có bảng tần số ghép nhóm:

Cân nặng(m)	[15;20)	[20;25)	[25;30)	[30;35)	[35;40)	[40;45)	[45;50)	[50;55)
Số học sinh	1	0	0	1	10	17	0	1

(Học sinh dựa vào bảng trên trả lời câu 6, câu 7, câu 8)

Câu 6: Giá trị đại diện của nhóm [25;30) là

- A) 27,5. B) 25. C) 30. D) 27.

Câu 7: Xác định số trung bình cộng của mẫu số liệu ghép nhóm trên.

- A) 25. B) 25,8. C) 30. D) 27.

Câu 8: Mốt của mẫu số liệu ghép nhóm trên là bao nhiêu

- A) 40,5. B) 42,5. C) 41,5. D) 41,25.

Anh Ba ghi nhận lại kết quả ném lao của mình ở cự li 30 lần , và có bảng sau:

Cự li (m)	[69,2;70)	[70;70,8)	[70,8;71,6)	[71,6;72,4)	[72,4;73,2)
Số lần	4	2	9	10	5

(Học sinh dựa vào bảng trên trả lời câu 9, câu 10)

Câu 9: Cự li trung bình mỗi lần ném của anh Ba

- A) 73,5. B) 42,5. C) 41,5. D) 71,5.

Câu 10: Khả năng anh Ba ném được khoảng bao nhiêu mét là cao nhất?

- A) 71,7. B) 71,75. C) 71,8. D) 71,5.

E. HƯỚNG DẪN GIẢI



BÀI TẬP TỰ LUYỆN

Bài 1:

- a) Mẫu số liệu đã cho là mẫu số liệu ghép nhóm.
- b) Có ba nhóm là: Dưới 15 tuổi, Từ 15 đến dưới 65 tuổi, Từ 65 tuổi trở lên.
Có 23371882 người dưới 15 tuổi; 65420451 người từ 15 đến dưới 65 tuổi
và 7416651 người từ 65 tuổi trở lên.
- Dân số Việt Nam năm 2019 là $23371882 + 65420451 + 7416651 = 96208984$ người.

Bài 2: Mẫu số liệu được chia thành 7 nhóm

- Có 6 nhân viên có thời gian đi từ nhà đến nơi làm việc là từ 15 đến dưới 20 phút
- Có 14 nhân viên có thời gian đi từ nhà đến nơi làm việc là từ 20 đến dưới 25 phút

- Có 25 nhân viên có thời gian đi từ nhà đến nơi làm việc là từ 25 đến dưới 30 phút
- Có 37 nhân viên có thời gian đi từ nhà đến nơi làm việc là từ 30 đến dưới 35 phút
- Có 21 nhân viên có thời gian đi từ nhà đến nơi làm việc là từ 35 đến dưới 40 phút
- Có 13 nhân viên có thời gian đi từ nhà đến nơi làm việc là từ 40 đến dưới 45 phút
- Có 9 nhân viên có thời gian đi từ nhà đến nơi làm việc là từ 45 đến dưới 50 phút

Bài 3: Bảng tần số ghép nhóm của mẫu số liệu:

Cân nặng	[40;45)	[45;50)	[50;55)	[55;60)	[60;65)
Số người	5	7	11	7	5

Bài 4: a) Bảng tần số ghép nhóm của mẫu số liệu:

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

b) Tỷ lệ học sinh mặc vừa cỡ M là: $(22 : 36) \times 100 = 61.11\%$

Tỷ lệ học sinh mặc vừa cỡ L là: $(8 : 36) \times 100 = 22.22\%$

Tỷ lệ học sinh mặc vừa cỡ XL là: $(6 : 36) \times 100 = 16.67\%$

Số lượng áo cỡ M nên may là: $61.11\% \times 500 = 306$

Số lượng áo cỡ L nên may là: $22.22\% \times 500 = 111$

Số lượng áo cỡ XL nên may là: $16.67\% \times 500 = 83$

Bài 5:

a)

Chiều cao(m)	[30;40)	[40;50)	[50;60)	[60;70)	[70;80)	[80;90)
Giá trị đại diện	35	45	55	65	75	85
Số cây	4	10	14	6	4	2

Giá trị trung bình của mẫu số liệu là

$$\bar{x} = \frac{35.4 + 45.10 + 55.14 + 65.6 + 75.4 + 85.2}{40} \approx 63,43$$

b) Nhóm chứa mốt: $[50;60)$

Mốt của mẫu số liệu:

$$M_0 = 50 + \frac{4}{4+8} \cdot 10 \approx 53,3$$

Bài 6:

a) Bảng tần số ghép nhóm cho mẫu số liệu là:

Cân nặng(m)	[15;20)	[20;25)	[25;30)	[30;35)	[35;40)	[40;45)	[45;50)	[50;55)
Giá trị đại diện	17,5	22,5	27,5	32,5	37,5	42,5	47,5	52,5
Số học sinh	1	0	0	1	10	17	0	1

b) Số trung bình cộng: $\bar{x} = \frac{17,5 + 32,5 + 37,5 \cdot 10 + 42,5 \cdot 17 + 52,5}{30} = 25,8$

c) Mốt của mẫu số liệu

$$M_0 = 40 + \frac{7}{7+17} \cdot 5 \approx 41,5$$

Bài 7:

Hàm lượng chất béo (g)	[2;6)	[6;10)	[10;14)	[14;18)	[18;22)	[22;26)
Giá trị đại diện	4	8	12	16	20	24
Tần số	2	6	10	13	16	13

a) Giá trị trung bình: $\bar{x} = \frac{4 \cdot 2 + 8 \cdot 6 + 12 \cdot 10 + 16 \cdot 13 + 20 \cdot 16 + 24 \cdot 13}{60} = 16,9$

b) Mốt của mẫu số liệu: $M_0 = 18 + \frac{3}{3+3} \cdot 4 = 20$

Bài 8:

a) Số xe trung bình đi qua trạm thu phí trong mỗi phút là 17,4 xe

b) Bảng tần số ghép nhóm

Số xe	[6;10]	[11;15]	[16;20]	[21;25]	[26;30]
Số lần	5	9	3	9	4

c) Hiệu chỉnh lại bảng số liệu ta có:

Số xe	[5,5;10,5)	[10,5;15,5)	[15,5;20,5)	[20,5;25,5)	[25,5;30,5)
-------	------------	-------------	-------------	-------------	-------------

Giá trị đại diện	8	13	18	23	28
Số lần	5	9	3	9	4

Trung bình số xe đi qua trạm thu phí mỗi phút xấp xỉ bằng:

$$\frac{8.5 + 13.9 + 18.3 + 23.9 + 28.4}{30} = 17,7$$

Bài 9: Số liệu trên được hiệu chỉnh như sau:

Số sách	[15,5;20,5)	[20,5;25,5)	[25,5;30,5)	[30,5;35,5)	[35,5;40,5)	[40,5;45,5)	[45,5;50,5)
Giá trị đại diện	18	23	28	33	38	43	48
Số ngày	3	6	15	27	22	14	5

Số sách được mượn trung bình mỗi ngày xấp xỉ bằng:

$$(18.3 + 23.6 + 28.15 + 33.27 + 38.33 + 43.14 + 48.5) : 92 = 34,6$$

Nhóm chứa một của mẫu số liệu là $[30,5;35,5)$

Do đó $u_m = 30,5; n_{m-1} = 15, n_{m+1} = 22, u_{m+1} - u_m = 35,5 - 30,5 = 5$

Mốt của mẫu số liệu ghép nhóm là:

$$M_0 = 30,5 + \frac{27 - 15}{(27 - 15) + (27 - 22)} \cdot 5 = 34$$

Bài 10: Chiều cao của 200 cây keo được thống kê như bảng sau:

Chiều cao	[8,5;8,8)	[8,8;9,1)	[9,1;9,4)	[9,4;9,7)	[9,7;10)
Giá trị đại diện	8,65	8,95	9,25	9,55	9,85
Số cây	20	35	60	55	30

Chiều cao trung bình của 200 cây xấp xỉ bằng:

$$(8,65.20 + 8,95.35 + 9,25.60 + 9,55.55 + 9,85.30) : 200 = 9,31 \text{ (m)}$$

Nhóm chứa một của mẫu số liệu là $[9,1;9,4)$

Do đó: $u_m = 9,1; n_{m-1} = 35; n_{m+1} = 55; u_{m+1} - u_m = 9,4 - 9,1 = 0,3$

Mốt của mẫu số liệu trên là:

$$M_0 = 9,1 + \frac{60 - 35}{(60 - 35) + (60 - 55)} \cdot 0,3 = 9,35 \text{ (m)}$$



BÀI TẬP TRẮC NGHIỆM

1.C

2.D

3.D

4.D

5.A

6.A

7.B

8.C

9.D

10.A

Bài 2

TRUNG VỊ VÀ TỨ PHÂN VỊ CỦA MẪU SỐ LIỆU GHÉP NHÓM

A. TÓM TẮT LÝ THUYẾT

I. Trung vị.

- Trong trường hợp mẫu số liệu có giá trị bất thường (rất lớn hoặc rất bé so với đa số các giá trị khác), người ta không dùng số trung bình để đo xu thế trung tâm mà dùng trung vị.
- Trung vị là giá trị chia đôi mẫu số liệu, nghĩa là trong mẫu số liệu được sắp xếp theo thứ tự không giảm thì giá trị trung vị ở vị trí chính giữa.

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

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

Khi đó

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

Ý nghĩa của trung vị của mẫu số liệu ghép nhóm

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

II. Tứ phân vị.

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

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

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

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

Khi đó

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

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

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

Khi đó

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

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

- Ba điểm tứ phân vị chia mẫu số liệu đã sắp xếp theo thứ tự không giảm thành bốn phần đều nhau. Giống như với trung vị, nói chung không thể xác định chính xác các điểm tứ phân vị của mẫu số liệu ghép nhóm.
- Bộ ba tứ phân vị của mẫu số liệu ghép nhóm là giá trị xấp xỉ cho tứ phân vị của mẫu số liệu gốc và được sử dụng làm giá trị đo xu thế trung tâm của mẫu số liệu.
- Tứ phân vị thứ nhất và thứ ba đo xu thế trung tâm của nửa dưới (các dữ liệu nhỏ hơn Q_2) và nửa trên (các dữ liệu lớn hơn Q_2) của mẫu số liệu.

B. CÁC DẠNG TOÁN.

DẠNG 1:

XÁC ĐỊNH TRUNG VỊ CỦA MẪU SỐ LIỆU GHÉP NHÓM

Phương pháp:

+ Nắm vững công thức cơ bản

Ví dụ 1. Kết quả khảo sát cân nặng của 1 thùng táo ở một lô hàng cho trong bảng sau:

Cân nặng (g)	[150;155)	[155;160)	[160;165)	[165;170)	[170;175)
Số quả táo	4	7	12	6	2

Hãy tìm trung vị của mẫu số liệu ghép nhóm trên.

Lời giải

Gọi $x_1; x_2; \dots; x_{31}$ là cân nặng của 31 quả bơ xếp theo thứ tự không giảm.

Do $x_1; x_2; \dots; x_4 \in [150;155)$; $x_5; x_6; \dots; x_{11} \in [155;160)$ nên trung vị của mẫu số liệu $x_1; x_2; \dots; x_{31}$ là $x_{16} \in [160;165)$.

Ta xác định được $n = 31, n_m = 12, C = 4 + 7 = 11, u_m = 160, u_{m+1} = 165$.

Vậy trung vị của mẫu số liệu ghép nhóm là

$$M_e = 160 + \frac{\frac{31}{2} - 11}{12} \cdot (165 - 160) = 161,875.$$

Ví dụ 2. Trong tuần lễ bảo vệ môi trường, các học sinh khối 12 tiến hành thu nhặt vỏ lon nước ngọt để tái chế. Nhà trường thống kê kết quả thu nhặt vỏ lon nước ngọt của học sinh khối 12 ở bảng sau:

Số vỏ lon	[11;15]	[16;20]	[21;25]	[26;30]	[31;35]
Số học sinh	58	87	54	44	23

Hãy tìm trung vị của mẫu số liệu ghép nhóm trên.

Lời giải

Do số vỏ chai là số nguyên nên ta hiệu chỉnh lại như sau:

Số vỏ lon	[10,5;15,5)	[15,5;20,5)	[20,5;25,5)	[25,5;30,5)	[30,5;35,5)
Số học sinh	58	87	54	44	23

Số học sinh tham gia thu nhặt vỏ lon nước ngọt là

$$n = 58 + 87 + 54 + 44 + 23 = 266.$$

Gọi $x_1; x_2; \dots; x_{266}$ lần lượt là số vỏ chai 266 học sinh khối 12 thu nhặt được xếp theo thứ tự không giảm.

Do $x_1; x_2; \dots; x_{58} \in [10, 5; 15, 5); x_{59}; x_{59}; \dots; x_{145} \in [15, 5; 20, 5)$ nên trung vị của mẫu số liệu $x_1; x_2; \dots; x_{240}$ là $\frac{1}{2}(x_{133} + x_{134}) \in [15, 5; 20, 5)$.

Ta xác định được $n = 266, n_m = 87, C = 58, u_m = 15, 5, u_{m+1} = 20, 5$ và $u_{m+1} = 20, 5$.

Trung vị của mẫu số liệu ghép nhóm là

$$M_e = 15, 5 + \frac{\frac{266}{2} - 58}{87} \cdot (20, 5 - 15, 5) \approx 19, 81.$$

DẠNG 2:

XÁC ĐỊNH TỨ PHÂN VỊ CỦA MẪU SỐ LIỆU GHEP NHÓM

Phương pháp:

+ Nắm vững công thức cơ bản

Ví dụ 3. Tiền lương nhận được trong 1 giờ làm việc của nhân viên công ty A được thống kê theo mẫu số liệu ghép nhóm sau (đơn vị: ngàn đồng):

Các lớp tiền lương	Số nhân viên
[50; 60)	8
[60; 70)	10
[70; 80)	16
[80; 90)	14
[90; 100)	10
[100; 110)	5
[110; 120)	2

Hãy xác định các tứ phân vị của mẫu số liệu trên.

Lời giải

Gọi $x_1 \leq x_2 \leq \dots \leq x_{65}$ là tiền lương của 65 nhân viên nhận được trong 1 giờ.

Tứ phân vị thứ hai của dãy số liệu $x_1; x_2; \dots; x_{65}$ là $x_{33} \in [70; 80)$. Do đó tứ phân vị thứ hai của mẫu số liệu ghép nhóm là

$$Q_2 = 70 + \frac{\frac{2 \cdot 65}{4} - (8+10)}{16} \cdot (80 - 70) = 79,0625$$

Tứ phân vị thứ nhất của dãy số liệu $x_1; x_2; \dots; x_{65}$ là $x_{17} \in [60; 70)$. Do đó tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là

$$Q_1 = 60 + \frac{\frac{1 \cdot 65}{4} - 8}{10} \cdot (70 - 60) = 68,25$$

Tứ phân vị thứ ba của dãy số liệu $x_1; x_2; \dots; x_{65}$ là $x_{49} \in [90; 100)$. Do đó tứ phân vị thứ ba của mẫu số liệu ghép nhóm là

$$Q_3 = 90 + \frac{\frac{3 \cdot 65}{4} - (8+10+16+14)}{10} \cdot (100 - 90) = 90,75.$$

Chú ý: Nếu tứ phân vị thứ k là $\frac{1}{2}(x_m + x_{m+1})$, trong đó $x_m \in [u_{j-1}; u_j)$ và $x_{m+1} \in [u_j; u_{j+1})$ thuộc hai nhóm liên tiếp thì ta lấy $Q_k = u_j$.

Ví dụ 4. Mức lương hàng tháng ở 1 công ty được Công đoàn thu thập theo bảng sau(đơn vị triệu đồng):

Mức lương	[6;10)	[10;15)	[15;20)	[20;25)	[25;30)
Nhân viên	17	38	27	21	7

- Hãy ước lượng các tứ phân vị của mẫu số liệu ghép nhóm trên.
- Chủ tịch Công đoàn muốn đề nghị hỗ trợ cho nhóm 25% số nhân viên có mức lương thấp nhất và ước lượng rằng số nhân viên này không ít hơn 10. Nhận định của chủ tịch có hợp lí hay không?

Lời giải

a) Gọi $x_1 \leq x_2 \leq \dots \leq x_{110}$ là mức lương của 110 nhân viên nhận được công ty trả trong 1 tháng.

Tứ phân vị thứ hai của dãy số liệu $x_1 \leq x_2 \leq \dots \leq x_{110}$ là $\frac{1}{2}(x_{55} + x_{56})$. Do

$x_{55} \in [10;15)$ và $x_{56} \in [15;20)$. Nên đó tứ phân vị thứ hai của mẫu số liệu ghép nhóm là $Q_2 = 15$

Tứ phân vị thứ nhất của dãy số liệu $x_1; x_2; \dots; x_{110}$ là $\frac{1}{2}(x_{27} + x_{28})$. Do

$x_{27} \in [10; 15)$ và $x_{28} \in [10; 15)$ đó tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là

$$Q_1 = 10 + \frac{\frac{1 \cdot 110}{4} - 17}{38} \cdot (15 - 10) \approx 11,38$$

Tứ phân vị thứ ba của dãy số liệu $x_1; x_2; \dots; x_{110}$ là $\frac{1}{2}(x_{82} + x_{83})$. Do

$x_{82} \in [15; 20)$ và $x_{83} \in [20; 25)$ đó tứ phân vị thứ ba của mẫu số liệu ghép nhóm là $Q_3 = 20$.

b) Do tứ phân vị thứ nhất $\approx 11,38$ nên nhận định trên là không hợp lí.

C. BÀI TẬP TỰ LUYỆN.

Bài 1: Lương tháng của một số giáo viên THPT được ghi lại như sau (đơn vị: triệu đồng):

12,2	9,8	10,9	6,7	13,6	9,2	12,5	9,6	11,7	12,7	10,0	10,0
8,3	11,1	11,9	8,4	6,7	13,8	13,1	6,5	10,7	8,9	11,2	13,2

a) Tìm tứ phân vị của dãy số liệu trên.

b) Tổng hợp lại dãy số liệu trên vào bảng tần số ghép nhóm theo mẫu sau:

Lương tháng (triệu đồng)	[6;8)	[8;10)	[10;12)	[12;14)
Số nhân viên	?	?	?	?

c) Hãy ước lượng tứ phân vị của số liệu ở bảng tần số ghép nhóm trên.

Bài 2: Một học sinh làm bài kiểm tra môn Toán gồm 30 câu trong 1 tháng được thống kê lại như sau:

8	18	22	11	25	23	21	13	14	15
18	25	10	11	24	12	14	14	6	8

a) Tìm tứ phân vị của dãy số liệu trên.

b) Tổng hợp lại dãy số liệu trên vào bảng tần số ghép nhóm theo mẫu sau:

Điểm số	[6;10]	[11;15]	[16;20]	[21;25]
Số trận	?	?	?	?

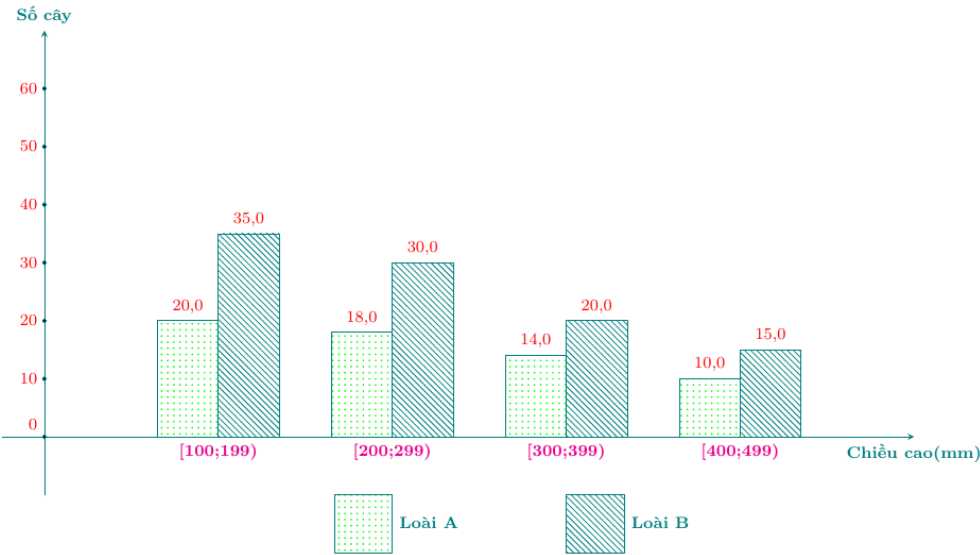
c) Hãy ước lượng tứ phân vị của số liệu ở bảng tần số ghép nhóm trên.

Bài 3: Trong một cuộc đua Marathon được tổ chức ở thành phố A người ta thống kê lại được như sau

Thời gian	[120;140)	[140;160)	[160;180)	[180;200)	[200;220)
Số người	4	6	10	15	25

Hãy ước lượng số trung bình và tứ phân vị của mẫu số liệu ghép nhóm trên.

Bài 4: Chiều cao của hai loài hoa được một người thống kê theo biểu đồ sau:



a) Hãy so sánh chiều cao của loài A và loài B theo số trung bình và trung vị.

b) Hãy ước lượng tứ phân vị thứ nhất và thứ ba của chiều cao của loài A và loài B.

D. BÀI TẬP TRẮC NGHIỆM

Câu 1: Số lượng huy chương vàng tại Sea Games 31 được thống kê

Số huy chương	$[0;10)$	$[10;50)$	$[50;100)$	$[100;210)$
Quốc gia	5	2	3	1

Xác định trung vị của mẫu số liệu ghép nhóm trên

- A) 10 B) 15 C) 20 D) 30.

Câu 2: Đo cân nặng của 1 lớp gồm 40 học sinh lớp 12B

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 hai của mẫu số liệu ghép nhóm thuộc khoảng nào sau đây?

- A) $[40;45]$ B) $[45;50]$ C) $[50;55]$ D) $[55;60]$

Câu 3: Đo cân nặng của 1 lớp gồm 40 học sinh lớp 12B

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 4: Thống kê điểm thi đánh giá năng lực của một trường THPT qua thang điểm 120 môn Toán

Điểm	$[0;20)$	$[20;40)$	$[40;60)$	$[60;80)$	$[80;100]$
Số học sinh	25	35	37	15	8

Trung vị của mẫu số liệu ghép nhóm thuộc khoảng nào sau đây?

- A) $[40;45]$ B) $[45;50]$ C) $[50;55]$ D) $[55;60]$

Câu 5: Thống kê điểm thi đánh giá năng lực của một trường THPT qua thang điểm 120 môn Toán

Điểm	$[0;20)$	$[20;40)$	$[40;60)$	$[60;80)$	$[80;100]$
Số học sinh	25	35	21	31	8

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

- A) $[60; 65]$ B) $[65; 70]$ C) $[50; 55]$ D) $[55; 60]$

Câu 6: Đo chiều cao các em học sinh khối 10 ta thu được kết quả

Chiều cao (cm)	Số học sinh
$[150; 152)$	5
$[152; 154)$	18
$[154; 156)$	40
$[156; 158)$	26
$[158; 160)$	8
$[160; 162)$	3

Trung vị của mẫu số liệu ghép nhóm trên là

- A) 154,5 B) 156,5 C) 153,5 D) 155,5

Câu 7: Đo chiều cao các em học sinh khối 10 ta thu được kết quả

Chiều cao (cm)	Số học sinh
$[150; 152)$	5
$[152; 154)$	18
$[154; 156)$	40
$[156; 158)$	26
$[158; 160)$	8
$[160; 162)$	3

Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm trên là

- A) 152,2 B) 153,3 C) 154,1 D) 151,5

Câu 8: Chiều dài của 60 lá dương xỉ trưởng thành

Lớp của chiều dài (cm)	Tần số
$[10; 20)$	8
$[20; 30)$	18
$[30; 40)$	24
$[40; 50)$	10

Tứ phân vị thứ hai của mẫu số liệu ghép nhóm trên là

- A) $\frac{95}{3}$ B) $\frac{91}{3}$ C) $\frac{89}{3}$ D) $\frac{93}{3}$

Câu 9: Khối lượng của 30 củ khoai lang thu hoạch ở 1 hộ gia đình

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

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

A) 99,5

B) 102,5

C) 100,5

D) 101,5

Câu 10: Khối lượng của 30 củ khoai lang thu hoạch ở 1 hộ gia đình

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

Trung vị của mẫu số liệu ghép nhóm trên là

A) 90

B) 95

C) 98

D) 97

Câu 11: Chiều cao của học sinh lớp 5 được mô tả như bảng sau

Lớp chiều cao	Tần số
[98; 103)	6
[103; 108)	7
[108; 113)	9
[113; 118)	5
[118; 123)	6
[123; 128)	4
[128; 133)	2
[133; 138)	2
[138; 143)	3
[143; 148]	1

Tứ phân vị thứ nhất của bảng phân bố ghép nhóm trên thuộc khoảng nào sau đây ?

A) [95;100]

B) [101;107]

C) [108;112]

D) [113;115]

Câu 12: Điểm thi của 32 học sinh trong kì thi Tiếng Anh (thang 100 điểm) được phân bố như sau

Lớp điểm	Tần số n_i
[40; 50)	4
[50; 60)	6
[60; 70)	10
[70; 80)	6
[80; 90)	4
[90; 100]	2

Tứ phân vị thứ ba của bảng phân bố ghép nhóm bằng

- A) $\frac{310}{4}$
- B) $\frac{315}{3}$
- C) $\frac{220}{4}$
- D) $\frac{230}{3}$

E. HƯỚNG DẪN GIẢI



BÀI TẬP TỰ LUYỆN

Bài 1: a) Các tứ phân vị là $Q_2 = 10,8; Q_1 = 9,05; Q_3 = 12,35$.

b) Bảng tần số ghép nhóm theo mẫu :

Lương tháng	[6;8)	[8;10)	[10;12)	[12;14)
Số giáo viên	3	6	8	7

c) Tứ phân vị thứ nhất là $Q_1 = 8 + \frac{\frac{24}{2} - 3}{6} \cdot (10 - 8) = 9$.

Tứ phân vị thứ hai là $Q_2 = 10 + \frac{\frac{24}{2} - 9}{8} \cdot (12 - 10) = 10,75$.

Tứ phân vị thứ ba là $Q_3 = 12 + \frac{\frac{3 \cdot 24}{3} - 17}{7} \cdot (14 - 12) \approx 12,29$.

Bài 2: a) Các tứ phân vị là $Q_2 = 14, Q_1 = 11, Q_3 = 21,5$.

b) Bảng tần số ghép nhóm theo mẫu :

Điểm	[6;10]	[11;15]	[16;20]	[21;25]
Tần số	4	8	2	6

$$\text{c) Tứ phân vị thứ nhất là } Q_1 = 11 + \frac{\frac{1.20}{4} - 4}{8} \cdot (15 - 11) = 11,5.$$

Tứ phân vị thứ hai là $Q_2 = 15,5$.

$$\text{Tứ phân vị thứ ba là } Q_3 = 21 + \frac{\frac{3.20}{6} - 14}{6} \cdot (25 - 21) \approx 21,67.$$

Bài 3: Ta có $n = 4 + 6 + 10 + 15 + 25 = 60$

$$\text{Trung vị là } M_e = 180 + \frac{\frac{60}{2} - 20}{15} \cdot (200 - 180) = \frac{580}{3} \approx 193,33.$$

$$\text{Tứ phân vị thứ nhất là } Q_1 = 160 + \frac{\frac{60}{4} - 10}{10} \cdot (180 - 160) = 170.$$

$$\text{Tứ phân vị thứ ba là } Q_3 = 200 + \frac{3 \cdot \frac{60}{4} - 35}{25} \cdot (220 - 200) = 208.$$

Bài 4: a) Bảng giá trị đại diện

Chiều cao đại diện (mm)	149,5	249,5	349,5	449,5
Loài A	20	18	14	10
Loài B	35	30	20	15

Chiều trung bình của loài A xấp xỉ bằng

$$\bar{x}_A = \frac{149,5 \cdot 20 + 249,5 \cdot 18 + 349,5 \cdot 14 + 449,5 \cdot 10}{62} \approx 272,08.$$

Chiều trung bình của loài B xấp xỉ bằng

$$\bar{x}_B = \frac{149,5.35 + 249,5.30 + 349,5.20 + 449,5.15}{100} = 201,5.$$

Chiều cao trung bình của loài A lớn hơn loài B .

Gọi $x_1 \leq x_2 \leq \dots \leq x_{85}$ là chiều cao của 62 cây loài A .

Nên trung vị của mẫu số liệu là $x_{31} \in [200; 299)$.

$$\text{Suy ra trung vị là } M_e = 200 + \frac{\frac{62}{2} - 20}{18} \cdot (299 - 200) = 260,5.$$

Gọi $y_1 \leq y_2 \leq \dots \leq y_{100}$ là chiều cao của 100 loài B .

Nên trung vị của mẫu số liệu là $\frac{1}{2}(x_{50} + x_{51}) \in [200; 299)$.

$$\text{Suy ra trung vị là } M_e = 200 + \frac{\frac{100}{2} - 35}{30} \cdot (299 - 200) = 249,5.$$

Chiều cao trung bình của loài A lớn hơn B .

b) Tứ phân vị thứ nhất của dãy số liệu $x_1 \leq x_2 \leq \dots \leq x_{62}$ là $x_{16} \in [100; 199)$.

Do đó tứ phân vị thứ nhất của loài A là

$$Q_1 = 100 + \frac{\frac{1.62}{4} - 0}{20} \cdot (199 - 100) = 176,725.$$

Tứ phân vị thứ ba của dãy số liệu $x_1 \leq x_2 \leq \dots \leq x_{62}$ là $x_{47} \in [300; 399)$.

Do đó tứ phân vị thứ ba của loài A là

$$Q_3 = 300 + \frac{\frac{3.62}{4} - 38}{14} \cdot (399 - 300) \approx 360,11.$$

Tứ phân vị thứ nhất của dãy số liệu $y_1 \leq y_2 \leq \dots \leq y_{100}$ là $y_{25} \in [100; 199)$.

Do đó tứ phân vị thứ nhất của loài B là

$$Q_1 = 100 + \frac{\frac{1.62}{4} - 0}{35} \cdot (199 - 100) \approx 143,84.$$

Tứ phân vị thứ ba của dãy số liệu $y_1 \leq y_2 \leq \dots \leq y_{100}$ là $y_{75} \in [300; 399)$.

Do đó tứ phân vị thứ ba của loài B là

$$Q_3 = 300 + \frac{\frac{3.62}{4} - 65}{20} \cdot (399 - 300) = 208,425.$$



BÀI TẬP TRẮC NGHIỆM

1.C	2.C	3.B	4.A	5.B	6.D	7.C	8.A	9.B	10.B
11.B	12.D								

Câu 1: Chọn C. Ta có $Q_2 = 10 + \frac{5,5 - 5}{2} \cdot 40 = 20$.

Câu 2: Chọn B. Ta có $Q_2 = 50 + \frac{\frac{40}{2} - 17}{7} \cdot 5 \approx 52,14$.

Câu 3: Chọn B. Ta có $Q_1 = 50 + \frac{\frac{40}{4} - 4}{13} \cdot 5 \approx 47,3$.

Câu 4: Chọn A. Ta có $Q_2 = 40$.

Câu 5: Chọn B. Ta có $Q_1 = 50 + \frac{\frac{40}{4} - 4}{13} \cdot 5 \approx 47,3$.

Câu 6: Chọn D. Ta có $Q_2 = 154 + \frac{50 - 23}{40} \cdot 2 \approx 155,35$.

Câu 7: Chọn C. Ta có $Q_1 = 154 + \frac{25 - 23}{40} \cdot 2 \approx 154,1$.

Câu 8: Chọn A. Ta có $Q_2 = 30 + \frac{30 - 26}{24} \cdot 10 = \frac{95}{3} \approx 31,67$.

Câu 9: Chọn B. Ta có $Q_1 = 50 + \frac{\frac{40}{4} - 4}{13} \cdot 5 \approx 47,3$.

Câu 10: Chọn B. Ta có $Q_2 = 90 + \frac{15 - 9}{12} \cdot 10 = 95$.

Câu 11: Chọn B. Ta có $Q_1 = 103 + \frac{\frac{45}{4} - 6}{7} \cdot 5 = 106,75$.

Câu 12: Chọn D. Ta có $Q_3 = 70 + \frac{24 - 20}{6} \cdot 10 = \frac{230}{3}$.

Chương V

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

BÀI TẬP CUỐI CHƯƠNG V

A. TÓM TẮT LÝ THUYẾT

I. Số liệu ghép nhóm

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

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

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

II. Số trung bình, mốt của mẫu số liệu ghép nhóm

- Số trung bình** của mẫu số liệu ghép nhóm:

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

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

$$M_0 = u_m + \frac{n_m - n_{m-1}}{(n_m - n_{m-1}) + (n_m - n_{m+1})} \cdot (u_{m+1} - u_m)$$

Trong đó: nhóm chứa mốt là $[u_m; u_{m+1})$; n_{m-1} là tần số của nhóm $[u_{m-1}; u_m)$; n_m là tần số của nhóm $[u_m; u_{m+1})$; n_{m+1} là tần số của nhóm $[u_{m+1}; u_{m+2})$

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

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

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

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

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

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

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

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

Khi đó:

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

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

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

Khi đó:

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

B. CÁC DẠNG TOÁN.

XÁC ĐỊNH CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CHO MẪU SỐ LIỆU GHEP NHÓM: SỐ TRUNG BÌNH, MỐT, TRUNG VỊ, TỬ PHÂN VỊ

Phương pháp:

+ Nắm vững khái niệm cơ bản

Ví dụ 1. Cho bảng tần số ghép nhóm số liệu thống kê chiều cao của 40 mẫu cây ở một vườn thực vật (đơn vị: centimét).

Nhóm	Tần số
[30; 40)	4
[40; 50)	10
[50; 60)	14
[60; 70)	6
[70; 80)	4
[80; 90)	2
	$n = 40$

a) Xác định số trung bình cộng, trung vị, tứ phân vị của mẫu số liệu ghép nhóm trên.

b) Mốt của mẫu số liệu trên là bao nhiêu?

Lời giải

a)

Nhóm	Giá trị đại diện	Tần số
[30; 40)	35	4
[40; 50)	45	10
[50; 60)	55	14
[60; 70)	65	6
[70; 80)	75	4
[80; 90)	85	2
		$n = 40$

❖ Tìm giá trị trung bình

Giá trị trung bình của mẫu số liệu là

$$\bar{x} = \frac{35.4 + 45.10 + 55.14 + 65.6 + 75.4 + 85.2}{40} = 55,5$$

❖ Tìm trung vị

- Khoảng chứa trung vị $[50; 60)$

Trung vị của mẫu số liệu là

$$M_e = 50 + \frac{20-14}{14} \cdot 10 \approx 54,29$$

❖ Tìm tứ phân vị

Nhóm chứa trung vị thứ nhất Q_1 và nhóm chứa trung vị thứ hai Q_3 lần lượt là $[40; 50)$, $[60; 70)$.

Khi đó:

$$Q_1 = 40 + \frac{10-4}{10} \cdot 10 = 46$$

$$Q_3 = 60 + \frac{30-28}{6} \cdot 10 = 63,3$$

Vậy tứ phân vị của mẫu số liệu là $Q_1 = 46$, $Q_2 = 54,29$, $Q_3 = 63,3$.

❖ Tìm mốt

Nhóm chứa mốt: $[50; 60)$

Mốt của mẫu số liệu:

$$M_0 = 50 + \frac{4}{4+8} \cdot 10 \approx 53,3$$

Ví dụ 2. Mẫu số liệu sau ghi lại cân nặng của 30 bạn học sinh (đơn vị: kilogram)

17 40 39 40,5 42 51 41,5 39 41 30
40 42 40,5 39,5 41 40,5 37 39,5 40 41
38,5 39,5 40 41 39 40,5 40 38,5 39,5 41,5

a) Lập bảng tần số ghép nhóm cho mẫu số liệu trên có tám nhóm ứng với tám nửa khoảng

$[15; 20), [20; 25), [25; 30), [30; 35), [35; 40), [40; 45), [45; 50), [50; 55)$.

b) Xác định số trung bình cộng, trung vị, tứ phân vị của mẫu số liệu ghép nhóm trên.

c) Mốt của mẫu số liệu ghép nhóm trên là bao nhiêu

Lời giải

a)

Nhóm	Giá trị đại diện	Tần số
$[15; 20)$	17,5	1
$[20; 25)$	22,5	0

[25;30)	27,5	0
[30;35)	32,5	1
[35;40)	37,5	10
[40;45)	42,5	17
[45;50)	47,5	0
[50;55)	52,5	1
		$n = 30$

b)

Số trung bình cộng: $\bar{x} = \frac{17,5 + 32,5 + 37,5 \cdot 10 + 42,5 \cdot 17 + 52,5}{30} = 40$

Trung vị: $Q_2 = 40 + \frac{15-12}{17} \cdot 5 \approx 40,9$

Trung vị Q_1 và Q_3 :

$$Q_1 = 35 + \frac{7,5-2}{10} \cdot 5 = 37,75$$

$$Q_3 = 40 + \frac{22,5-12}{17} \cdot 5 \approx 43,1$$

Tứ phân vị: $Q_1 = 37,75$, $Q_2 = 40,9$, $Q_3 = 43,1$

c) Mốt của mẫu số liệu

$$M_0 = 40 + \frac{7}{7+17} \cdot 5 \approx 41,5$$

C. BÀI TẬP TỰ LUYỆN.

Bài 1: 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

a) Xác định giá trị trung bình, trung vị, tứ phân vị của mẫu số liệu.

b) Tìm mốt của mẫu số liệu.

Bài 2: Để chuẩn bị cho đồ án tốt nghiệp, một sinh viên y khoa đã khảo sát huyết áp tối đa của một số bệnh nhân và lập được bảng tần số ghép nhóm sau:

Huyết áp	Tần số
[90;110)	6
[110;130)	20
[130;150)	35
[150;170)	45
[170;190)	30
[190;210)	16

- a) Xác định trung bình, trung vị, tứ phân vị và một của mẫu số liệu.
b) Hãy giải thích vì sao trong trường hợp này, cả ba giá trị trung bình, trung vị và một tìm được đều đại diện tốt cho huyết áp của những bệnh nhân được khảo sát.

Bài 3: 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. Họ tổ chức dữ liệu thu được vào bảng

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

Hãy tính (làm tròn kết quả đến hàng phần mười)

- a) Trung vị của mẫu số liệu.
b) Trung bình của mẫu số liệu. Trong trường hợp này thì trung bình hay trung vị đại diện tốt hơn cho mẫu số liệu?

D. BÀI TẬP TRẮC NGHIỆM

(Sử dụng mẫu số liệu này cho các câu từ câu 1 đến câu 5)

Khảo sát thời gian chạy bộ trong một ngày của một số học sinh khối 11 thu được mẫu số liệu ghép nhóm sau:

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

Câu 1: Giá trị đại diện của nhóm [20;40) là

- A) 10. B) 20. C) 30. D) 40.

Câu 2: Mẫu số liệu ghép nhóm này có một là

- A) 59. B) 40. C) 52. D) 53.

Câu 3: Nhóm chứa mốt của mẫu số liệu này là

- A) [20; 40) B) [40; 60) C) [60; 80) D) [80; 100)

Câu 4: Nhóm chứa tứ phân vị thứ nhất là

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

Câu 5: Nhóm chứa trung vị là

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

(Sử dụng mẫu số liệu sau cho các câu từ câu 6 đến câu 8)

Người ta tiến hành phỏng vấn 40 người về một mẫu áo khoác. Người điều tra yêu cầu cho điểm mẫu áo đó theo thang điểm là 100. Kết quả được trình bày trong bảng ghép nhóm sau:

Nhóm	[50; 60)	[60; 70)	[70; 80)	[80; 90)	[90; 100)	
Tần số	4	5	23	6	2	$N = 40$

Câu 6: Trung vị của mẫu số liệu ghép nhóm trên gần nhất với giá trị

- A) 74. B) 75. C) 76. D) 77.

Câu 7: Tứ phân vị của mẫu số liệu ghép nhóm trên (làm tròn đến hàng đơn vị) là

- A) $Q_1 \approx 71, Q_2 \approx 76, Q_3 \approx 78.$ B) $Q_1 \approx 71, Q_2 \approx 75, Q_3 \approx 78.$
C) $Q_1 \approx 70, Q_2 \approx 76, Q_3 \approx 79.$ D) $Q_1 \approx 70, Q_2 \approx 75, Q_3 \approx 79.$

Câu 8: Mốt của mẫu số liệu ghép nhóm trên (làm tròn kết quả đến hàng đơn vị) là

- A) 73. B) 74. C) 75. D) 76.

Câu 9: Dưới đây là một mẫu số liệu cho ở dạng bảng tần số ghép nhóm

Nhóm	[0; 50)	[50; 100)	[100; 150)	[150; 200)	
Tần số	6	8	7	6	$N = 27$

Trung bình của mẫu số liệu là một số thỏa mãn điều kiện

- A) $\bar{x} \approx 112,8.$ B) $\bar{x} \approx 107,8.$ C) $\bar{x} \approx 102,5.$ D) $\bar{x} \approx 85,5.$

Câu 10: Dưới đây là một mẫu số liệu cho ở dạng bảng tần số ghép nhóm

Nhóm ghép	Tần số	Nhóm ghép	Tần số
[62,5; 67,5)	4	[82,5; 87,5)	22
[67,5; 72,5)	7	[87,5; 92,5)	5
[72,5; 77,5)	10	[92,5; 97,5)	10
[77,5; 82,5)	26	[97,5; 102,5)	16

A) $77,5 \leq M_e < 82,5$.

B) $82,5 \leq M_e < 87,5$.

C) $87,5 \leq M_e < 92,5$.

D) $92,5 \leq M_e < 97,5$.

E. HƯỚNG DẪN GIẢI



BÀI TẬP TỰ LUYỆN

Bài 1:

Hàm lượng chất béo (g)	[2;6)	[6;10)	[10;14)	[14;18)	[18;22)	[22;26)
Giá trị đại diện	4	8	12	16	20	24
Tần số	2	6	10	13	16	13

a) Giá trị trung bình: $\bar{x} = \frac{4.2 + 8.6 + 12.10 + 16.13 + 20.16 + 24.13}{60} = 16,9$

Trung vị: $M_e = 14 + \frac{30-18}{13}.4 = 17,7$

Trung vị thứ nhất và thứ ba:

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

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

Tứ phân vị của mẫu số liệu là $Q_1 = 12,8; Q_2 = 17,7; Q_3 = 21,5$

b) Mốt của mẫu số liệu: $M_0 = 18 + \frac{3}{3+3}.4 = 20$

Bài 2:

Huyết áp	Giá trị đại diện	Tần số
[90;110)	100	6
[110;130)	120	20
[130;150)	140	35
[150;170)	160	45
[170;190)	180	30
[190;210)	200	16

a) Giá trị trung bình:

$$\bar{x} = \frac{100.6 + 120.20 + 140.35 + 160.45 + 180.30 + 200.16}{152} = 155,9$$

$$\text{Trung vị: } M_e = 150 + \frac{76 - 61}{45} \cdot 20 = 156,7$$

Trung vị thứ nhất và thứ ba:

$$Q_1 = 130 + \frac{38 - 26}{35} \cdot 20 = 136,9$$

$$Q_3 = 170 + \frac{114 - 106}{30} \cdot 20 = 175,3$$

Tứ phân vị của mẫu số liệu: $Q_1 = 136,9$; $Q_2 = 156,7$; $Q_3 = 175,3$

$$\text{Mốt: } M_0 = 150 + \frac{10}{10 + 15} \cdot 20 = 158$$

b) Vì ba giá trị này cùng thuộc vào một nhóm ghép có tần số cao nhất và chúng gần như không có sự chênh lệch nhiều.

Bài 3:

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

a) Trung vị của mẫu số liệu

$$M_e = 9 + \frac{42,5 - 24}{31} \cdot 4 \approx 11,4$$

b) Trung bình của mẫu số liệu là $\bar{x} = 13,2$

Trung bình thuộc nhóm $[13;17)$ cho thấy trong 85 số liệu đã có ít nhất 55 số liệu nhỏ hơn số trung bình. Suy ra, trong trường hợp này thì trung vị là số đại diện tốt hơn cho mẫu số liệu.

BÀI TẬP TRẮC NGHIỆM

1.C	2.C	3.B	4.B	5.C	6.B	7.D	8.B	9.C	10.B
-----	-----	-----	-----	-----	-----	-----	-----	-----	------

D. BÀI TẬP TRẮC NGHIỆM

Câu 1: Chọn C. $\frac{20 + 40}{2} = 30$

Câu 2: Chọn C. $M_0 = 40 + \frac{3}{3+2} \cdot 20 = 52$

Câu 3: Chọn B.

Câu 4: Chọn B. $\frac{N}{4} = \frac{42}{4} = 10,5$; $5+9=14$

Nhóm chứa trung vị thứ nhất $[20;40)$

Câu 5: Chọn C. $\frac{N}{2} = \frac{42}{2} = 21$; $5+9+12=26$

Nhóm chứa trung vị $[40;60)$

Câu 6: Chọn B. $M_e = 70 + \frac{20-9}{23} \cdot 10 = 75$

Câu 7: Chọn D. $Q_1 = 70 + \frac{10-9}{23} \cdot 10 = 70$ và $Q_3 = 70 + \frac{30-9}{23} \cdot 10 = 79$

Câu 8: Chọn B. $M_0 = 70 + \frac{18}{18+27} \cdot 10 = 74$

Câu 9: Chọn C. $\bar{x} = \frac{25.6+75.8+125.7+175.6}{27} = 99,1$

Câu 10: Chọn B. $\frac{N}{2} = \frac{100}{2} = 50$

$$4+7+10+26+22=69$$

Nhóm chứa trung vị là $[82,5;87,5)$