



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

BÀI 1: SỐ TRUNG BÌNH VÀ MÓT CỦA MẪU SỐ LIỆU GHÉP NHÓM

I LÝ THUYẾT.

1. SỐ LIỆU GHÉP NHÓM

Một số loại số liệu điều tra có thể nhận rất nhiều những giá trị khác nhau, hoặc khó xác định được giá trị chính xác, ví dụ như chiều cao, cân nặng, tuổi thọ, ... Để thuận tiện cho việc lưu trữ và xử lý các loại số liệu này, người ta thường ghép các số liệu gần nhau lại thành nhóm. Mẫu số liệu ghép hóm 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 $n = n_1 + n_2 + ... + n_k$.
- Giá trị chính giữa của 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 số 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.

Ví dụ. Bảng thống kê sau cho biết thời gian chạy (phút) của 30 vận động viên (VĐV) trong một giải chạy Marathon.

Thời gian	129	130	133	134	135	136	138	141	142	143	144	145
Số VĐV	1	2	1	1	1	2	3	3	4	5	2	5

Hãy chuyển mẫu số liệu trên sang mẫu số liệu ghép nhóm gồm sáu nhóm có độ dài bằng nhau và bằng 3.

Lời giải

Giá trị nhỏ nhất là 129, giá trị lớn nhất là 145 nên khoảng biến thiên là $145 - 129 = 16$. Tổng độ dài của sáu nhóm là 18. Để cho đối xứng, ta chọn đầu mút trái của nhóm đầu tiên là 127,5 và đầu mút phải của nhóm cuối cùng là 145,5 ta được các nhóm là $[127,5; 130,5)$, $[130,5; 133,5)$, ..., $[142,5; 145,5)$. Đếm số giá trị thuộc mỗi nhóm, ta có mẫu số liệu ghép nhóm như sau:

Thời gian	$[127,5; 130,5)$	$[130,5; 133,5)$	$[133,5; 136,5)$	$[136,5; 139,5)$	$[139,5; 142,5)$	$[142,5; 145,5)$
Số VĐV	3	1	4	3	7	12

2. 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 :

$$\bar{x} = \frac{n_1 c_1 + n_2 c_2 + \dots + n_k c_k}{n}, \text{ trong đó } n = n_1 + n_2 + \dots + n_k.$$

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.

Ví dụ 1. Tìm cân nặng trung bình của học sinh lớp 11D cho trong bảng sau:

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

Lời giải

Trong mỗi khoảng cân ặng, giá trị đại diện 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

$$\bar{x} = \frac{10.43 + 7.48 + 16.53 + 4.58 + 2.63 + 3.68}{42} \approx 51,81 \text{ (kg)}$$

Ví dụ 2. Tìm hiệu 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

Tính thời gian xem tivi trung bình trong tuần trước của các bạn học sinh này.

Lời giải

Trong mỗi khoảng thời gian, giá trị đại diện trung bình cộng của giá trị hai đầu mút nên ta có bảng sau:

Thời gian (giờ)	2,5	7,5	12,5	17,5	22,5
Số học sinh	8	16	4	2	2

Tổng số học sinh là $n = 32$. Thời gian xem tivi trung bình của học sinh là

$$\bar{x} = \frac{2,5.8 + 7,5.16 + 12,5.4 + 17,5.2 + 22,5.2}{32} \approx 8,44 (h)$$

3. MỐT

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

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

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

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

Ý nghĩa của một của mẫu số liệu ghép nhóm

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

Ví dụ 1. Bảng số liệu ghép nhóm sau cho biết chiều cao (cm) của 50 học sinh lớp 11A.

Khoảng chiều cao (cm)	[145;150)	[150;155)	[155;160)	[160;165)	[165;170)
Số học sinh	7	14	10	10	9

Tính một của mẫu số liệu ghép nhóm này. Có thể kết luận gì từ giá trị được?

Lời giải

Tần số lớn nhất là 14 nên nhóm chứa một là nhóm $[150;155)$.

Ta có $j = 2; a_2 = 150; m_2 = 14; m_1 = 7; m_3 = 10; h = 5$. Do đó

$$M_o = 150 + \frac{14 - 7}{(14 - 7) + (14 - 10)} . 5 \approx 153,18.$$

Số học sinh có chiều cao khoảng 153,18 cm là nhiều nhất.

Ví dụ 2. 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	2	10	6	4	3

Tính một của mẫu số liệu ghép nhóm này.

Lời giải

Tần số lớn nhất là 10 nên nhóm chứa một là nhóm $[10,5;20,5)$.

Ta có $j = 2; a_2 = 10,5; m_2 = 10; m_1 = 2; m_3 = 6; h = 10$.

Do đó $M_0 = 10,5 + \frac{10-2}{(10-2)+(10-6)} \cdot 10 \approx 17,17$.



HỆ THỐNG BÀI TẬP TỰ LUẬN.

Câu 1: Trong các mẫu số liệu sau, mẫu nào là mẫu số liệu ghép nhóm? Đọc và giải thích mẫu số liệu ghép nhóm đó.

a) Số tiền mà sinh viên chi cho thanh toán cước điện thoại trong tháng.

Số tiền (nghìn đồng)	[0; 50)	[50; 100)	[100; 150)	[150; 200)	[200; 250)
Số sinh viên	5	12	23	17	3

b) Thống kê nhiệt độ tại một điểm trong 40 ngày, ta có bảng số liệu sau

Nhiệt độ (°C)	[19; 22)	[22; 25)	[25; 28)	[28; 31)
Số ngày	7	15	12	6

Câu 2: Cho mẫu số liệu về số tiền điện phải trả của 50 gia đình trong một tháng ở một khu phố (đơn vị: nghìn đồng).

Giá trị	[375; 450)	[450; 525)	[525; 600)	[600; 675)	[675; 750)	[750; 825)
Số lượng gia đình	6	15	10	6	9	4

Đọc và giải thích mẫu số liệu này.

Câu 3: Cho mẫu số liệu về khối lượng của 30 củ khoai tây thu hoạch ở một nông trường (đơn vị: gam).

Giá trị	[70; 80)	[80; 90)	[90; 100)	[100; 110)	[110; 120)
Số lượng củ khoai	3	6	12	6	3

Đọc và giải thích mẫu số liệu này.

Câu 4: Số sản phẩm một công nhân làm được trong một ngày được cho như sau

18 25 39 12 54 27 46 25 19 8 36 22
20 19 17 44 5 18 23 28 25 34 46 27 16

Hãy chuyển mẫu số liệu trên sang dạng ghép nhóm với bảy nhóm có độ dài bằng nhau.

Câu 5: Thời gian ra sân (giờ) của một số cựu cầu thủ ở giải ngoại hạng Anh qua các thời kì được cho như sau:

653 632 609 572 565 535 516 514
508 505 504 504 503 499 496 492

Hãy chuyển mẫu số liệu trên sang dạng ghép nhóm với bảy nhóm có độ dài bằng nhau.

CHUYÊN ĐỀ V – TOÁN – 11 – CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CỦA MẪU SỐ LIỆU GHEP NHÓM

Câu 6: 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

Hãy chuyển mẫu số liệu trên sang mẫu số liệu ghép nhóm gồm 8 nhóm có độ dài bằng nhau và bằng 5.

Câu 7: Trong các mẫu số liệu sau, mẫu nào là mẫu số liệu ghép nhóm? Đọc và giải thích mẫu số liệu ghép nhóm đó.

a) Số tiền mà sinh viên chi cho thanh toán cước điện thoại trong tháng.

Số tiền (nghìn đồng)	[0; 50)	[50; 100)	[100; 150)	[150; 200)	[200; 250)
Số sinh viên	5	12	23	17	3

b) Thống kê nhiệt độ tại một 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

Câu 8: Cho mẫu số liệu về số tiền điện phải trả của 50 gia đình trong một tháng ở một khu phố (đơn vị: nghìn đồng).

Giá trị	[375; 450)	[450; 525)	[525; 600)	[600; 675)	[675; 750)	[750; 825)
Số lượng gia đình	6	15	10	6	9	4

Đọc và giải thích mẫu số liệu này.

Câu 9: Cho mẫu số liệu về khối lượng của 30 củ khoai tây thu hoạch ở một nông trường (đơn vị: gam).

Giá trị	[70; 80)	[80; 90)	[90; 100)	[100; 110)	[110; 120)
Số lượng củ khoai	3	6	12	6	3

Đọc và giải thích mẫu số liệu này.

Câu 10: Số sản phẩm một công nhân làm được trong một ngày được cho như sau

18 25 39 12 54 27 46 25 19 8 36 22
20 19 17 44 5 18 23 28 25 34 46 27 16

Hãy chuyển mẫu số liệu trên sang dạng ghép nhóm với bảy nhóm có độ dài bằng nhau.

Câu 11: Thời gian ra sân (giờ) của một số cựu cầu thủ ở giải ngoại hạng Anh qua các thời kì được cho như sau:

653 632 609 572 565 535 516 514
508 505 504 504 503 499 496 492

Hãy chuyển mẫu số liệu trên sang dạng ghép nhóm với bảy nhóm có độ dài bằng nhau.

**CHUYÊN ĐỀ V – TOÁN – 11 – CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CỦA MẪU SỐ
LIỆU GHEP NHÓM**

Câu 12: 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

Hãy chuyển mẫu số liệu trên sang mẫu số liệu ghép nhóm gồm 8 nhóm có độ dài bằng nhau và bằng 5.

Câu 13: Thống kê điểm trung bình môn Toán của một số học sinh lớp 11 được cho ở bảng sau:

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

Hãy ước lượng số trung bình và môđ của mẫu số liệu ở bảng tần số ghép nhóm trên.

Câu 14: Để kiểm tra thời gian sử dụng pin của một chiếc điện thoại mới, chị An thống kê thời gian sử dụng điện thoại của mình từ lúc sạc đầy pin cho đến khi hết pin ở bảng sau:

Thời gian sử dụng (giờ)	$[7; 9)$	$[9; 11)$	$[11; 13)$	$[13; 15)$	$[15; 17)$
Số lần	2	5	7	6	3

Hãy ước lượng thời gian sử dụng trung bình từ lúc chị An sạc đầy pin điện thoại cho tới khi hết pin.

Câu 15: Tổng số lượng mưa trong tháng 8 đo được tại một trạm quan trắc đặt tại Vũng Tàu từ năm 2002 đến năm 2020 được ghi lại như dưới đây (đơn vị: mm)

121,8 158,3 334,9 200,9 165,6 161,5 194,3 220,7 189,8 243,2
165,9 165,9 134 173 169, 189, 254 168 255

(Nguồn: Tổng cục Thống kê)

a) Xác định số trung bình và tứ phân vị của mẫu số liệu trên.

b) Hoàn thành bảng tần số ghép nhóm theo mẫu sau:

Tổng lượng mưa trong tháng 8 (mm)	$[120; 175)$	$[175; 230)$	$[230; 285)$	$[285; 340)$
So năm	?	?	?	?

c) Hãy ước lượng số trung bình và môđ của mẫu số liệu ở bảng tần số ghép nhóm trên.

Câu 16: Bảng sau thống kê số ca nhiễm mới SARS-CoV-2 mỗi ngày trong tháng 12/2021 tại Việt Nam.

Ngày	Số ca	Ngày	Số ca	Ngày	Số ca	Ngày	Số ca
1	15 139	9	15 965	17	15 685	25	16 046
2	14 295	10	15 474	18	16 363	26	15 667
3	14 254	11	16 830	19	16 586	27	15 310
4	14 598	12	15 264	20	15 420	28	14 866
5	14 927	13	16 035	21	16 806	29	14 299
6	15 215	14	15 871	22	17 044	30	20 454
7	14 433	15	16 192	23	16 860	31	17 004
8	15 223	16	15 720	24	16 633		

(Nguồn: worldmeters.info)

a) Xác định số trung bình và tứ phân vị của mẫu số liệu trên.

b) Hoàn thành bảng tần số ghép nhóm theo mẫu sau:

So ca (nghìn)	[14; 15,5)	[15,5; 17)	[17; 18,5)	[18,5; 20)	[20; 21,5)
So ngày	?	?	?	?	?

c) Hãy ước lượng số trung bình, tứ phân vị và một của mẫu số liệu ở bảng tần số ghép nhóm trên.



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

BÀI 1: SỐ TRUNG BÌNH VÀ MÓT CỦA MẪU SỐ LIỆU GHÉP NHÓM



LÝ THUYẾT.

1. SỐ LIỆU GHÉP NHÓM

Một số loại số liệu điều tra có thể nhận rất nhiều những giá trị khác nhau, hoặc khó xác định được giá trị chính xác, ví dụ như chiều cao, cân nặng, tuổi thọ, ... Để thuận tiện cho việc lưu trữ và xử lý các loại số liệu này, người ta thường ghép các số liệu gần nhau lại thành nhóm. Mẫu số liệu ghép hóm 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 của mỗi nhóm được dùng làm giá trị đại diện cho nhóm ấy. Ví dụ nhóm $[u_1; u_2)$ có giá trị đại diện là $\frac{1}{2}(u_1 + u_2)$.
- Hiệu $u_{j+1} - u_j$ được gọi là độ dài của nhóm $[u_j; u_{j+1})$.

Một số quy tắc ghép nhóm của mẫu số liệu

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

- Sử dụng từ $k = 5$ đến $k = 20$ nhóm. Cỡ mẫu càng lớn thì cần càng nhiều nhóm số liệu.

Các nhóm có cùng độ dài bằng L thỏa mãn $R < k.L$, trong đó R là khoảng biến thiên, k là số nhóm.

- Giá trị nhỏ nhất của mẫu số 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.

Ví dụ. Bảng thống kê sau cho biết thời gian chạy (phút) của 30 vận động viên (VĐV) trong một giải chạy Marathon.

Thời gian	129	130	133	134	135	136	138	141	142	143	144	145
Số VĐV	1	2	1	1	1	2	3	3	4	5	2	5

Hãy chuyển mẫu số liệu trên sang mẫu số liệu ghép nhóm gồm sáu nhóm có độ dài bằng nhau và bằng 3.

Lời giải

Giá trị nhỏ nhất là 129, giá trị lớn nhất là 145 nên khoảng biến thiên là $145 - 129 = 16$. Tổng độ dài của sáu nhóm là 18. Để cho đối xứng, ta chọn đầu mút trái của nhóm đầu tiên là 127,5 và đầu mút phải của nhóm cuối cùng là 145,5 ta được các nhóm là $[127,5; 130,5)$, $[130,5; 133,5)$, ..., $[142,5; 145,5)$. Đếm số giá trị thuộc mỗi nhóm, ta có mẫu số liệu ghép nhóm như sau:

Thời gian	$[127,5; 130,5)$	$[130,5; 133,5)$	$[133,5; 136,5)$	$[136,5; 139,5)$	$[139,5; 142,5)$	$[142,5; 145,5)$
Số VĐV	3	1	4	3	7	12

2. 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 :

$$\bar{x} = \frac{n_1 c_1 + n_2 c_2 + \dots + n_k c_k}{n}, \text{ trong đó } n = n_1 + n_2 + \dots + n_k.$$

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.

Ví dụ 1. Tìm cân nặng trung bình của học sinh lớp 11D cho trong bảng sau:

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

Lời giải

Trong mỗi khoảng cân ặng, giá trị đại diện 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

$$\bar{x} = \frac{10.43 + 7.48 + 16.53 + 4.58 + 2.63 + 3.68}{42} \approx 51,81 \text{ (kg)}$$

Ví dụ 2. Tìm hiệu 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

Tính thời gian xem tivi trung bình trong tuần trước của các bạn học sinh này.

Lời giải

Trong mỗi khoảng thời gian, giá trị đại diện trung bình cộng của giá trị hai đầu mút nên ta có bảng sau:

Thời gian (giờ)	2,5	7,5	12,5	17,5	22,5
Số học sinh	8	16	4	2	2

Tổng số học sinh là $n = 32$. Thời gian xem tivi trung bình của học sinh là

$$\bar{x} = \frac{2,5.8 + 7,5.16 + 12,5.4 + 17,5.2 + 22,5.2}{32} \approx 8,44 (h)$$

3. MỐT

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

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

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

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

Ý nghĩa của một của mẫu số liệu ghép nhóm

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

Ví dụ 1. Bảng số liệu ghép nhóm sau cho biết chiều cao (cm) của 50 học sinh lớp 11A.

Khoảng chiều cao (cm)	[145;150)	[150;155)	[155;160)	[160;165)	[165;170)
Số học sinh	7	14	10	10	9

Tính một của mẫu số liệu ghép nhóm này. Có thể kết luận gì từ giá trị được?

Lời giải

Tần số lớn nhất là 14 nên nhóm chứa một là nhóm $[150;155)$.

Ta có $j = 2; a_2 = 150; m_2 = 14; m_1 = 7; m_3 = 10; h = 5$. Do đó

$$M_o = 150 + \frac{14 - 7}{(14 - 7) + (14 - 10)} . 5 \approx 153,18.$$

Số học sinh có chiều cao khoảng 153,18 cm là nhiều nhất.

Ví dụ 2. 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	2	10	6	4	3

Tính một của mẫu số liệu ghép nhóm này.

Lời giải

Tần số lớn nhất là 10 nên nhóm chứa một là nhóm $[10, 5; 20, 5)$.

Ta có $j = 2; a_2 = 10, 5; m_2 = 10; m_1 = 2; m_3 = 6; h = 10$.

Do đó $M_0 = 10, 5 + \frac{10 - 2}{(10 - 2) + (10 - 6)} \cdot 10 \approx 17, 17$.



HỆ THỐNG BÀI TẬP TỰ LUẬN.

Câu 1: Trong các mẫu số liệu sau, mẫu nào là mẫu số liệu ghép nhóm? Đọc và giải thích mẫu số liệu ghép nhóm đó.

a) Số tiền mà sinh viên chi cho thanh toán cước điện thoại trong tháng.

Số tiền (nghìn đồng)	[0; 50)	[50; 100)	[100; 150)	[150; 200)	[200; 250)
Số sinh viên	5	12	23	17	3

b) Thống kê nhiệt độ tại một điểm trong 40 ngày, ta có bảng số liệu sau

Nhiệt độ (°C)	[19; 22)	[22; 25)	[25; 28)	[28; 31)
Số ngày	7	15	12	6

Lời giải

Cả hai mẫu số liệu trên đều là mẫu số liệu ghép nhóm.

a) Có năm nhóm là

Dưới 50 nghìn đồng có 5 sinh viên.

Từ 50 đến dưới 100 nghìn đồng có 2 sinh viên.

Từ 100 đến dưới 150 nghìn đồng có 23 sinh viên.

Từ 150 đến dưới 200 nghìn đồng có 17 sinh viên.

Từ 200 đến dưới 250 nghìn đồng có 3 sinh viên.

b) Có bốn nhóm là

Từ 19°C đến dưới 22°C có 7 ngày.

Từ 22°C đến dưới 25°C có 15 ngày.

Từ 25°C đến dưới 28°C có 12 ngày.

Từ 28°C đến dưới 31°C có 6 ngày.

Câu 2: Cho mẫu số liệu về số tiền điện phải trả của 50 gia đình trong một tháng ở một khu phố (đơn vị: nghìn đồng).

Giá trị	[375; 450)	[450; 525)	[525; 600)	[600; 675)	[675; 750)	[750; 825)
Số lượng gia đình	6	15	10	6	9	4

Đọc và giải thích mẫu số liệu này.

Lời giải

Mẫu số liệu đã cho là mẫu số liệu ghép nhóm.

Có tất cả 6 nhóm là: từ 375 nghìn đồng đến dưới 450 nghìn đồng có 6 gia đình, từ 450 nghìn đồng đến dưới 525 nghìn đồng có 15 gia đình, từ 525 nghìn đồng đến dưới 600 nghìn đồng có 10 gia đình, từ 600 nghìn đồng đến dưới 675 nghìn đồng có 6 gia đình, từ 675 nghìn đồng đến dưới 750 nghìn đồng có 9 gia đình và từ 750 nghìn đồng đến dưới 825 nghìn đồng có 4 gia đình.

Câu 3: Cho mẫu số liệu về khối lượng của 30 củ khoai tây thu hoạch ở một nông trường (đơn vị: gam).

Giá trị	[70;80)	[80;90)	[90;100)	[100;110)	[110;120)
Số lượng củ khoai	3	6	12	6	3

Đọc và giải thích mẫu số liệu này.

Lời giải

Mẫu số liệu đã cho là mẫu số liệu ghép nhóm.

Có tất cả 5 nhóm là: từ 70 gam đến dưới 80 gam có 3 củ, từ 80 gam đến dưới 90 gam có 6 củ, từ 90 gam đến dưới 100 gam có 12 củ và từ 100 gam đến dưới 110 gam có 6 củ, từ 110 gam đến dưới 120 gam có 3 củ.

Câu 4: Số sản phẩm một công nhân làm được trong một ngày được cho như sau

18 25 39 12 54 27 46 25 19 8 36 22
20 19 17 44 5 18 23 28 25 34 46 27 16

Hãy chuyển mẫu số liệu trên sang dạng ghép nhóm với bảy nhóm có độ dài bằng nhau.

Lời giải

Khoảng biến thiên là $54 - 5 = 49$.

Ta chia thành các nhóm sau $[4,5; 13)$; $[13; 21,5)$; $[21,5; 30)$; ... ; $[47; 55,5)$.

Đếm số giá trị của mỗi nhóm, ta có bảng ghép nhóm sau:

Số sản phẩm	$[4,5; 13)$	$[13; 21,5)$	$[21,5; 30)$	$[30; 38,5)$	$[38,5; 47)$	$[47; 55,5)$
Số công nhân	3	7	8	2	4	1

Câu 5: Thời gian ra sân (giờ) của một số cựu cầu thủ ở giải ngoại hạng Anh qua các thời kì được cho như sau:

653 632 609 572 565 535 516 514
508 505 504 504 503 499 496 492

Hãy chuyển mẫu số liệu trên sang dạng ghép nhóm với bảy nhóm có độ dài bằng nhau.

Lời giải

Khoảng biến thiên là $653 - 492 = 161$.

Ta chia thành các nhóm sau $[492; 515)$; $[515; 538)$; $[538; 561)$; ... ; $[47; 55,5)$.

Đếm số giá trị của mỗi nhóm, ta có bảng ghép nhóm sau:

Thời gian (giờ)	$[492; 515)$	$[515; 538)$	$[538; 561)$	$[561; 584)$	$[584; 607)$	$[607; 630)$	$[630; 653]$
Số cầu thủ	9	2	0	2	0	1	2

Câu 6: 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

Hãy chuyển mẫu số liệu trên sang mẫu số liệu ghép nhóm gồm 8 nhóm có độ dài bằng nhau và bằng 5.

Lời giải

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

Câu 7: Thống kê điểm trung bình môn Toán của một số học sinh lớp 11 được cho ở bảng sau:

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

Hãy ước lượng số trung bình và một của mẫu số liệu ở bảng tần số ghép nhóm trên.

Lời giải

Bảng tần số ghép nhóm theo giá trị đại diện là

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

Điểm trung bình môn Toán của một số học sinh lớp 11 là

$$\bar{x} = \frac{8.6,75 + 10.7,25 + 16.7,75 + 24.8,25 + 13.8,75 + 7.9,25 + 4.9,75}{82} \approx 8,12$$

Một

Mốt M_0 chứa trong nhóm $[8; 8,5)$

Do đó: $u_m = 8; u_{m+1} = 8,5 \Rightarrow u_{m+1} - u_m = 8,5 - 8 = 0,5$

$$n_{m-1} = 16; n_m = 24; n_{m+1} = 13$$

$$M_0 = 8 + \frac{24-16}{(24-16) + (24-13)}(8,5-8) = 8,21$$

Câu 8: Để kiểm tra thời gian sử dụng pin của một chiếc điện thoại mới, chị An thống kê thời gian sử dụng điện thoại của mình từ lúc sạc đầy pin cho đến khi hết pin ở bảng sau:

Thời gian sử dụng (giờ)	$[7; 9)$	$[9; 11)$	$[11; 13)$	$[13; 15)$	$[15; 17)$
Số lần	2	5	7	6	3

Hãy ước lượng thời gian sử dụng trung bình từ lúc chị An sạc đầy pin điện thoại cho tới khi hết pin.

Lời giải

Thời gian sử dụng trung bình:

$$\bar{x} = \frac{2.8 + 5.10 + 7.12 + 6.14 + 3.16}{2 + 5 + 7 + 6 + 3} \approx 12,26$$

Câu 9: Tổng số lượng mưa trong tháng 8 đo được tại một trạm quan trắc đặt tại Vũng Tàu từ năm 2002 đến năm 2020 được ghi lại như dưới đây (đơn vị: mm)

121,8 158,3 334,9 200,9 165,6 161,5 194,3 220,7 189,8 243,2
165,9 165,9 134 173 169, 189, 254 168 255

(Nguồn: Tổng cục Thống kê)

a) Xác định số trung bình và tứ phân vị của mẫu số liệu trên.

b) Hoàn thành bảng tần số ghép nhóm theo mẫu sau:

Tổng lượng mưa trong tháng 8 (mm)	[120; 175)	[175; 230)	[230; 285)	[285; 340)
So năm	?	?	?	?

c) Hãy ước lượng số trung bình và một của mẫu số liệu ở bảng tần số ghép nhóm trên.

Lời giải

a)

Số trung bình:

$$\bar{x} = \frac{121,8 + 158,3 + 334,9 + 200,9 + 165,6 + 161,5 + 194,3 + 220,7 + 189,8 + \dots + 255}{19} \approx 19288$$

Tứ phân vị:

Xếp mẫu số liệu không giảm ta được:

121,8 134 158,3 161,5 165,6 165,9 165,9 168 169 173
189 189,8 194,3 200,9 202,7 243,2 254 255 334,9

Từ đó ta có:

Tứ phân vị thứ nhất của mẫu số liệu là: 165,6.

Tứ phân vị thứ hai của mẫu số liệu là: 173.

Tứ phân vị thứ ba của mẫu số liệu là: 202,7.

b) Hoàn thành bảng tần số ghép nhóm theo mẫu sau:

Tổng lượng mưa trong tháng 8 (mm)	[120; 175)	[175; 230)	[230; 285)	[285; 340)
Số năm	10	5	3	1

Giá trị đại diện của các lớp:

$$c_1 = \frac{120 + 175}{2} = 147,5; \quad c_2 = \frac{175 + 230}{2} = 202,5$$

$$c_3 = \frac{230 + 285}{2} = 257,5; \quad c_4 = \frac{285 + 340}{2} = 312,5$$

Tần số các lớp: $n_1 = 10; n_2 = 5; n_3 = 3; n_4 = 1$

$$\text{Số trung bình: } \bar{x} = \frac{n_1 c_1 + n_2 c_2 + n_3 c_3 + n_4 c_4}{n_1 + n_2 + n_3 + n_4} = \frac{7145}{38} \approx 188,02$$

Mốt

Mốt M_0 chứa trong nhóm $[120; 175)$

Do đó: $u_m = 120; u_{m+1} = 175 \Rightarrow u_{m+1} - u_m = 175 - 120 = 55$

$$n_{m-1} = 0; n_m = 10; n_{m+1} = 5$$

$$M_0 = 120 + \frac{10 - 0}{(10 - 0) + (10 - 5)} (175 - 120) = \frac{470}{3} = 156,67$$

Câu 10: Bảng sau thống kê số ca nhiễm mới SARS-CoV-2 mỗi ngày trong tháng 12/2021 tại Việt Nam.

Ngày	Số ca	Ngày	Số ca	Ngày	Số ca	Ngày	Số ca
1	15 139	9	15 965	17	15 685	25	16 046
2	14 295	10	15 474	18	16 363	26	15 667
3	14 254	11	16 830	19	16 586	27	15 310
4	14 598	12	15 264	20	15 420	28	14 866
5	14 927	13	16 035	21	16 806	29	14 299
6	15 215	14	15 871	22	17 044	30	20 454
7	14 433	15	16 192	23	16 860	31	17 004
8	15 223	16	15 720	24	16 633		

(Nguồn: worldmeters.info)

a) Xác định số trung bình và tứ phân vị của mẫu số liệu trên.

b) Hoàn thành bảng tần số ghép nhóm theo mẫu sau:

So ca (nghìn)	[14; 15,5)	[15,5; 17)	[17; 18,5)	[18,5; 20)	[20; 21,5)
So ngày	?	?	?	?	?

c) Hãy ước lượng số trung bình và mốt của mẫu số liệu ở bảng tần số ghép nhóm trên.

Lời giải

a)

$$\text{Số trung bình: } \bar{x} = \frac{14254 + 14295 + \dots + 20454 + 17004}{31} \approx 15821.$$

Tứ phân vị:

Xếp mẫu số liệu không giảm ta được:

14254 14295 14299 14433 14598 14866 14927 15139 15215 15223
 15264 15310 15420 15474 15668 15685 15720 15871 15965 16035
 16046 16192 16363 16586 16633 16806 16830 16860 17004 17044
 20454

Từ đó ta có:

Tứ phân vị thứ nhất của mẫu số liệu là: 15139.

Tứ phân vị thứ hai của mẫu số liệu là: 15685.

Tứ phân vị thứ ba của mẫu số liệu là: 16586.

b) Hoàn thành bảng tần số ghép nhóm theo mẫu sau:

Số ca (nghìn)	[14; 15,5)	[15,5; 17)	[17; 18,5)	[18,5; 20)	[20; 21,5)
Số ngày	13	15	2	0	1

c) Hãy ước lượng số trung bình và một của mẫu số liệu ở bảng tần số ghép nhóm trên.

Số trung bình: $\bar{x} = \frac{14,75.13 + 16,25.15 + 17,75.2 + 19,25.0 + 20,75.1}{31} = \frac{1967}{124} \approx 15,86.$

Mốt

Mốt M_0 chứa trong nhóm $[15,5; 17)$

Do đó: $u_m = 15,5; u_{m+1} = 17 \Rightarrow u_{m+1} - u_m = 17 - 15,5 = 1,5$

$n_{m-1} = 13; n_m = 15; n_{m+1} = 2$

$M_0 = 15,5 + \frac{15-13}{(15-13)+(15-2)}(17-15,5) = \frac{157}{10} = 15,7$



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

BÀI 2: TRUNG VỊ VÀ TỬ PHÂN VỊ CỦA MẪU SỐ LIỆU GHÉP NHÓM



LÝ THUYẾT.

1. TRUNG VỊ

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

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

Khi đó

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

Ví dụ 1. 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

Tính trung vị của mẫu số liệu ghép nhóm này.

Lời giải

Cỡ mẫu là $n = 3 + 12 + 15 + 24 + 2 = 56$.

Gọi $x_1, \dots, x_{56}$ là thời gian vào internet của 56 học sinh và giả sử dãy này được sắp xếp theo thứ

tự tăng dần. Khi đó, trung vị là $\frac{x_{28} + x_{29}}{2}$. Do 2 giá trị x_{28}, x_{29} thuộc nhóm $[15, 5; 18, 5)$ nên

nhóm này chứa trung vị. Do đó, $p = 3; a_3 = 15, 5; m_3 = 15; m_1 + m_2 = 3 + 12 = 15; a_4 - a_3 = 3$ và ta

$$\text{có } M_e = 15, 5 + \frac{\frac{56}{2} - 15}{15} \cdot 3 = 18, 1.$$

Ví dụ 2. Ghi lại tốc độ bóng trong 200 lần giao bóng của một vận động viên môn quần vợt cho kết quả như bảng sau:

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

Tính trung vị của mẫu số liệu ghép nhóm này.

Lời giải

Cỡ mẫu là $n = 200$.

Gọi $x_1, \dots, x_{200}$ là tốc độ giao bóng của 200 lần giao bóng và giả sử dãy này được sắp xếp theo thứ tự tăng dần. Khi đó, trung vị là $\frac{x_{100} + x_{101}}{2}$. Do 2 giá trị x_{100}, x_{101} thuộc nhóm $[165; 170)$ nên nhóm này chứa trung vị. Do đó,

$p = 4; a_4 = 165; m_3 = 43; m_1 + m_2 + m_3 = 18 + 28 + 35 = 81; a_5 - a_4 = 5$ và ta có

$$M_e = 165 + \frac{\frac{200}{2} - 81}{43} \cdot 5 = 167,21.$$

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.

2. 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 là 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;

$$\bullet \quad C = n_1 + n_2 + \dots + n_{j-1}.$$

Khi đó

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

Ví dụ 1. Tìm tứ phân vị thứ nhất Q_1 và tứ phân vị thứ ba Q_3 của mẫu số liệu ghép nhóm cho trong Ví dụ 2.

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

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

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

$$Q_3 = 18, 5 + \frac{\frac{3 \cdot 56}{4} - 30}{24} \cdot 3 = 20.$$

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

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.



HỆ THỐNG BÀI TẬP TỰ LUẬN.

Câu 1: Thống kê điểm trung bình môn Toán của một số học sinh lớp 11 được cho ở bảng sau:

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

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

Câu 2: Tổng số lượng mưa trong tháng 8 đo được tại một trạm quan trắc đặt tại Vũng Tàu từ năm 2002 đến năm 2020 được ghi lại như dưới đây (đơn vị: mm)

121,8 158,3 334,9 200,9 165,6 161,5 194,3 220,7 189,8 243,2
165,9 165,9 134 173 169, 189, 254 168 255

(Nguồn: Tổng cục Thống kê)

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

Câu 3: Bảng sau thống kê số ca nhiễm mới SARS-CoV-2 mỗi ngày trong tháng 12/2021 tại Việt Nam.

Ngày	Số ca	Ngày	Số ca	Ngày	Số ca	Ngày	Số ca
1	15 139	9	15 965	17	15 685	25	16 046
2	14 295	10	15 474	18	16 363	26	15 667
3	14 254	11	16 830	19	16 586	27	15 310
4	14 598	12	15 264	20	15 420	28	14 866
5	14 927	13	16 035	21	16 806	29	14 299
6	15 215	14	15 871	22	17 044	30	20 454
7	14 433	15	16 192	23	16 860	31	17 004
8	15 223	16	15 720	24	16 633		

(Nguồn: worldmeters.info)

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



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

BÀI 2: TRUNG VỊ VÀ TỬ PHÂN VỊ CỦA MẪU SỐ LIỆU GHÉP NHÓM



LÝ THUYẾT.

1. TRUNG VỊ

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

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

Khi đó

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

Ví dụ 1. 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

Tính trung vị của mẫu số liệu ghép nhóm này.

Lời giải

Cỡ mẫu là $n = 3 + 12 + 15 + 24 + 2 = 56$.

Gọi $x_1, \dots, x_{56}$ là thời gian vào internet của 56 học sinh và giả sử dãy này được sắp xếp theo thứ

tự tăng dần. Khi đó, trung vị là $\frac{x_{28} + x_{29}}{2}$. Do 2 giá trị x_{28}, x_{29} thuộc nhóm $[15, 5; 18, 5)$ nên

nhóm này chứa trung vị. Do đó, $p = 3; a_3 = 15, 5; m_3 = 15; m_1 + m_2 = 3 + 12 = 15; a_4 - a_3 = 3$ và ta

$$\text{có } M_e = 15, 5 + \frac{\frac{56}{2} - 15}{15} \cdot 3 = 18, 1.$$

Ví dụ 2. Ghi lại tốc độ bóng trong 200 lần giao bóng của một vận động viên môn quần vợt cho kết quả như bảng sau:

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

Tính trung vị của mẫu số liệu ghép nhóm này.

Lời giải

Cỡ mẫu là $n = 200$.

Gọi $x_1, \dots, x_{200}$ là tốc độ giao bóng của 200 lần giao bóng và giả sử dãy này được sắp xếp theo thứ tự tăng dần. Khi đó, trung vị là $\frac{x_{100} + x_{101}}{2}$. Do 2 giá trị x_{100}, x_{101} thuộc nhóm $[165; 170)$ nên nhóm này chứa trung vị. Do đó,

$p = 4; a_4 = 165; m_3 = 43; m_1 + m_2 + m_3 = 18 + 28 + 35 = 81; a_5 - a_4 = 5$ và ta có

$$M_e = 165 + \frac{\frac{200}{2} - 81}{43} \cdot 5 = 167,21.$$

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.

2. 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 là 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;

$$\bullet \quad C = n_1 + n_2 + \dots + n_{j-1}.$$

Khi đó

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

Ví dụ 1. Tìm tứ phân vị thứ nhất Q_1 và tứ phân vị thứ ba Q_3 của mẫu số liệu ghép nhóm cho trong Ví dụ 2.

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

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

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

$$Q_3 = 18, 5 + \frac{\frac{3 \cdot 56}{4} - 30}{24} \cdot 3 = 20.$$

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

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.

HỆ THỐNG BÀI TẬP TỰ LUẬN.

Câu 1: Thống kê điểm trung bình môn Toán của một số học sinh lớp 11 được cho ở bảng sau:

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

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

Lời giải

Tứ phân vị thứ hai. Nhóm [8; 8,5)

$$Q_2 = 8 + \frac{\frac{2.82}{4} - (8+10+16)}{24} (8,5 - 8) \approx 8,15$$

Tứ phân vị thứ nhất. Nhóm [7,5; 8)

$$Q_1 = 7,5 + \frac{\frac{2.82}{4} - (8+10)}{16} (8 - 7,5) \approx 7,58$$

Tứ phân vị thứ ba. Nhóm [8,5; 9)

$$Q_3 = 8,5 + \frac{\frac{3.82}{4} - (8+10+16+24)}{16} (9 - 8,5) \approx 8,63$$

Câu 2: Tổng số lượng mưa trong tháng 8 đo được tại một trạm quan trắc đặt tại Vũng Tàu từ năm 2002 đến năm 2020 được ghi lại như dưới đây (đơn vị: mm)

121,8	158,3	334,9	200,9	165,6	161,5	194,3	220,7	189,8	243,2
165,9	165,9	134	173	169,	189,	254	168	255	

(Nguồn: Tổng cục Thống kê)

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

Lời giải

Tứ phân vị thứ nhất. Nhóm [120; 175)

$$Q_1 = 120 + \frac{\frac{1.19}{4} - 0}{10} (175 - 120) = \frac{1169}{8} \approx 146,125$$

Tứ phân vị thứ hai. Nhóm [175; 230)

$$Q_2 = 175 + \frac{\frac{2.19}{4} - (0+10)}{5} (230 - 175) = \frac{339}{2} \approx 169,5$$

Tứ phân vị thứ ba. Nhóm [230; 285)

$$Q_3 = 230 + \frac{\frac{3.19}{4} - (0+10+5)}{3} (285 - 230) = \frac{865}{4} \approx 216,25$$

Câu 3: Bảng sau thống kê số ca nhiễm mới SARS-CoV-2 mỗi ngày trong tháng 12/2021 tại Việt Nam.

Ngày	Số ca	Ngày	Số ca	Ngày	Số ca	Ngày	Số ca
1	15 139	9	15 965	17	15 685	25	16 046
2	14 295	10	15 474	18	16 363	26	15 667
3	14 254	11	16 830	19	16 586	27	15 310
4	14 598	12	15 264	20	15 420	28	14 866
5	14 927	13	16 035	21	16 806	29	14 299
6	15 215	14	15 871	22	17 044	30	20 454
7	14 433	15	16 192	23	16 860	31	17 004
8	15 223	16	15 720	24	16 633		

(Nguồn: worldmeters.info)

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

Lời giải

Tứ phân vị thứ nhất. Nhóm [14;15,5)

$$Q_1 = 14 + \frac{\frac{1.31}{4} - 0}{13} (15,5 - 14) = \frac{1549}{104} \approx 14,89$$

Tứ phân vị thứ hai. Nhóm [15,5;17)

$$Q_2 = 15,5 + \frac{\frac{2.31}{4} - (0+13)}{15} (17 - 15,5) = \frac{63}{4} \approx 15,75$$

Tứ phân vị thứ ba. Nhóm [17;18,5)

$$Q_3 = 15,5 + \frac{\frac{3.31}{4} - (0+13+15)}{2} (18,5 - 17) = \frac{215}{16} \approx 13,44$$

CHƯƠNG



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



HỆ THỐNG BÀI TẬP TRẮC NGHIỆM.

Câu 1: Điều tra về chiều cao của học sinh khối lớp 11, ta được mẫu số liệu sau:

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

Mẫu số liệu ghép nhóm đã cho có tất cả bao nhiêu nhóm?

- A. 5. B. 6. C. 7. D. 12.

Câu 2: Điều tra về chiều cao của học sinh khối lớp 11, 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
		$N = 100$

Giá trị đại diện của nhóm thứ tư là

- A. 156,5. B. 157. C. 157,5. D. 158.

**CHUYÊN ĐỀ V – TOÁN – 11 – CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CỦA MẪU SỐ
LIỆU GHEP NHÓM**

Câu 3: Đo chiều cao (tính bằng cm) của 500 học sinh trong một trường THPT ta thu được kết quả như sau:

Chiều cao	[150;154)	[154;158)	[158;162)	[162;166)	[166;170)
Số học sinh	25	50	200	175	50

Mẫu số liệu ghép nhóm đã cho có tất cả bao nhiêu nhóm?

- A.** 5. **B.** 6. **C.** 7. **D.** 12.

Câu 4: Đo chiều cao (tính bằng cm) của 500 học sinh trong một trường THPT ta thu được kết quả như sau:

Chiều cao	[150;154)	[154;158)	[158;162)	[162;166)	[166;170)
Số học sinh	25	50	200	175	50

Giá trị đại diện của nhóm [162;166) là

- A.** 162. **B.** 164. **C.** 166. **D.** 4.

Câu 5: Đo cân nặng của một số học sinh lớp 11D cho trong bảng sau:

Cân nặng (kg)	[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

Giá trị đại diện của nhóm [60,5;65,5) là

- A.** 55,5. **B.** 58. **C.** 60,5. **D.** 5.

Câu 6: Tìm hiểu thời gia 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.

Câu 7: 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

Có bao nhiêu học sinh truy cập Internet mỗi buổi tối có thời gian từ 18,5 phút đến dưới 21,5 phút?

- A.** 24. **B.** 15. **C.** 2. **D.** 20.

**CHUYÊN ĐỀ V – TOÁN – 11 – CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CỦA MẪU SỐ
LIỆU GHÉP NHÓM**

Câu 8: 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.

Câu 9: 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 một của mẫu số liệu trên là

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

Câu 10: 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 11: 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 trung vị của mẫu số liệu trên là

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

Câu 12: 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 13: 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 14: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một cửa hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

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

Số trung bình của mẫu số liệu trên thuộc khoảng nào trong các khoảng dưới đây?

- A.** [7; 9). **B.** [9; 11). **C.** [11; 13). **D.** [13; 15).

Câu 15: 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

Trung vị 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 16: 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

Mốt 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 17: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một cửa hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

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

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

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

Câu 18: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một cửa hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

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

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

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



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



HỆ THỐNG BÀI TẬP TRẮC NGHIỆM.

Câu 1: Điều tra về chiều cao của học sinh khối lớp 11, ta được mẫu số liệu sau:

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

Mẫu số liệu ghép nhóm đã cho có tất cả bao nhiêu nhóm?

- A. 5. B. 6. C. 7. D. 12.

Lời giải

Mẫu số liệu ghép nhóm đã cho có tất cả 6 nhóm.

Câu 2: Điều tra về chiều cao của học sinh khối lớp 11, 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
		$N = 100$

Giá trị đại diện của nhóm thứ tư là

- A. 156,5. B. 157. C. 157,5. D. 158.

Lời giải

Giá trị đại diện của nhóm thứ tư là $\frac{156+158}{2}=157$.

Câu 3: Đo chiều cao (tính bằng cm) của 500 học sinh trong một trường THPT ta thu được kết quả như sau:

Chiều cao	[150;154)	[154;158)	[158;162)	[162;166)	[166;170)
Số học sinh	25	50	200	175	50

Mẫu số liệu ghép nhóm đã cho có tất cả bao nhiêu nhóm?

- A.** 5. **B.** 6. **C.** 7. **D.** 12.

Lời giải

Mẫu số liệu ghép nhóm đã cho có tất cả 5 nhóm.

Câu 4: Đo chiều cao (tính bằng cm) của 500 học sinh trong một trường THPT ta thu được kết quả như sau:

Chiều cao	[150;154)	[154;158)	[158;162)	[162;166)	[166;170)
Số học sinh	25	50	200	175	50

Giá trị đại diện của nhóm $[162;166)$ là

- A.** 162. **B.** 164. **C.** 166. **D.** 4.

Lời giải

Ta có bảng sau

Lớp chiều cao	Giá trị đại diện	Số học sinh
[150;154)	152	25
[154;158)	156	50
[158;162)	160	200
[162;166)	164	175
[166;170)	168	50

Câu 5: Đo cân nặng của một số học sinh lớp 11D cho trong bảng sau:

Cân nặng (kg)	[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

Giá trị đại diện của nhóm $[60, 5; 65, 5)$ là

- A.** 55,5. **B.** 58. **C.** 60,5. **D.** 5.

Lời giải

CHUYÊN ĐỀ V – TOÁN – 11 – CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CỦA MẪU SỐ LIỆU GHEP NHÓM

Trong mỗi khoảng cân ặng, giá trị đại diện trung bình cộng của giá trị hai đầu mút nên ta có bảng sau:

Cân nặng (kg)	[40,5; 45,5)	[45,5; 50,5)	[50,5; 55,5)	[55,5; 60,5)	[60,5; 65,5)	[65,5; 70,5)
Giá trị đại diện	43	48	53	58	63	68
Số học sinh	10	7	16	4	2	3

Câu 6: Tìm hiểu thời gia 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 7: 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

Có bao nhiêu học sinh truy cập Internet mỗi buổi tối có thời gian từ 18,5 phút đến dưới 21,5 phút?

A. 24.

B. 15.

C. 2.

D. 20.

Lời giải

Số học sinh truy cập Internet mỗi buổi tối có thời gian từ 18,5 phút đến dưới 21,5 phút là 24.

Câu 8: 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 9: 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 một của mẫu số liệu trên là

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

Lời giải

Mốt M_0 chứa trong nhóm [40; 60)

Câu 10: 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: $u_m = 40; u_{m+1} = 60 \Rightarrow u_{m+1} - u_m = 60 - 40 = 20$

$n_{m-1} = 9; n_m = 12; n_{m+1} = 10$

$$M_0 = 40 + \frac{12 - 9}{(12 - 9) + (12 - 10)}(60 - 40) = 52$$

Câu 11: 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 trung vị 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 trung vị của mẫu số liệu trên là $Q_2 = \frac{x_{21} + x_{22}}{2}$

Mà $x_{21}, x_{22} \in [40; 60)$

Vậy nhóm chứa trung vị của mẫu số liệu trên là nhóm [40; 60)

Câu 12: 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 13: 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 14: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một cửa hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

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

Số trung bình của mẫu số liệu trên thuộc khoảng nào trong các khoảng dưới đây?

A. $[7; 9)$. **B.** $[9; 11)$. **C.** $[11; 13)$. **D.** $[13; 15)$.

Lời giải

Bảng tần số ghép nhóm theo giá trị đại diện là

Doanh thu	$[5; 7)$	$[7; 9)$	$[9; 11)$	$[11; 13)$	$[13; 15)$
Giá trị đại diện	6	8	10	12	14
Số ngày	2	7	7	3	1

Số trung bình: $\bar{x} = \frac{2.6 + 7.8 + 7.10 + 3.12 + 1.14}{20} = 9,4$

Câu 15: 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

Trung vị 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

Goi $x_1, x_2, \dots, x_{20}$ là doanh thu bán hàng trong 20 ngày xếp theo thứ tự không giảm.

Khi đó: $x_1, x_2 \in [5; 7)$, $x_3, \dots, x_9 \in [7; 9)$, $x_{10}, \dots, x_{16} \in [9; 11)$, $x_{17}, \dots, x_{19} \in [11; 13)$, $x_{20} \in [13; 15)$

Do đó, trung vị của mẫu số liệu thuộc nhóm $[9; 11)$

Câu 16: 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

Một cửa 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

Có 2 nhóm chứa một cửa mẫu số liệu trên đó là $[7; 9)$ và $[9; 11)$, do đó:

Xét nhóm $[7; 9)$ ta có:

$$M_0 = 7 + \frac{7-2}{(7-2) + (7-7)}(9-7) = 9$$

Xét nhóm $[9; 11)$ ta có:

$$M'_0 = 9 + \frac{7-7}{(7-7) + (7-3)}(11-9) = 9$$

Vậy một cửa mẫu số liệu là 9.

Câu 17: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một cửa hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

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

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

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

Lời giải

Goi $x_1, x_2, \dots, x_{20}$ là doanh thu bán hàng trong 20 ngày xếp theo thứ tự không giảm.

Khi đó: $x_1, x_2 \in [5; 7)$, $x_3, \dots, x_9 \in [7; 9)$, $x_{10}, \dots, x_{16} \in [9; 11)$, $x_{17}, \dots, x_{19} \in [11; 13)$, $x_{20} \in [13; 15)$

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

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

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

A. 10.

B. 11.

C. 12.

D. 13.

Lời giải

Gọi $x_1, x_2, \dots, x_{20}$ là doanh thu bán hàng trong 20 ngày xếp theo thứ tự không giảm.

Khi đó: $x_1, x_2 \in [5; 7)$, $x_3, \dots, x_9 \in [7; 9)$, $x_{10}, \dots, x_{16} \in [9; 11)$, $x_{17}, \dots, x_{19} \in [11; 13)$, $x_{20} \in [13; 15)$

Do đó, tứ phân vị thứ ba của mẫu số liệu thuộc nhóm $[9; 11)$

$$n = 20, n_m = 7, C = 9, u_m = 9, u_{m+1} = 11$$

$$Q_3 = 9 + \frac{\frac{3.20}{4} - 9}{7} (11 - 9) \approx 10,71 \approx 11$$