

CHUYÊN ĐỀ 12. CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM MẪU SỐ LIỆU GHÉP NHÓM

A. KIẾN THỨC CƠ BẢN CẦN NẮM

1. Mẫu số liệu ghép nhóm

a) Bảng tần số ghép nhóm

- Mẫu số liệu ghép nhóm là mẫu số liệu cho dưới dạng bảng tần số của các nhóm số liệu.
- Mỗi nhóm số liệu là tập hợp gồm các giá trị của số liệu được ghép nhóm theo một tiêu chí xác định dưới dạng $[a; b)$, trong đó a là đầu mút trái, b là đầu mút phải.
- Hiệu $b - a$ được gọi là độ dài nhóm $[a; b)$.
- Tần số của một nhóm là số số liệu trong mẫu số liệu thuộc vào nhóm đó. Tần số nhóm 1, nhóm 2, nhóm 3, ..., nhóm k được kí hiệu lần lượt là $n_1, n_2, n_3, \dots, n_k$.
- Bảng tần số ghép nhóm được trình bày dưới dạng bảng thống kê như bảng 1, trong đó mẫu số liệu gồm n số liệu được chia thành k nhóm tương ứng với k nửa khoảng $[a_1; a_2), [a_2; a_3), [a_3; a_4), \dots, [a_k; a_{k+1})$, với $a_1 < a_2 < \dots < a_k < a_{k+1}$ và $n = n_1 + n_2 + n_3 + \dots + n_k$.

Bảng 1: Bảng tần số ghép nhóm

Nhóm	$[a_1; a_2)$	$[a_2; a_3)$	$[a_3; a_4)$	...	$[a_k; a_{k+1})$
Tần số	n_1	n_2	n_3	...	n_k

b) Ghép nhóm mẫu số liệu, tần số tích lũy

Để chuyển mẫu số liệu không ghép nhóm sang mẫu số liệu ghép nhóm, ta làm như sau:

- Chia miền giá trị của mẫu số liệu thành một số nhóm theo tiêu chí cho trước.
- Đếm số giá trị của mẫu số liệu thuộc mỗi nhóm (tần số) và lập bảng tần số ghép nhóm.

Lưu ý: Khi ghép nhóm số liệu, ta nên chia các nhóm có độ dài như nhau. Đầu mút của nhóm có thể không phải là giá trị của mẫu số liệu. Nhóm cuối cùng có thể là $[a_k; a_{k+1}]$.

Tần số tích lũy của mỗi nhóm bằng tần số của nhóm đó cộng với tần số của nhóm phía trước. Bảng tần số tích lũy được trình bày như Bảng 2.

Bảng 2: Bảng tần số tích lũy

Nhóm	Tần số	Tần số tích lũy
$[a_1; a_2)$	n_1	n_1
$[a_2; a_3)$	n_2	$n_1 + n_2$
$[a_3; a_4)$	n_3	$n_1 + n_2 + n_3$
...	...	...
$[a_k; a_{k+1})$	n_k	$n_1 + n_2 + n_3 + \dots + n_k$

2. Các số đặt trưng đo xu thế trung tâm cho mẫu số liệu ghép nhóm

a) Số trung bình của mẫu số liệu ghép nhóm

Cho mẫu số liệu ghép nhóm như bảng 3, trong đó giá trị trung điểm x_i của nửa khoảng (tính bằng trung bình cộng của hai đầu mút) ứng với nhóm i .

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

$$\bar{x} = \frac{n_1 \cdot x_1 + n_2 \cdot x_2 + n_3 \cdot x_3 + \dots + n_k \cdot x_k}{n}$$

Trong đó $n = n_1 + n_2 + n_3 + \dots + n_k$.

Bảng 3

Nhóm	Giá trị đại diện	Tần số
$[a_1; a_2)$	x_1	n_1
$[a_2; a_3)$	x_2	n_2
$[a_3; a_4)$	x_3	n_3
$\dots$	$\dots$	$\dots$
$[a_k; a_{k+1})$	x_k	n_k
		$n = n_1 + n_2 + n_3 + \dots + n_k$

Ý nghĩa của số trung bình của mẫu số liệu ghép nhóm là giá trị trung bình xấp xỉ của các số liệu trong mẫu. Nó có thể được xem như giá trị trung tâm của mẫu và đôi khi được dùng làm một đại diện cho mẫu.

b) Trung vị của mẫu số liệu ghép nhóm

- Nhóm chứa trung vị của mẫu số liệu ghép nhóm là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng $\frac{n}{2}$, trong đó n là cỡ mẫu.

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;
- $C = n_1 + n_2 + \dots + n_{m-1}$.

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

Khi đó

Ý nghĩa của trung vị của mẫu số liệu ghép nhóm: là giá trị xấp xỉ cho trung vị của mẫu số liệu và có thể sử dụng làm giá trị đại diện cho mẫu số liệu.

c) Tứ phân vị của mẫu số liệu ghép nhóm

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;

$$C = n_1 + n_2 + \dots + n_{m-1}$$

•

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

Khi đó

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

$$[u_j; u_{j+1})$$

- Giả sử nhóm

chứa tứ phân vị thứ ba;

$$n_j$$

- là tần số của nhóm chứa tứ phân vị thứ ba;

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

•

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

Khi đó:

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

- Tứ phân vị của mẫu số liệu ghép nhóm là giá trị xấp xỉ của tứ phân vị của mẫu số liệu.
- Các tứ phân vị Q_1, Q_2, Q_3 chia mẫu số liệu ghép nhóm thành bốn phần có số số liệu bằng nhau. Các tứ phân vị cho ta một hình ảnh về sự phân bố của mẫu số liệu. Dựa vào các tứ phân vị, ta có thể biết số liệu tập trung ít hay nhiều quanh trung vị.

d) Một của mẫu số liệu ghép nhóm

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

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

định bởi công thức

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

Ý nghĩa của một của mẫu số liệu ghép nhóm

Một của một mẫu số liệu ghép nhóm cho biết rằng những giá trị xấp xỉ với một xuất hiện nhiều nhất trong mẫu. Nó cũng thể hiện xu thế tập trung của mẫu số liệu.

B. BÀI TẬP VẬN DỤNG

BÀI TOÁN 1: BẢNG SỐ LIỆU GHÉP NHÓM – TRUNG BÌNH CỘNG

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) Thông kê nhiệt độ tại một địa 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 đều là mẫu số liệu ghép nhóm

a) Có 5 sinh viên chi dưới 50 nghìn đồng cho việc thanh toán cước điện thoại trong tháng

Có 12 sinh viên chi từ 50 đến dưới 100 nghìn đồng cho việc thanh toán cước điện thoại trong tháng

Có 23 sinh viên chi từ 100 đến dưới 150 nghìn đồng cho việc thanh toán cước điện thoại trong tháng

Có 17 sinh viên chi từ 150 đến dưới 200 nghìn đồng cho việc thanh toán cước điện thoại trong tháng

Có 3 sinh viên chi từ 200 đến dưới 250 nghìn đồng cho việc thanh toán cước điện thoại trong tháng

Như vậy, đa số sinh viên chi từ 100 đến dưới 150 nghìn đồng mỗi tháng cho cước điện thoại và có ít sinh viên chi trên 200 nghìn đồng cho cước điện thoại mỗi tháng

b) Có 7 ngày có nhiệt độ từ 19°C đến dưới 22°C

Có 15 ngày có nhiệt độ từ 22°C đến dưới 25°C

Có 12 ngày có nhiệt độ từ 25°C đến dưới 28°C

Có 6 ngày có nhiệt độ từ 28°C đến dưới 31°C

Câu 2: 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 sang dạng ghép nhóm với sáu nhóm có độ dài bằng nhau.

Lời giải

Giá trị nhỏ nhất là: 5

Giá trị lớn nhất là 54

Do đó khoảng biến thiên là $54 - 5 = 49$

Để chia thành 6 nhóm với độ dài bằng nhau ta lấy điểm đầu mút phải trái của nhóm đầu tiên là 3 và đầu mút phải của nhóm cuối cùng là 57 với độ dài mỗi nhóm là 9.

Ta được mẫu số liệu ghép nhóm như sau:

Số sản phẩm	[3;12)	[12;21)	[21;30)	[30;39)	[39;48)	[48;57)
Số công nhân	2	8	8	2	4	1

Câu 3: 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.

(Theo: <https://www.premierleague.com/>)

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

Lời giải

Giá trị lớn nhất là: 653

Giá trị bé nhất là: 492

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

Để chia thành 7 nhóm có độ dài bằng nhau, ta lấy điểm đầu mút trái của nhóm đầu tiên là 492, điểm đầu mút phải của nhóm cuối là 653 với độ dài mỗi nhóm là 23

Ta có mẫu số liệu ghép nhóm sau:

Thời gian ra sân (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 4: Quãng đường (km) từ nhà đến nơi làm việc của 40 công nhân một nhà máy được ghi lại như sau:

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

a) Ghép nhóm dãy số liệu trên thành các khoảng có độ rộng bằng nhau, khoảng đầu tiên là $[0;5)$. Tìm giá trị đại diện cho mỗi nhóm.

b) Tính số trung bình của mẫu số liệu không ghép nhóm và mẫu số liệu ghép nhóm. Giá trị nào chính xác hơn?

c) Xác định nhóm chứa một của mẫu số liệu ghép nhóm thu được.

Lời giải

a)

Quãng đường (km)	[0;5)	[5;10)	[10;15)	[15;20)	[20;25)	[25;30)	[30;35)
Số công nhân	6	10	11	9	1	1	2
Giá trị đại diện	2.5	7.5	12.5	17.5	22.5	27.5	32.5

b) Với mẫu số liệu không ghép nhóm:

$$\bar{x} = (5 + 3 + 10 + 20 + 25 + 11 + 13 + 7 + 12 + 31 + 19 + 10 + 12 + 17 + 18 + 11 + 32 + 17 + 16 + 2 + 7 + 9 + 7 + 8 + 3 + 5 + 12 + 15 + 18 + 3 + 12 + 14 + 2 + 9 + 6 + 15 + 15 + 7 + 6 + 12) : 40 = 11.9$$

Với mẫu số liệu ghép nhóm:

$$\bar{x} = \frac{2.5 \times 6 + 7.5 \times 10 + 12.5 \times 11 + 17.5 \times 9 + 22.5 + 27.5 + 32.5 \times 2}{40} = 12.5$$

Số trung bình của mẫu số liệu không ghép nhóm chính xác hơn

c) 11 là tần số lớn nhất nên nhóm chứa một là [10;15)

Câu 5: Thống kê chỉ số chất lượng không khí (AQI) tại một địa điểm vào các ngày trong tháng 6/2022 được cho trong bảng sau:

Chỉ số AQI	[0;50)	[50;100)	[100;150)	[150;200)	Trên 200
Số ngày	5	11	7	4	3

a) Đọc và giải thích mẫu số liệu ghép nhóm.

b) Chất lượng không khí được xem là tốt nếu AQI nhỏ hơn 50, là trung bình nếu AQI từ 50 đến dưới 100. Trong tháng 6/2022 tại địa điểm này có bao nhiêu ngày chất lượng không khí dưới mức trung bình?

Lời giải

a) Trong tháng 6/2022 có 5 ngày chỉ số AQI dưới 50; 11 ngày chỉ số AQI từ 50 đến 100; 7 ngày chỉ số AQI từ 100 đến dưới 150; 4 ngày chỉ số AQI từ 150 đến dưới 200; 3 ngày chỉ số AQI trên 200.

b) Số ngày chất lượng không khí dưới mức trung bình là: $7 + 4 + 3 = 14$.

Câu 6: Trẻ sơ sinh được xem là nhẹ cân nếu cân nặng khi sinh dưới 2 kg, là thừa cân nếu cân nặng khi sinh trên 4 kg, là cân nặng trung bình nếu cân nặng khi sinh từ 2 kg đến 4 kg. Thống kê cân nặng (tính theo kg) của 15 trẻ sơ sinh tại một bệnh viện cho kết quả như sau:

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

a) Tìm số trẻ nhẹ cân, thừa cân, có cân nặng trung bình trong 15 trẻ sơ sinh trên.

b) Xây dựng mẫu số liệu ghép nhóm cho mẫu số liệu trên.

Lời giải

1,12,2

a) Số trẻ nhẹ cân, cân nặng trung bình, thừa cân tương ứng là .

b) Mẫu số liệu ghép nhóm:

Cân nặng	Dưới 2	[2; 4]	Trên 4
Số trẻ	1	12	2

Câu 7: Thời gian hoàn thành bài kiểm tra Toán 45 phút của các bạn trong lớp được cho như sau:

Thời gian (phút)	[25; 30)	[30; 35)	[35; 40)	[40; 45)
Số học sinh	2	7	10	25

a) Nêu các nhóm số liệu và tần số tương ứng.

b) Có bao nhiêu học sinh hoàn thành bài kiểm tra trước khi hết giờ trên 5 phút?

Lời giải

a) Các nhóm số liệu [25;30), [30;35), [35;40), [40; 45] với tần số tương ứng là 2, 7, 10, 25 .

b) Số học sinh hoàn thành bài kiểm tra trước khi hết giờ ít nhất 5 phút là $2 + 7 + 10 = 19$ **Câu 8:** Cơ cấu dân số Việt Nam năm 2020 theo độ tuổi được cho trong bảng sau:

Độ tuổi	Dưới 5 tuổi	5 - 14	15 - 24	25 - 64	Trên 65
Số người (triệu)	7,89	14,68	13,32	53,78	7,66

(Theo: <http://ourworldindata.org>)

Chọn 80 là giá trị đại diện cho nhóm trên 65 tuổi. Tính tuổi trung bình của người Việt Nam năm 2020.

Lời giải

Độ tuổi	Dưới 5 tuổi	5-14	15 - 24	25 - 64	Trên 65
Số người (triệu)	7.89	14.68	13.32	53.78	7.66
Giá trị đại diện	2.5	9.5	19.5	44.5	80

Tuổi trung bình của người Việt Nam năm 2020:

$$\bar{x} = \frac{7.89 \times 2.5 + 14.68 \times 9.5 + 13.32 \times 19.5 + 53.78 \times 44.5 + 7.66 \times 80}{7.89 + 14.68 + 13.32 + 53.78 + 7.66} = 35$$

Câu 9: 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.

Lời giải

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

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:

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

Câu 10: 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

Câu 11: 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 $R = 63,6 - 45,1 = 18,5$

$$L > \frac{R}{k} = \frac{18,5}{5} = 3,7$$

Độ dài mỗi nhóm

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

Câu 12: 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 $R = 29 - 10 = 19$

$$L > \frac{R}{k} = \frac{19}{5} = 3,8$$

Độ dài mỗi nhóm

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

Câu 13: 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 $n_1c_1 + n_2c_2 + n_3c_3 + n_4c_4 + n_5c_5$

$$\bar{x} = \frac{n_1c_1 + n_2c_2 + n_3c_3 + n_4c_4 + n_5c_5}{40}$$

c) Tính

Lời giải

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

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

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

Câu 14: 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) 0)	[160;165))	[165;170) 0)	[170;175) 5)
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 .

Câu 15: 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 $R = 63,6 - 45,1 = 18,5$

$$L > \frac{R}{k} = \frac{18,5}{5} = 3,7$$

Độ dài mỗi nhóm

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:

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

CHUYÊN ĐỀ 2. MỐT- TRUNG VỊ VÀ TỨ PHÂN VỊ

Câu 16: Phỏng vấn một số học sinh khối 11 về thời gian (giờ) ngủ của một buổi tối, thu được bảng số liệu ở bên.

a) So sánh thời gian ngủ trung bình của các bạn học sinh nam và nữ.

b) Hãy cho biết 75% học sinh khối 11 ngủ ít nhất bao nhiêu giờ?

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

Lời giải

a)

Thời gian	Giá trị đại diện	Số học sinh nam	Số học sinh nữ
[4;5)	4.5	6	4
[5;6)	5.5	10	8
[6;7)	6.5	13	10
[7;8)	7.5	9	11
[8;9)	8.5	7	8

Thời gian ngủ trung bình của các bạn nam

$$\bar{x}_{nam} = \frac{4.5 \times 6 + 5.5 \times 10 + 6.5 \times 13 + 7.5 \times 9 + 8.5 \times 7}{6 + 10 + 13 + 9 + 7} = 6.52$$

Thời gian ngủ trung bình của các bạn nữ

$$\bar{x}_{n\ddot{u}} = \frac{4.5 \times 4 + 5.5 \times 8 + 6.5 \times 10 + 7.5 \times 11 + 8.5 \times 8}{4 + 8 + 10 + 11 + 8} = 6.77$$

$6.77 > 6.52$. Như vậy thời gian ngủ trung bình của các bạn nữ nhiều hơn các bạn nam

b) Cỡ mẫu $n = 86$

Tứ phân vị thứ nhất Q_1 là $\frac{x_{21} + x_{22}}{2}$. Do x_{21}, x_{22} đều thuộc nhóm $[5;6)$ nên nhóm này chứa Q_1 .
Do đó, $p = 2, a_2 = 5, m_2 = 18; m_1 = 10, a_3 - a_2 = 1$ và ta có:

$$Q_1 = 5 + \frac{\frac{86}{4} - 10}{18} \times 1 = 5.64$$

Vậy 75% học sinh khối 11 ngủ ít nhất 5.64 giờ

Câu 17: Tuổi thọ (năm) của 50 bình ắc quy ô tô được cho như sau:

Tuổi thọ (năm)	[2; 2,5)	[2,5; 3)	[3; 3,5)	[3,5; 4)	[4; 4,5)	[4,5; 5)
Tần số	4	9	14	11	7	5

a) Xác định một và giải thích ý nghĩa.

b) Tính tuổi thọ trung bình của 50 bình ắc quy ô tô này.

Lời giải

a) 14 là tần số lớn nhất nên một thuộc nhóm $[3;3.5)$, ta có

$$j = 3, a_3 = 3, m_3 = 14, m_2 = 9, m_4 = 11, h = 0.5$$

$$M_o = 3 + \frac{14 - 9}{(14 - 9) + (14 - 11)} \times 0.5 = 3.31$$

Do đó:

b)

Tuổi thọ (năm)	[2;2.5)	[2.5;3)	[3;3.5)	[3.5;4)	[4;4.5)	[4.5;5)
Tần số	4	9	14	11	7	5
Giá trị đại diện	2.25	2.75	3.25	3.75	4.25	4.75

$$\bar{x} = \frac{4 \times 2.25 + 9 \times 2.75 + 14 \times 3.25 + 11 \times 3.75 + 7 \times 4.25 + 5 \times 4.75}{50} = 3.48$$

Tuổi thọ trung bình:

Câu 18: Điểm thi môn Toán (thang điểm 100, điểm được làm tròn đến 1) của 60 thí sinh được cho trong bảng sau:

Điểm	0 - 9	10 - 19	20 - 29	30 - 39	40 - 49
Số thí sinh	1	2	4	6	15
Điểm	50 - 59	60 - 69	70 - 79	80 - 89	90 - 99
Số thí sinh	12	10	6	3	1

a) Hiệu chỉnh để thu được mẫu số liệu ghép nhóm dạng Bảng 3.2.

b) Tìm các tứ phân vị và giải thích ý nghĩa của chúng.

Lời giải

a)

Điểm	[0;10)	[10;20)	[20;30)	[30;40)	[40;50)	[50;60)	[60;70)	[70;80)	[80;90)	[90;100)
Số thí sinh	1	2	4	6	15	12	10	6	3	1

b) Cỡ mẫu $n = 60$

Tứ phân vị thứ nhất Q_1 là $\frac{x_{15} + x_{16}}{2}$. Do x_{15}, x_{16} đều thuộc nhóm [40;50) nên nhóm này chứa Q_1 .

Do đó, $p = 5; a_5 = 40, m_5 = 15, m_1 + m_2 + m_3 + m_4 = 1 + 2 + 4 + 6 = 13, 4$

$4a_6 - a_5 = 10$ và ta có:

$$Q_1 = 40 + \frac{\frac{60}{4} - 13}{15} \times 10 = 41.33$$

Ý nghĩa: Có 25% số giá trị nhỏ hơn 41.3

Tứ phân vị thứ hai tức M_e là $\frac{x_{30} + x_{31}}{2}$. Do x_{30}, x_{31} đều thuộc nhóm [50;60) nên nhóm này chứa M_e .

Do đó, $p = 6; a_6 = 50, m_6 = 12, m_1 + m_2 + m_3 + m_4 + m_5 = 1 + 2 + 4 + 6 + 15 = 28, a_7 - a_6 = 10$ và ta có:

$$M_e = 50 + \frac{\frac{60}{2} - 28}{12} \times 10 = 51.67$$

Ý nghĩa: Có 50% số giá trị nhỏ hơn 51.67

Tứ phân vị thứ ba Q_3 là $\frac{x_{45} + x_{46}}{2}$. Do x_{45}, x_{46} đều thuộc nhóm [60;70) nên nhóm này chứa Q_3 .

Do đó, $p = 7; a_7 = 60, m_7 = 10$ $m_1 + m_2 + m_3 + m_4 + m_5 + m_6 = 1 + 2 + 4 + 6 + 15 + 12 = 40$
 $a_8 - a_7 = 10$ và ta có:

$$Q_3 = 60 + \frac{\frac{60 \times 3}{4} - 40}{10} \times 10 = 65$$

Ý nghĩa: Có 75% số giá trị nhỏ hơn 65

Tứ phân vị thứ hai tức M_e là $\frac{x_{30} + x_{31}}{2}$. Do x_{30}, x_{31} đều thuộc nhóm $[50; 60)$ nên nhóm này chứa M_e .

Do đó, $p = 6; a_6 = 50, m_6 = 12$

$$m_1 + m_2 + m_3 + m_4 + m_5 = 1 + 2 + 4 + 6 + 15 = 28$$

$a_7 - a_6 = 10$ và ta có:

$$Q_3 = 60 + \frac{\frac{60 \times 3}{4} - 40}{10} \times 10 = 65$$

Ý nghĩa: Có 75% số giá trị nhỏ hơn 65

Câu 19: Người ta ghi lại tuổi thọ của một số con ong cho kết quả như sau:

Tuổi thọ (ngày)	[0; 20)	[20; 40)	[40; 60)	[60; 80)	[80; 100)
Số lượng	5	12	23	31	29

Tìm một của mẫu số liệu. Giải thích ý nghĩa của giá trị nhận được.

Lời giải

Tần số lớn nhất là 31 nên nhóm chứa một là $[60; 80)$. Ta có:

$$j = 4, a_4 = 60, m_4 = 31, m_3 = 23, m_5 = 29, h = 20.$$

$$\text{Do đó } M_o = 60 + \frac{31 - 23}{(31 - 23) + (31 - 29)} \times 20 = 76$$

Ý nghĩa: Đa số các con ong có tuổi thọ là 76 ngày

Câu 20: Một bảng xếp hạng đã tính điểm chuẩn hoá cho chỉ số nghiên cứu của một số trường đại học ở Việt Nam và thu được kết quả sau:

Điểm	Dưới 20	[20; 30)	[30; 40)	[40; 60)	[60; 80)	[80; 100)
Số trường	4	19	6	2	3	1

Xác định điểm ngưỡng để đưa ra danh sách 25% trường đại học có chỉ số nghiên cứu tốt nhất Việt Nam.

Lời giải

Điểm ngưỡng để đưa ra danh sách 25% trường đại học có chỉ số nghiên cứu tốt nhất Việt Nam là tứ phân vị thứ ba

Ta có: cỡ mẫu $n = 35$

Tứ phân vị thứ ba Q_3 là x_{27} mà x_{27} thuộc nhóm $[30;40)$ nên nhóm này chứa Q_3 . Do đó, $p = 3, a_3 = 30, m_3 = 6, m_1 + m_2 = 4 + 19 = 23, a_4 - a_3 = 10$ và ta có:

$$Q_3 = 30 + \frac{\frac{3 \times 35}{4} - 23}{6} \times 10 = 35.42$$

Vậy để đưa ra danh sách 25% trường đại học có chỉ số nghiên cứu tốt nhất Việt Nam ta lấy các trường có điểm chuẩn hóa trên 35.42

Câu 21: 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

- Tính cự li trung bình của mỗi lần ném.
- 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	?	?	?	?	?

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

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

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

Cự li trung bình mỗi lần ném xấp xỉ bằng

$$[71,6; 72,4)$$

d) Nhóm chứa một của mẫu số liệu trên là

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

Do đó:

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

Câu 22: 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 một của mẫu số liệu ghép nhóm trên.

Lời giải

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:

$$(18 \cdot 3 + 23 \cdot 6 + 28 \cdot 15 + 33 \cdot 27 + 38 \cdot 22 + 43 \cdot 14 + 48 \cdot 5) : 92 = 34,6$$

$$[30,5; 35,5)$$

Nhóm chứa một của mẫu số liệu là

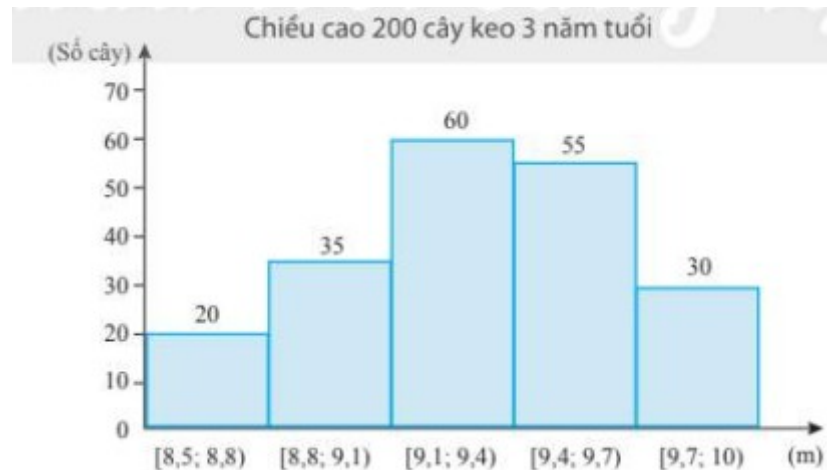
$$u_m = 30,5; n_{m-1} = 15, n_{m+1} = 22, u_{m+1} - u_m = 35,5 - 30,5 = 5$$

Do đó

Mốt của mẫu số liệu ghép nhóm là:

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

Câu 23: 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.

Lời giải

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:

$$(8,65 \cdot 20 + 8,95 \cdot 35 + 9,25 \cdot 60 + 9,55 \cdot 55 + 9,85 \cdot 30) : 200 = 9,31 \text{ (m)}$$

$$[9,1; 9,4)$$

Nhóm chứa mốt của mẫu số liệu là

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

Do đó:

Mốt của mẫu số liệu trên là:

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

Câu 24: Lương tháng của một số nhân viên một văn phòng được ghi lại như sau (đơn vị: triệu đồng):

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

- 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.

Lời giải

- a) Tứ phân vị thứ nhất là: 9,01.

Tứ phân vị thứ hai là: 10,8.

Tứ phân vị thứ ba là: 12,35.

- b)

Lương tháng (triệu đồng)	[6;8)	[8;10)	[10;12)	[12;14)
Số nhân viên	3	6	8	7

$$x_1, x_2, x_3, \dots, x_{24}$$

- c) Gọi lần lượt là số nhân viên theo thứ tự không gian.

$$x_1, \dots, x_3 \in [6; 8); x_4, \dots, x_9 \in [8; 10); x_{10}, \dots, x_{17} \in [10; 12); x_{18}, \dots, x_{24} \in [12; 14)$$

Do

;

$$\frac{1}{2}(x_{12} + x_{13}) \quad [10; 12)$$

Tứ phân vị thứ hai của dãy số liệu là thuộc nhóm nên tứ phân vị thứ hai

$$Q_2 = 10 + \frac{\frac{24}{2} - 9}{8} (12 - 10) = 10,75$$

của mẫu số liệu là

$$\frac{1}{2}(x_6 + x_7) \quad [8; 10)$$

Tứ phân vị thứ nhất của dãy số liệu là thuộc nhóm nên tứ phân vị thứ nhất

$$Q_1 = 8 + \frac{\frac{24}{4} - 3}{6} (10 - 8) = 9$$

của mẫu số liệu là

$$\frac{1}{2}(x_{18} + x_{19}) \quad [12; 14)$$

Tứ phân vị thứ ba của dãy số liệu là thuộc nhóm nên tứ phân vị thứ ba của

$$Q_3 = 12 + \frac{\frac{3 \cdot 24}{4} - 17}{7}(14 - 12) = 12,3$$

mẫu số liệu là

Câu 25: Kiểm tra điện lượng của một số viên pin tiểu do một hãng sản xuất thu được kết quả sau:

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

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

Lời giải

Điện lượng (nghìn mAh)	[0,9;0,95)	[0,95;1,0)	[1,0;1,05)	[1,05;1,1)	[1,1;1,15)
Giá trị đại diện	0,925	0,975	1,025	1,075	1,125
Số viên pin	10	20	35	15	5

Số trung bình của dãy số liệu xấp xỉ bằng:

$$(0,925 \cdot 10 + 0,975 \cdot 20 + 1,025 \cdot 35 + 1,075 \cdot 15 + 1,125 \cdot 5) : 85 = 1,016$$

$$[1,0; 1,05)$$

Nhóm chứa một của dãy số liệu là:

$$M_0 = 1,0 + \frac{35 - 20}{(35 - 20) + (35 - 15)} \cdot (1,05 - 1,0) = 1,02$$

$$x_1, x_2, x_3, \dots, x_{85}$$

Gọi lần lượt là số viên pin theo thứ tự không gian.

$$x_1, \dots, x_{10} \in [0,9; 0,95); x_{11}, \dots, x_{30} \in [0,95; 1,0); x_{31}, \dots, x_{65} \in [1,0; 1,05)$$

Do

$$x_{66}, \dots, x_{80} \in [1,05; 1,1); x_{81}, \dots, x_{85} \in [1,1; 1,15)$$

$$\frac{1}{2}(x_{42} + x_{43}) \quad [1, 0; 1, 05)$$

Tứ phân vị thứ hai của dãy số liệu là thuộc nhóm nên tứ phân vị thứ

$$Q_2 = 1,0 + \frac{\frac{85}{2} - 30}{35}(1,05 - 1,0) = 1,02$$

hai của mẫu số liệu là

$$\frac{1}{2}(x_{21} + x_{22}) \quad [0,95; 1,0)$$

Tứ phân vị thứ nhất của dãy số liệu là thuộc nhóm nên tứ phân vị thứ

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

nhất của mẫu số liệu là

$$\frac{1}{2}(x_{63} + x_{64}) \quad [1, 0; 1, 05)$$

Tứ phân vị thứ ba của dãy số liệu là thuộc nhóm nên tứ phân vị thứ ba

$$Q_3 = 1,0 + \frac{\frac{3.85}{4} - 30}{35}(1,05 - 1,0) = 1,048$$

của mẫu số liệu là

Câu 26: 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, tứ phân vị và một của mẫu số liệu ghép nhóm trên.

Lời giải

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

Số trung bình của mẫu số liệu xấp xỉ bằng:

$$(6,75.8 + 7,25.10 + 7,75.16 + 8,25.24 + 8,75.13 + 9,25.7 + 9,75.4) : 82 = 8,12$$

$$[8; 8,5)$$

Nhóm chứa một của mẫu số li là:

Một của mẫu số liệu là:

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

$$x_1, x_2, x_3, \dots, x_{85}$$

Gọi lần lượt là tần số theo thứ tự không gian.

$$x_1, \dots, x_8 \in [6, 5; 7); x_9, \dots, x_{18} \in [7; 7, 5); x_{19}, \dots, x_{34} \in [7, 5; 8)$$

Do ;

$$x_{35}, \dots, x_{58} \in [8; 8, 5); x_{59}, \dots, x_{71} \in [8, 5; 9); \dots$$

$$\frac{1}{2}(x_{41} + x_{42}) \quad [8; 8, 5)$$

Tứ phân vị thứ hai của dãy số liệu là thuộc nhóm nên tứ phân vị thứ hai

$$Q_2 = 8 + \frac{\frac{82}{2} - 34}{24} (8,5 - 8) = 8,15$$

của mẫu số liệu là

$$\frac{1}{2}(x_{20} + x_{21}) \quad [7, 5; 8)$$

Tứ phân vị thứ nhất của dãy số liệu là thuộc nhóm nên tứ phân vị thứ nhất

$$Q_1 = 7,5 + \frac{\frac{82}{4} - 18}{16} (8 - 7,5) = 7,58$$

của mẫu số liệu là

$$\frac{1}{2}(x_{61} + x_{62}) \quad [8, 5; 9)$$

Tứ phân vị thứ ba của dãy số liệu là thuộc nhóm nên tứ phân vị thứ ba của

$$Q_3 = 8,5 + \frac{\frac{3 \cdot 82}{4} - 58}{13} (9 - 8,5) = 8,63$$

mẫu số liệu là

Câu 27: Để kiểm tra thời gian sử dụng pin của 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

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

25%

b) Chị An cho rằng có khoảng số lần sạc điện thoại chỉ dùng được dưới 10 giờ. Nhận định của chị An có hợp lí không?

Lời giải

Thời gian sử dụng (giờ)	[7;9)	[9;11)	[11;13)	[13;15)	[15;17)
Giá trị đại diện	8	10	12	14	16
Số lần	2	5	7	6	3

a) Thời gian sử dụng trung bình xấp xỉ bằng:

$$(8.2 + 10.5 + 12.7 + 14.6 + 16.3) : 23 = 12,3$$

(giờ)

$$x_1, x_2, x_3, \dots, x_{23}$$

b) Gọi lần lượt là số lần sử dụng theo thứ tự không gian.

$$x_1, x_2 \in [7;9); x_3, \dots, x_7 \in [9;11); x_8, \dots, x_{14} \in [11;13)$$

Do

$$x_{15}, \dots, x_{20} \in [13;15), \dots$$

$$\frac{1}{2}(x_5 + x_6)$$

Tứ phân vị thứ nhất của dãy số liệu là thuộc nhóm [9;11) nên tứ phân vị thứ nhất

$$Q_1 = 9 + \frac{\frac{23}{4} - 2}{5}(11 - 9) = 10,5$$

của mẫu số liệu là

25%

Do ba điểm tứ phân vị chia mẫu số liệu thành 4 phần, mỗi phần chứa số lượng các số liệu nên ta thấy nhận định của chị An là hợp lí.

Câu 28:

121,8	158,3	334,9	200,9	165,6	161,5	194,3	220,7	189,8	234,2
165,9	165,9	134	173	169	189	254	168	255	

(Nguồn: Tổng cục Thống kê)

năm 2020 được ghi lại như dưới đây (đơn vị: mm):

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

b) Hoàn thiện 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	?	?	?	?

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

Lời giải

a) Số trung bình của mẫu số liệu trên là 192,5.

Mốt của mẫu số liệu trên là 165,9.

$$x_1, x_2, x_3, \dots, x_{19}$$

Gọi lần lượt là số năm theo thứ tự không gian.

$$x_{10} = 173$$

Tứ phân vị thứ hai của mẫu số liệu là .

$$\frac{1}{2}(x_4 + x_5) = 163,55$$

Tứ phân vị thứ nhất của mẫu số liệu là .

$$\frac{1}{2}(x_{14} + x_{15}) = 210,8$$

Tứ phân vị thứ ba của mẫu số liệu là .

b)

Tổng lượng mưa trong tháng 8	[120;175)	[175;230)	[230;285)	[285;340)
Giá trị đại diện	147,5	202,5	257,5	312,5
Số năm	10	5	3	1

c) Số trung bình của dãy số liệu xấp xỉ bằng:

$$(147,5 \cdot 10 + 202,5 \cdot 5 + 257,5 \cdot 3 + 312,5 \cdot 1) : 19 = 188$$

Nhóm chứa mốt của dãy số liệu là:

$$M_0 = 120 + \frac{10 - 0}{(10 - 0) + (10 - 5)} \cdot (175 - 120) = 156,7$$

$$\text{Do } x_1, \dots, x_{10} \in [120; 175); x_{11}, \dots, x_{15} \in [175; 230); x_{16}, \dots, x_{18} \in [230; 285); x_{19} \in [285; 340)$$

Tứ phân vị thứ hai của dãy số liệu là x_{10} thuộc nhóm $[120; 175)$ nên tứ phân vị thứ hai của mẫu

$$Q_2 = 120 + \frac{\frac{19}{2} - 0}{10} (175 - 120) = 172,5$$

số liệu là .

$$\frac{1}{2}(x_4 + x_5) \quad [120; 175)$$

Tứ phân vị thứ nhất của dãy số liệu là thuộc nhóm nên tứ phân vị thứ

$$Q_1 = 120 + \frac{\frac{19}{4} - 0}{10}(175 - 120) = 146,125$$

nhất của mẫu số liệu là .

$$\frac{1}{2}(x_{14} + x_{15}) \quad [175; 230)$$

Tứ phân vị thứ ba của dãy số liệu là thuộc nhóm nên tứ phân vị thứ ba

$$Q_3 = 175 + \frac{\frac{3 \cdot 19}{4} - 10}{5}(230 - 175) = 221,75$$

của mẫu số liệu là .

Câu 29: 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: worldometers.info)

a) Xác định số trung bình và tứ phân vị của mẫu số liệu trên. Mẫu số liệu có bao nhiêu giá trị ngoại lệ?

b) Hoàn thiện 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	?	?	?	?	?

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

Lời giải

a) Số trung bình của mẫu số liệu là: 15821.

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

Tứ phân vị thứ nhất của mẫu số liệu là 15033.

Tứ phân vị thứ ba của mẫu số liệu là 16474,6.

b)

Số ca (nghìn)	[14;15,5)	[15,5;17)	[17;18,5)	[18,5;20)	[20;21,5)
Giá trị đại diện	14,75	16,25	17,75	19,25	20,75
Số ngày	14	14	2	0	1

c) Số trung bình của dãy số liệu xấp xỉ bằng:

$$(14,75.14 + 16,25.14 + 17,75.2 + 19,25.0 + 20,75.1) : 31 = 15,81$$

$$x_1, x_2, x_3, \dots, x_{31}$$

Gọi lần lượt là số ngày theo thứ tự không gian.

$$x_1, \dots, x_{14} \in [14;15,5); x_{15}, \dots, x_{28} \in [15,5;17); x_{29}, x_{30} \in [17;18,5)$$

Do

Tứ phân vị thứ hai của dãy số liệu là x_{16} thuộc nhóm $[15,5;17)$ nên tứ phân vị thứ hai của mẫu

$$Q_2 = 15,5 + \frac{\frac{31}{2} - 14}{14} (17 - 15,5) = 15,7$$

số liệu là

Tứ phân vị thứ nhất của dãy số liệu là $\frac{1}{2}(x_7 + x_8)$ thuộc nhóm $[14;15,5)$ nên tứ phân vị thứ nhất

$$Q_1 = 14 + \frac{\frac{31}{4} - 0}{14} (15,5 - 14) = 14,8$$

của mẫu số liệu là

Tứ phân vị thứ ba của dãy số liệu là $\frac{1}{2}(x_{23} + x_{24})$ thuộc nhóm $[15,5;17)$ nên tứ phân vị thứ ba

$$Q_3 = 15,5 + \frac{\frac{3.31}{4} - 14}{14} (17 - 15,5) = 16,5$$

của mẫu số liệu là

Câu 30: 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 một 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.

Câu 31: 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 một 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 một 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

Câu 32: 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.

Câu 33: 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

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

Cự li trung bình mỗi lần ném xấp xỉ bằng

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:

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

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

Câu 34: 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 táo 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

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

Câu 35: 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 $x_1; x_2; \dots; x_{58} \in [10,5;15,5)$; $x_{59}; x_{55}; \dots; x_{145} \in [15,5;20,5)$ nên trung vị của mẫu số liệu $x_1; x_2; \dots; x_{266}$ 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

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

Câu 36: 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 $Q_k = u_j$.

Câu 37: 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

a) Hãy ước lượng các tứ phân vị của mẫu số liệu ghép nhóm trên.

b) 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 lý 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 lý.

Câu 38: Mẫu số liệu đây ghi lại tốc độ của 40 ô tô khi đi qua một trạm đo tốc độ (đơn vị: km/h)

48, 5 43 50 55 45 60 53 55,5 44 65

51 62,5 41 44,5 57 57 68 49 46,5 53,5

61 49,5 54 62 59 56 47 50 60 61

49,5 52,5 57 47 60 55 45 47,5 48 61,5

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

$$[40; 45), [45; 50), [50; 55), [55; 60), [60; 65), [65; 70)$$

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) Bảng tần số ghép nhóm bao gồm cả giá trị đại diện là:

Nhóm	Giá trị đại diện	Tần số
$[40; 45)$	42,5	4
$[45; 50)$	47,5	11
$[50; 55)$	52,5	7
$[55; 60)$	57,5	8
$[60; 65)$	62,5	8

$[65; 70)$	67,5	n
------------	------	-----

b) Trung bình cộng là:

$$\bar{x} = \frac{42,5 \cdot 4 + 47,5 \cdot 11 + 52,5 \cdot 7 + 57,5 \cdot 8 + 62,5 \cdot 8 + 67,5 \cdot 2}{40} = 53,875$$

Trung vị là:

Có bảng ghép nhóm bao gồm cả tần số tích lũy là

Nhóm	Tần số	Tần số tích lũy
$[40; 45)$	4	4
$[45; 50)$	11	15
$[50; 55)$	7	22
$[55; 60)$	8	30
$[60; 65)$	8	38
$[65; 70)$	2	40

Số phần tử của mẫu là $n = 40$. Ta có: $\frac{n}{2} = \frac{40}{2} = 20 \Rightarrow$ Nhóm 3 là nhóm đầu tiên có tần số tích lũy lớn hoặc bằng 20.

Xét nhóm 3 là nhóm $[50; 55)$ có $r = 50; d = 5; n_3 = 7$ và nhóm 2 là nhóm $[45; 50)$ có $cf_2 = 15$.

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

Áp dụng công thức, ta có trung vị của mẫu số liệu là:

Q_1 là:

Số phần tử của mẫu là $n = 40$.

$$\text{Ta có } \frac{n}{4} = \frac{40}{4} = 10$$

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

Xét nhóm 2 là nhóm $[45; 50)$ có $r = 45; d = 5; n\{2\} = 11$ và nhóm 1 là nhóm $[40; 45)$ có $cf_1 = 4$

$$Q_1 = 45 + \left(\frac{10 - 4}{11} \right) \cdot 5 \approx 47,7 \text{ (km / h)}$$

Áp dụng công thức, ta có Q_1 của mẫu số liệu là

Q_2 là: Có $Q_2 = M_e \approx 53,6 \text{ (km / h)}$

Q_3 là:

Ta có $\frac{3n}{4} = 30$. Suy ra nhóm 4 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 30. Xét nhóm 4 là nhóm $[55; 60)$ có $r = 55; d = 5; n_4 = 8$ và nhóm 3 là nhóm $[50; 55)$ có $cf_3 = 22$

Áp dụng công thức, ta có Q_3 của mẫu số liệu là:

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

c) Một của mẫu số liệu là

$$\Rightarrow M_o = 45 + \left(\frac{11 - 4}{2 \cdot 11 - 4 - 7} \right) \cdot 5 \approx 43,2$$

Có nhóm 2 là nhóm có tần số lớn nhất

Câu 39: Mẫu số liệu sau ghi lại cân nặng của 30 bạn học sinh (đơn vị: kilôgam):

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) Một của mẫu số liệu ghép nhóm trên là bao nhiêu?

Lời giải

a) Bảng tần số ghép nhóm bao gồm cả giá trị đại diện là

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	030

b) Trung bình cộng là:

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

Trung vị là:

Có bảng ghép nhóm bao gồm cả tần số tích lũy là:

Nhóm	Tần số	Tần số tích lũy
[15; 20)	1	1
[20; 25)	0	1
[25; 30)	0	1
[30; 35)	1	2
[35; 40)	10	12
[40; 45)	17	29
[45; 50)	0	29
[50; 55)	1	0

Số phần tử của mẫu là $n = 30$. Ta có:

$$M_c = 40 + \left(\frac{15-12}{17} \right) \cdot 5 \approx 40,9 \text{ (kilôgam)}$$

Áp dụng công thức, ta có trung vị của mẫu số liệu là:

Q_1 là:

Số phần tử của mẫu là $n = 30$.

Ta có $\frac{n}{4} = \frac{30}{4} = 7,5$. Suy ra nhóm 5 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 7,5.

Xét nhóm 5 là nhóm $[35; 40)$ có $r = 35; d = 5; n_s = 10$ và nhóm 4 là nhóm $[30; 35)$ có $cf_4 = 2$

$$Q_1 = 35 + \left(\frac{7,5-2}{10} \right) \cdot 5 = 37,75 \text{ (kilôgam)}$$

Áp dụng công thức, ta có Q_1 của mẫu số liệu là

Q_2 là: Có $Q_2 = M_c \approx 40,9$ (kilôgam)

Q_3 là:

Ta có $\frac{3n}{4} = 22,5$. Suy ra nhóm 6 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 22,5.

Xét nhóm 6 là nhóm $[40; 45)$ có $r = 40; d = 5; n_6 = 17$ và nhóm 5 là nhóm $[35; 40)$ có $cf_5 = 12$

Áp dụng công thức, ta có Q_3 của mẫu số liệu là:

$$Q_3 = 40 + \left(\frac{22,5-12}{17} \right) \cdot 5 = 43,1 \text{ (kilôgam)}$$

c) Mốt của mẫu số liệu là:

$$M_o = 40 + \left(\frac{17-10}{2 \cdot 17-10} \right) \cdot 5 \approx 41,46$$

Có nhóm 6 là nhóm có tần số lớn nhất

Câu 40: Mẫu số liệu dưới đây ghi lại độ dài quãng đường di chuyển trong một tuần (đơn vị: kilômét) của 40 chiếc ô tô:

100	105	115	116	130	135	138	132	135	120
125	128	120	124	140	140	146	145	142	142
145	148	150	150	159	155	151	156	155	151
154	152	153	160	162	175	176	165	188	198

a) Lập bảng tần số ghép nhóm bao gồm cả tần số tích lũy với năm nhóm ứng với năm nửa khoảng:

$$[100;120), [120;140), [140;160), [160;180), [180;200)$$

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

Nhóm	Tần số	Tần số tích lũy
$[100;120)$	4	4
$[120;140)$	15	19
$[140;160)$	14	33
$[160;180)$	5	38
$[180;200)$	2	40

$$\bar{x} = \frac{110 \cdot 4 + 130 \cdot 15 + 150 \cdot 14 + 170 \cdot 5 + 190 \cdot 2}{40} = 143$$

b) Trung bình cộng là:

$$M_e = 140 + \left(\frac{20-19}{14} \right) \cdot 20 \approx 141$$

• Trung vị là:

$$Q_1 = 120 + \left(\frac{10-4}{15} \right) \cdot 20 = 128$$

• Q_1 là

Q_2 là: Có $Q_2 = M_e \approx 141$

$$Q_3 \text{ là: } Q_3 = 140 + \left(\frac{30-19}{15} \right) \cdot 20 = 155,6$$

c) Một của mẫu số liệu là:

$$\Rightarrow M_o = 120 + \left(\frac{15-4}{2 \cdot 15 - 4 - 14} \right) \cdot 20 \approx 138,3$$

Có nhóm 2 là nhóm có tần số lớn nhất

Câu 41: 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) Một 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

$$\bar{x} = \frac{35 \cdot 4 + 45 \cdot 10 + 55 \cdot 14 + 65 \cdot 6 + 75 \cdot 4 + 85 \cdot 2}{40} = 55,5$$

⌘ Tìm trung vị

- Khoảng chứa trung vị [50; 60)

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

Trung vị của mẫu số liệu là

⌘ 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 một

Nhóm chứa một: $[50; 60)$

$$\text{Một của mẫu số liệu: } M_0 = 50 + \frac{4}{4+8} \cdot 10 \approx 53,3$$

Câu 42: 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) Một 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)

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

Số trung bình cộng:

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

Trung vị Q_1 và Q_3 :

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

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

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

c) Một của mẫu số liệu

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

Câu 43: Một bưu tá thống kê lại số bưu phẩm gửi đến một cơ quan mỗi ngày trong tháng 6/2022 ở bảng sau:

30	32	28	34	37	26	44	24	22	38
34	20	30	27	28	34	38	32	42	39
43	42	32	26	36	32	37	24	29	32

a) Tính số trung bình và một của mẫu số liệu trên.

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

Số bưu phẩm	[20; 24]	[25; 29]	[30; 34]	[35; 39]	[40; 44]
Số ngày	?	?	?	?	?

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

Lời giải

a) Số trung bình của mẫu số liệu là 32,4. Một của mẫu số liệu là 32.

b) Bảng tần số ghép nhóm.

Số bưu phẩm	[20; 24]	[25; 29]	[30; 34]	[35; 39]	[40; 44]
Số ngày	4	6	10	6	4

c) Do số bưu phẩm là số nguyên nên ta hiệu chỉnh lại bảng tần số ghép nhóm như sau:

Số bưu phẩm	[19,5; 24,5)	[24,5; 29,5)	[29,5; 34,5)	[34,5; 39,5)	[39,5; 44,5)
Gía trị đại diện	22	27	32	37	42
Số ngày	4	6	10	6	4

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x} = \frac{4 \cdot 22 + 6 \cdot 27 + 10 \cdot 32 + 6 \cdot 37 + 4 \cdot 42}{30} = 32.$$

[29,5; 34,5)

Nhóm chứa một của mẫu số liệu trên là nhóm

$$u_m = 29,5; n_{m-1} = 6; n_m = 10; n_{m+1} = 6; u_{m+1} - u_m = 34,5 - 29,5 = 5$$

Do đó:

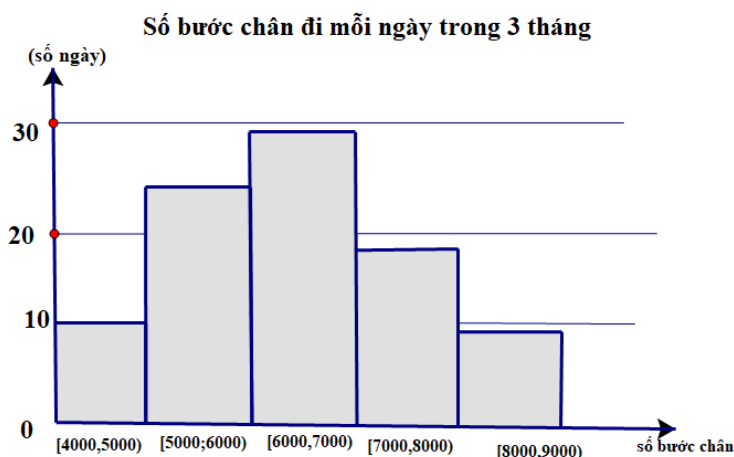
$$M_0 = 29,5 + \frac{10 - 6}{(10 - 6) + (10 - 6)} \cdot 5 = 32$$

Mốt của mẫu số liệu ghép nhóm là

Câu 44: Thảo thống kê lại số bước chân bạn đi mỗi ngày trong 3 tháng. Kết quả được biểu diễn ở biểu đồ ở bên.

a) Hãy lập bảng tần số ghép nhóm, kèm theo giá trị đại diện biểu diễn dữ liệu thống kê trên.

b) Hãy ước lượng số trung bình và một của mẫu số liệu ghép nhóm trên.



Lời giải

a) Bảng tần số ghép nhóm:

Số bước chân	Giá trị đại diện	Số ngày
$[4000; 5000)$	4500	12
$[5000; 6000)$	5500	24
$[6000; 7000)$	6500	29
$[7000; 8000)$	7500	18
$[8000; 9000)$	8500	9

b) Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x} = \frac{12 \cdot 4500 + 24 \cdot 5500 + 29 \cdot 6500 + 18 \cdot 7500 + 9 \cdot 8500}{92} \approx 6369,57$$

Nhóm chứa một của mẫu số liệu trên là nhóm $[6000; 7000)$.

$$u_m = 6000; n_{m-1} = 24; n_m = 29; n_{m+1} = 18; u_{m+1} - u_m = 7000 - 6000 = 1000$$

Do đó:

Mốt của mẫu số liệu ghép nhóm là

$$M_o = 6000 + \frac{29 - 24}{(29 - 24) + (29 - 18)} \cdot 1000 = 6312,5.$$

Câu 45: Diện tích các tỉnh và thành phố khu vực Nam Bộ được thống kê ở bảng sau:

Tỉnh/ thành phố	Diện tích (km ²)	Tỉnh/ thành phố	Diện tích (km ²)
Bình Phước	6877	Vĩnh Long	1526
Tây Ninh	4041	Đồng Tháp	3384
Bình Dương	2695	An Giang	3537
Đồng Nai	5864	Kiên Giang	6349
Bà Rịa - Vũng Tàu	1981	Cần Thơ	1439
TP.Hồ Chí Minh	2061	Hậu Giang	1622
Long An	4495	Sóc Trăng	3312
Tiền Giang	2511	Bạc Liêu	2669
Bến Tre	2395	Cà Mau	5221
Trà Vinh	2358		

(Nguồn: Tổng cục Thống kê)

a) Hãy tính diện tích trung bình của mỗi tỉnh/thành phố khu vực Nam Bộ.

b) Dựa vào số liệu trên, hãy hoàn thiện bảng tần số ghép nhóm về diện tích các tỉnh khu vực Nam Bộ theo mẫu sau:

Diện tích (km ²)	[1000; 2500)	[2500; 4000)	[4000; 5500)	[5500; 7000)
Số tỉnh/thành phố	?	?	?	?

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

Lời giải

a) Diện tích trung bình của mỗi tỉnh/thành phố khu vực Nam Bộ là

$$(6877 + 4041 + 2695 + 5864 + 1981 + 2061 + 4495 + 2511 + 2395 + 2358 + 1526 + 3384 + 3537 + 6349 + 1439 + 1622 + 3312 + 2669 + 5221) : 19 \approx 3386,16 \text{ (km}^2\text{)}$$

b) Bảng tần số ghép nhóm:

Diện tích (km ²)	[1000; 2500)	[2500; 4000)	[4000; 5500)	[5500; 7000)
------------------------------	--------------	--------------	--------------	--------------

Số tỉnh/thành phố	7	6	3	3
-------------------	---	---	---	---

c) Bảng tần số ghép nhóm bao gồm giá trị đại diện của các nhóm như sau:

Diện tích (km ²)	[1000; 2500)	[2500; 4000)	[4000; 5500)	[5500; 7000)
Giá trị đại diện	1750	3250	4750	6250
Số tỉnh/ thành phố	7	6	3	3

Khi đó, số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x} = \frac{7 \cdot 1750 + 6 \cdot 3250 + 3 \cdot 4750 + 3 \cdot 6250}{19} \approx 3407,89.$$

[1000; 2500)

Nhóm chứa một của mẫu số liệu trên là nhóm

$$u_m = 1000; n_{m-1} = 0; n_m = 7; n_{m+1} = 6; u_{m+1} - u_m = 2500 - 1000 = 1500$$

Do đó:

Mốt của mẫu số liệu ghép nhóm là

$$M_0 = 1000 + \frac{7-0}{(7-0)+(7-6)} \cdot 1500 = 2312,5.$$

Câu 46: Nhân ngày hội đọc sách, các học sinh của một trường trung học phổ thông mang sách cũ đến tặng thư viện trường và trao đổi với các bạn học sinh khác. Bảng sau thống kê số sách cũ mà các bạn học sinh lớp 11 B mang đến trường.

Số sách	[1; 3]	[4; 6]	[7; 9]	[10; 12]	[13; 15]
Số học sinh	5	14	10	8	3

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

Lời giải

Bảng số liệu ghép nhóm hiệu chỉnh như sau:

Số sách	[0,5; 3,5)	[3,5; 6,5)	[6,5; 9,5)	[9,5; 12,5)	[12,5; 15,5)
Số học sinh	5	14	10	8	3

Số trung bình của mẫu số liệu ghép nhóm là 7,25.

$$\frac{145}{26} \approx 5,58$$

Mốt của mẫu số liệu ghép nhóm là

Câu 47: Bảng sau thống kê số lượt chở khách mỗi ngày của một lái xe taxi trong 30 ngày.

15	13	7	5	18	13	11	9	10	8	14	11	16	10	9
----	----	---	---	----	----	----	---	----	---	----	----	----	----	---

13	11	12	13	15	12	13	6	8	17	13	6	18	12	13
----	----	----	----	----	----	----	---	---	----	----	---	----	----	----

a) Hãy tính số trung bình và một của mẫu số liệu trên.

$$[4,5;7,5)$$

b) Hãy lập bảng tần số ghép nhóm của mẫu số liệu trên với nhóm đầu tiên là .

c) Hãy ước lượng số trung bình và một của mẫu số liệu ghép nhóm đó.

Lời giải

a) Số trung bình mẫu số liệu là 11,7. Một của mẫu số liệu là 13.

b) Bảng tần số ghép nhóm của mẫu số liệu đã cho như sau:

Số lượt khách	$[4,5;7,5)$	$[7,5;10,5)$	$[10,5;13,5)$	$[13,5;16,5)$	$[16,5;19,5)$
Số ngày	4	6	13	4	3

c) Số trung bình của mẫu số liệu ghép nhóm là 11,6.

$$\text{Một của mẫu số liệu ghép nhóm là } \frac{189}{16} \approx 11,8$$

Câu 48: Bảng sau thống kê cân nặng (đơn vị: kg) của một số con ngan đực 88 ngày tuổi ở một trang trại.

4,60	4,62	4,64	4,65	4,67	4,67	4,68	4,68	4,70	4,70
4,70	4,70	4,71	4,71	4,72	4,73	4,74	4,76	4,77	4,77
4,77	4,78	4,78	4,80	4,82	4,84	4,84	4,85	4,87	4,89
4,89	4,90	4,92	4,92	4,93	4,94	4,94	4,95	4,97	4,97
4,97	4,99	4,99	5,01	5,02	5,03	5,04	5,05	5,06	5,07

$$[4,6;4,7)$$

a) Hãy lập bảng tần số ghép nhóm của mẫu số liệu trên với nhóm đầu tiên là .

b) Hãy ước lượng số trung bình và một của mẫu số liệu ghép nhóm trên.

Lời giải

a) Bảng tần số ghép nhóm của mẫu số liệu đã cho như sau:

Cân nặng (kg)	$[4,6;4,7)$	$[4,7;4,8)$	$[4,8;4,9)$	$[4,9;5,0)$	$[5,0;5,1)$
Số con ngan	8	15	8	12	7

b) Số trung bình của mẫu số liệu ghép nhóm là 4,84. Một của mẫu số liệu ghép nhóm là 4,75.

Câu 49: Bảng sau thống kê chiều cao (đơn vị: cm) của một số cây giống sau khi nảy mầm được 2 tuần.

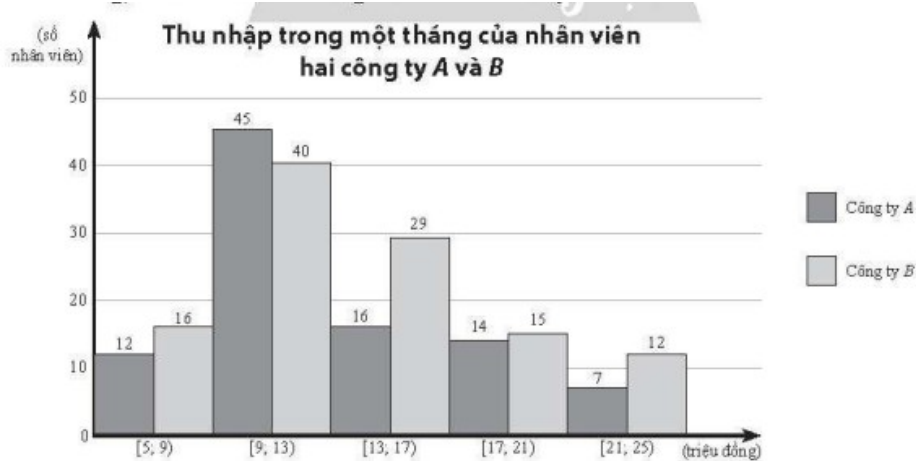
Chiều cao (cm)	$[6, 2; 6, 7)$	$[6, 7; 7, 2)$	$[7, 2; 7, 7)$	$[7, 7; 8, 2)$	$[8, 2; 8, 7)$
Số cây	10	21	28	12	9

Hãy ước lượng chiều cao trung bình và một của mẫu số liệu ghép nhóm trên.

Lời giải

Số trung bình của mẫu số liệu ghép nhóm là khoảng 7,38. Một của mẫu số liệu ghép nhóm là khoảng 7,35.

Câu 50: Thống kê lại thu nhập trong một tháng của nhân viên hai công ty A và B (đơn vị: triệu đồng) được thể hiện trong biểu đồ dưới đây.



Hãy so sánh thu nhập trung bình của nhân viên hai công ty theo số trung bình và một của mẫu số liệu ghép nhóm.

Lời giải

Ta lập bảng thống kê thu nhập của nhân viên hai công ty như sau:

Thu nhập (đơn vị: triệu đồng)	Giá trị đại diện	Số nhân viên công ty A	Số nhân viên công ty B
$[5; 9)$	7	12	16
$[9; 13)$	11	45	40
$[13; 17)$	15	16	29
$[17; 21)$	19	14	15
$[21; 25)$	23	7	12

Số trung bình thu thập trong một tháng của nhân viên công ty A là

$$\bar{x}_A = \frac{12 \cdot 7 + 45 \cdot 11 + 16 \cdot 15 + 14 \cdot 19 + 7 \cdot 23}{94} \approx 13,26$$

(triệu đồng)

Số trung bình thu thập trong một tháng của nhân viên công ty ^B là

$$\bar{x}_B = \frac{16 \cdot 7 + 40 \cdot 11 + 29 \cdot 15 + 15 \cdot 19 + 12 \cdot 23}{112} \approx 13,82$$

(triệu đồng)

Vậy nếu so sánh theo số trung bình của mẫu số liệu ghép nhóm thì thu nhập của nhân viên công ty ^A thấp hơn thu nhập của nhân viên công ty ^B.

Ta ước lượng một của thu nhập của nhân viên công ty ^A :

Nhóm một của mẫu số liệu là nhóm $[9;13)$.

Do đó: $u_m = 9; n_{m-1} = 12; n_m = 45; n_{m+1} = 16; u_{m+1} - u_m = 13 - 9 = 4$.

Mốt của mẫu số liệu ghép nhóm là

$$M_o = 9 + \frac{45 - 12}{(45 - 12) + (45 - 16)} \cdot 4 = \frac{345}{31} \approx 11,13.$$

Ta ước lượng một của thu nhập của nhân viên công ty ^B : Nhóm một của mẫu số liệu là nhóm $[9;13)$.

Do đó: $u_m = 9; n_{m-1} = 16; n_m = 40; n_{m+1} = 29; u_{m+1} - u_m = 13 - 9 = 4$.

Mốt của mẫu số liệu ghép nhóm là

$$M_o = 9 + \frac{40 - 16}{(40 - 16) + (40 - 29)} \cdot 4 = \frac{411}{35} \approx 11,74.$$

Vậy nếu so sánh theo một của mẫu số liệu ghép nhóm, thu nhập của nhân viên công ty ^A thấp hơn thu nhập của nhân viên công ty ^B.

Câu 51: Một bưu tá thống kê lại số bưu phẩm gửi đến một cơ quan mỗi ngày trong tháng 6/2022 ở bảng sau:

30	32	28	34	37	26	44	24	22	38
34	20	30	27	28	34	38	32	42	39
43	42	32	26	36	32	37	24	29	32

a) Tính số trung bình và một của mẫu số liệu trên.

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

Số bưu phẩm	$[20; 24]$	$[25; 29]$	$[30; 34]$	$[35; 39]$	$[40; 44]$
--------------------	------------	------------	------------	------------	------------

Số ngày	?	?	?	?	?
----------------	---	---	---	---	---

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

Lời giải

a) Số trung bình của mẫu số liệu là 32,4. Một của mẫu số liệu là 32.

b) Bảng tần số ghép nhóm.

Số bưu phẩm	[20; 24]	[25; 29]	[30; 34]	[35; 39]	[40; 44]
Số ngày	4	6	10	6	4

c) Do số bưu phẩm là số nguyên nên ta hiệu chỉnh lại bảng tần số ghép nhóm như sau:

Số bưu phẩm	[19,5; 24,5)	[24,5; 29,5)	[29,5; 34,5)	[34,5; 39,5)	[39,5; 44,5)
Giá trị đại diện	22	27	32	37	42
Số ngày	4	6	10	6	4

Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x} = \frac{4.22 + 6.27 + 10.32 + 6.37 + 4.42}{30} = 32.$$

$$[29,5; 34,5)$$

Nhóm chứa một của mẫu số liệu trên là nhóm

$$u_m = 29,5; n_{m-1} = 6; n_m = 10; n_{m+1} = 6; u_{m+1} - u_m = 34,5 - 29,5 = 5$$

Do đó:

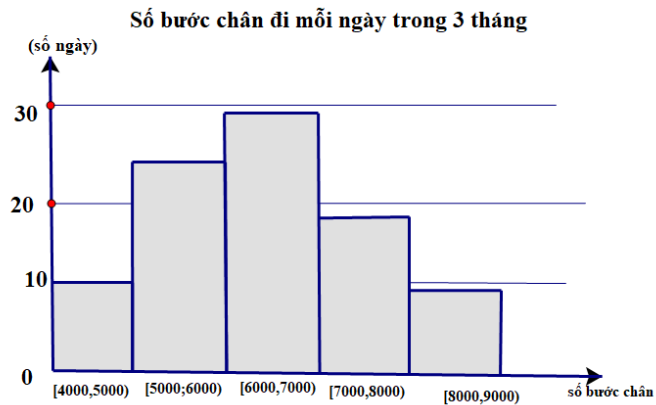
$$M_0 = 29,5 + \frac{10 - 6}{(10 - 6) + (10 - 6)} \cdot 5 = 32$$

Một của mẫu số liệu ghép nhóm là

Câu 52: Thảo thống kê lại số bước chân bạn đi mỗi ngày trong 3 tháng. Kết quả được biểu diễn ở biểu đồ ở bên.

a) Hãy lập bảng tần số ghép nhóm, kèm theo giá trị đại diện biểu diễn dữ liệu thống kê trên.

b) Hãy ước lượng số trung bình và một của mẫu số liệu ghép nhóm trên.

**Lời giải**

a) Bảng tần số ghép nhóm:

Số bước chân	Giá trị đại diện	Số ngày
$[4000; 5000)$	4500	12
$[5000; 6000)$	5500	24
$[6000; 7000)$	6500	29
$[7000; 8000)$	7500	18
$[8000; 9000)$	8500	9

b) Số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x} = \frac{12 \cdot 4500 + 24 \cdot 5500 + 29 \cdot 6500 + 18 \cdot 7500 + 9 \cdot 8500}{92} \approx 6369,57$$

Nhóm chứa một của mẫu số liệu trên là nhóm $[6000; 7000)$.

$$u_m = 6000; n_{m-1} = 24; n_m = 29; n_{m+1} = 18; u_{m+1} - u_m = 7000 - 6000 = 1000$$

Do đó:

Mốt của mẫu số liệu ghép nhóm là

$$M_o = 6000 + \frac{29 - 24}{(29 - 24) + (29 - 18)} \cdot 1000 = 6312,5.$$

Câu 53: Diện tích các tỉnh và thành phố khu vực Nam Bộ được thống kê ở bảng sau:

Tỉnh/ thành phố	Diện tích (km ²)	Tỉnh/ thành phố	Diện tích (km ²)
Bình Phước	6877	Vĩnh Long	1526
Tây Ninh	4041	Đồng Tháp	3384
Bình Dương	2695	An Giang	3537

Đồng Nai	5864	Kiên Giang	6349
Bà Rịa - Vũng Tàu	1981	Cần Thơ	1439
TP.Hồ Chí Minh	2061	Hậu Giang	1622
Long An	4495	Sóc Trăng	3312
Tiền Giang	2511	Bạc Liêu	2669
Bến Tre	2395	Cà Mau	5221
Trà Vinh	2358		

(Nguồn: Tổng cục Thống kê)

- a) Hãy tính diện tích trung bình của mỗi tỉnh/thành phố khu vực Nam Bộ.
- b) Dựa vào số liệu trên, hãy hoàn thiện bảng tần số ghép nhóm về diện tích các tỉnh khu vực Nam Bộ theo mẫu sau:

Diện tích (km^2)	[1000; 2500)	[2500; 4000)	[4000; 5500)	[5500; 7000)
Số tỉnh/thành phố	?	?	?	?

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

Lời giải

- a) Diện tích trung bình của mỗi tỉnh/thành phố khu vực Nam Bộ là

$$(6877 + 4041 + 2695 + 5864 + 1981 + 2061 + 4495 + 2511 + 2395 + 2358 + 1526 +$$

- b) Bảng tần số ghép nhóm:

Diện tích (km^2)	[1000; 2500)	[2500; 4000)	[4000; 5500)	[5500; 7000)
Số tỉnh/thành phố	7	6	3	3

- c) Bảng tần số ghép nhóm bao gồm giá trị đại diện của các nhóm như sau:

Diện tích (km^2)	[1000; 2500)	[2500; 4000)	[4000; 5500)	[5500; 7000)
Giá trị đại diện	1750	3250	4750	6250
Số tỉnh/ thành phố	7	6	3	3

Khi đó, số trung bình của mẫu số liệu ghép nhóm là

$$\bar{x} = \frac{7 \cdot 1750 + 6 \cdot 3250 + 3 \cdot 4750 + 3 \cdot 6250}{19} \approx 3407,89.$$

[1000; 2500)

Nhóm chứa một của mẫu số liệu trên là nhóm

$$u_m = 1000; n_{m-1} = 0; n_m = 7; n_{m+1} = 6; u_{m+1} - u_m = 2500 - 1000 = 1500$$

Do đó:

Mốt của mẫu số liệu ghép nhóm là

$$M_0 = 1000 + \frac{7-0}{(7-0)+(7-6)} \cdot 1500 = 2312,5.$$

Câu 54: Nhân ngày hội đọc sách, các học sinh của một trường trung học phổ thông mang sách cũ đến tặng thư viện trường và trao đổi với các bạn học sinh khác. Bảng sau thống kê số sách cũ mà các bạn học sinh lớp ¹¹B mang đến trường.

Số sách	[1;3]	[4;6]	[7;9]	[10;12]	[13;15]
Số học sinh	5	14	10	8	3

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

Lời giải

Bảng số liệu ghép nhóm hiệu chỉnh như sau:

Số sách	[0,5;3,5)	[3,5;6,5)	[6,5;9,5)	[9,5;12,5)	[12,5;15,5)
Số học sinh	5	14	10	8	3

Số trung bình của mẫu số liệu ghép nhóm là 7,25.

$$\frac{145}{26} \approx 5,58$$

Mốt của mẫu số liệu ghép nhóm là

Câu 55: Một kĩ thuật viên ghi lại cân nặng của 20 chi tiết máy ở bảng sau (đơn vị: gam):

5,63	5,58	5,42	5,58	5,56	5,54	5,55	5,40	5,60	5,56
5,46	5,51	5,58	5,48	5,61	5,50	5,54	5,64	5,43	5,63

a) Tính cân nặng trung bình của mỗi chi tiết máy.

$$[5,40;5,45)$$

b) Lập bảng tần số ghép nhóm của mẫu số liệu trên với nhóm đầu tiên là và ước lượng số trung bình của mẫu số liệu ghép nhóm trên.

Lời giải

$$5,54 \text{ g}$$

a) Cân nặng trung bình của mỗi chi tiết máy là

b) Bảng số liệu ghép nhóm của mẫu số liệu đã cho như sau:

Cân nặng	[5,40;5,45)	[5,45;5,50)	[5,50;5,55)	[5,55;5,60)	[5,60;5,65)
Số chi tiết	3	2	4	6	5

Số trung bình của mẫu số liệu ghép nhóm là 5,545.

Câu 56: Bảng sau thống kê số lượt chở khách mỗi ngày của một lái xe taxi trong 30 ngày.

15	13	7	5	18	13	11	9	10	8	14	11	16	10	9
13	11	12	13	15	12	13	6	8	17	13	6	18	12	13

a) Hãy tính số trung bình và một của mẫu số liệu trên.

$$[4,5;7,5)$$

b) Hãy lập bảng tần số ghép nhóm của mẫu số liệu trên với nhóm đầu tiên là .

c) Hãy ước lượng số trung bình và một của mẫu số liệu ghép nhóm đó.

Lời giải

a) Số trung bình mẫu số liệu là 11,7. Một của mẫu số liệu là 13.

b) Bảng tần số ghép nhóm của mẫu số liệu đã cho như sau:

Số lượt khách	$[4,5;7,5)$	$[7,5;10,5)$	$[10,5;13,5)$	$[13,5;16,5)$	$[16,5;19,5)$
Số ngày	4	6	13	4	3

c) Số trung bình của mẫu số liệu ghép nhóm là 11,6.

$$\text{Một của mẫu số liệu ghép nhóm là } \frac{189}{16} \approx 11,8$$

Câu 57: Bảng sau thống kê cân nặng (đơn vị: kg) của một số con ngan đực 88 ngày tuổi ở một trang trại.

4,60	4,62	4,64	4,65	4,67	4,67	4,68	4,68	4,70	4,70
4,70	4,70	4,71	4,71	4,72	4,73	4,74	4,76	4,77	4,77
4,77	4,78	4,78	4,80	4,82	4,84	4,84	4,85	4,87	4,89
4,89	4,90	4,92	4,92	4,93	4,94	4,94	4,95	4,97	4,97
4,97	4,99	4,99	5,01	5,02	5,03	5,04	5,05	5,06	5,07

$$[4,6;4,7)$$

a) Hãy lập bảng tần số ghép nhóm của mẫu số liệu trên với nhóm đầu tiên là .

b) Hãy ước lượng số trung bình và một của mẫu số liệu ghép nhóm trên.

Lời giải

a) Bảng tần số ghép nhóm của mẫu số liệu đã cho như sau:

Cân nặng (kg)	$[4,6;4,7)$	$[4,7;4,8)$	$[4,8;4,9)$	$[4,9;5,0)$	$[5,0;5,1)$
Số con ngan	8	15	8	12	7

b) Số trung bình của mẫu số liệu ghép nhóm là 4,84. Một của mẫu số liệu ghép nhóm là 4,75.

Câu 58: Bảng sau thống kê chiều cao (đơn vị: cm) của một số cây giống sau khi nảy mầm được 2 tuần.

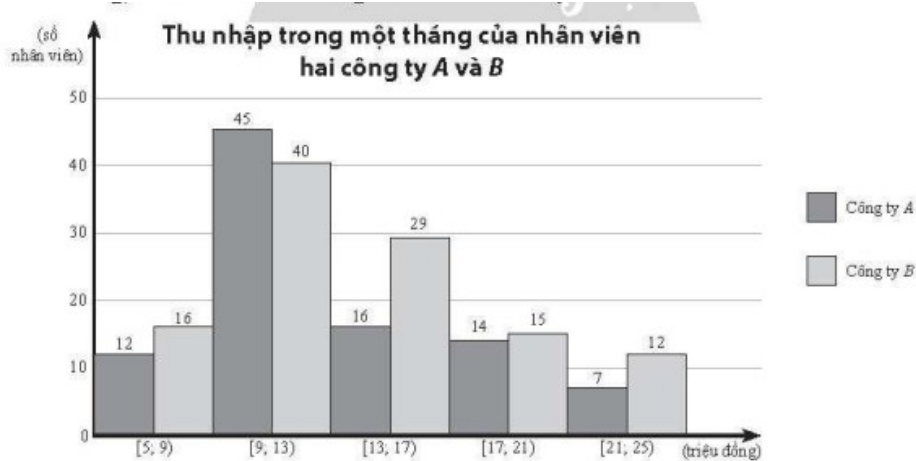
Chiều cao (cm)	$[6, 2; 6, 7)$	$[6, 7; 7, 2)$	$[7, 2; 7, 7)$	$[7, 7; 8, 2)$	$[8, 2; 8, 7)$
Số cây	10	21	28	12	9

Hãy ước lượng chiều cao trung bình và một của mẫu số liệu ghép nhóm trên.

Lời giải

Số trung bình của mẫu số liệu ghép nhóm là khoảng 7,38. Một của mẫu số liệu ghép nhóm là khoảng 7,35.

Câu 59: Thống kê lại thu nhập trong một tháng của nhân viên hai công ty A và B (đơn vị: triệu đồng) được thể hiện trong biểu đồ dưới đây.



Hãy so sánh thu nhập trung bình của nhân viên hai công ty theo số trung bình và một của mẫu số liệu ghép nhóm.

Lời giải

Ta lập bảng thống kê thu nhập của nhân viên hai công ty như sau:

Thu nhập (đơn vị: triệu đồng)	Giá trị đại diện	Số nhân viên công ty A	Số nhân viên công ty B
$[5; 9)$	7	12	16
$[9; 13)$	11	45	40
$[13; 17)$	15	16	29
$[17; 21)$	19	14	15
$[21; 25)$	23	7	12

Số trung bình thu thập trong một tháng của nhân viên công ty A là

$$\bar{x}_A = \frac{12 \cdot 7 + 45 \cdot 11 + 16 \cdot 15 + 14 \cdot 19 + 7 \cdot 23}{94} \approx 13,26$$

(triệu đồng)

Số trung bình thu thập trong một tháng của nhân viên công ty B là

$$\bar{x}_B = \frac{16 \cdot 7 + 40 \cdot 11 + 29 \cdot 15 + 15 \cdot 19 + 12 \cdot 23}{112} \approx 13,82$$

(triệu đồng)

Vậy nếu so sánh theo số trung bình của mẫu số liệu ghép nhóm thì thu nhập của nhân viên công ty A thấp hơn thu nhập của nhân viên công ty B .

Ta ước lượng một của thu nhập của nhân viên công ty A :

Nhóm một của mẫu số liệu là nhóm $[9;13)$.

Do đó: $u_m = 9; n_{m-1} = 12; n_m = 45; n_{m+1} = 16; u_{m+1} - u_m = 13 - 9 = 4$.

Mốt của mẫu số liệu ghép nhóm là

$$M_o = 9 + \frac{45 - 12}{(45 - 12) + (45 - 16)} \cdot 4 = \frac{345}{31} \approx 11,13.$$

Ta ước lượng một của thu nhập của nhân viên công ty B : Nhóm một của mẫu số liệu là nhóm $[9;13)$.

Do đó: $u_m = 9; n_{m-1} = 16; n_m = 40; n_{m+1} = 29; u_{m+1} - u_m = 13 - 9 = 4$.

Mốt của mẫu số liệu ghép nhóm là

$$M_o = 9 + \frac{40 - 16}{(40 - 16) + (40 - 29)} \cdot 4 = \frac{411}{35} \approx 11,74.$$

Vậy nếu so sánh theo một của mẫu số liệu ghép nhóm, thu nhập của nhân viên công ty A thấp hơn thu nhập của nhân viên công ty B .

Câu 60: Một công ty cung cấp nước sạch thống kê lượng nước các hộ gia đình trong một khu vực tiêu thụ trong một tháng ở bảng sau:

Lượng nước tiêu thụ (m^3)	$[3;6)$	$[6;9)$	$[9;12)$	$[12;15)$	$[15;18)$
Số hộ gia đình	24	57	42	29	8

a) Hãy ước lượng số trung bình, một và trung vị của mẫu số liệu ghép nhóm trên.

b) Công ty muốn gửi một thông báo khuyến nghị tiết kiệm nước đến 25% các hộ gia đình có lượng nước tiêu thụ cao nhất. Hỏi công ty nên gửi đến các hộ tiêu thụ từ bao nhiêu mét khối nước trở lên?

Lời giải

a) Cỡ mẫu $n = 160$.

Bảng tần số ghép nhóm của mẫu số liệu trên như sau:

Lượng nước tiêu thụ (m^3)	$[3;6)$	$[6;9)$	$[9;12)$	$[12;15)$	$[15;18)$
Giá trị đại diện	4,5	7,5	10,5	13,5	16,5
Số hộ gia đình	24	57	42	29	8

Số trung bình của mẫu số liệu ghép nhóm đã cho là

$$\bar{x} = \frac{1}{160}(24 \cdot 4,5 + 57 \cdot 7,5 + 42 \cdot 10,5 + 29 \cdot 13,5 + 8 \cdot 16,5) = 9,375.$$

$[6;9)$

Nhóm chứa một của mẫu số liệu ghép nhóm trên là nhóm .

$$u_m = 6; n_m = 57; n_{m-1} = 24; n_{m+1} = 42; u_{m+1} = 9$$

Do đó:

$$M_o = 6 + \frac{(57 - 24)}{(57 - 24) + (57 - 42)} \cdot (9 - 6) = 8,0625$$

Một của mẫu số liệu là .

$$x_1; x_2; \dots; x_{160}$$

Gọi là mẫu số liệu được xếp theo thứ tự không giảm.

$$x_1, \dots, x_{24} \in [3;6); x_{25}, \dots, x_{81} \in [6;9); x_{82}, \dots, x_{123} \in [9;12)$$

Ta có: ;

$$x_{124}, \dots, x_{152} \in [12;15); x_{153}, \dots, x_{160} \in [15;18).$$

Cỡ mẫu $n = 160$ là số chẵn nên trung vị là $M_e = \frac{1}{2}(x_{80} + x_{81})$. Do x_{80} và x_{81} thuộc nhóm $[6;9)$

$$M_e = 6 + \frac{\frac{160}{2} - 24}{57} \cdot (9 - 6) \approx 8,95$$

nên trung vị của mẫu số liệu là .

25%

b) các hộ gia đình có lượng nước tiêu thụ cao nhất có lượng nước tiêu thụ không nhỏ hơn

$$Q_3$$

, với là tứ phân vị thứ ba của mẫu số liệu.

Tứ phân vị thứ ba của dãy số liệu $x_1; x_2; \dots; x_{160}$ là $\frac{1}{2}(x_{120} + x_{121})$. Do x_{120} và x_{121} thuộc nhóm $[9;12)$

nên tứ phân vị thứ ba của mẫu số liệu là

$$Q_3 = 9 + \frac{\frac{160 \cdot 3}{4} - (24 + 57)}{42} \cdot (12 - 9) \approx 11,79.$$

Vậy công ty nên gửi thông báo tiết kiệm nước đến các hộ gia đình có lượng nước tiêu thụ từ $11,79 \text{ m}^3$

nước

trở lên.

Câu 61: Bảng sau thống kê khối lượng một số quả măng cụt được lựa chọn ngẫu nhiên trong một thùng hàng.

Khối lượng (gam)	$[80;82)$	$[82;84)$	$[84;86)$	$[86;88)$	$[88;90)$
Số quả	18	20	24	15	13

a) Hãy ước lượng số trung bình, mốt và trung vị của mẫu số liệu ghép nhóm trên.

b) Người ta muốn chia măng cụt trong thùng ra làm ba loại theo cân nặng, bao gồm: loại nhỏ, $25\%, 50\%$ và 25% loại vừa và loại to. Các loại này lần lượt chiếm khoảng và số măng cụt trong thùng. Hãy xác định ngưỡng cân nặng để phân loại quả.

Lời giải

a) Cỡ mẫu $n = 90$.

Bảng tần số ghép nhóm của mẫu số liệu trên như sau:

Khối lượng (gam)	$[80;82)$	$[82;84)$	$[84;86)$	$[86;88)$	$[88;90)$
Khối lượng đại diện	81	83	85	87	89
Số quả	18	20	24	15	13

Số trung bình của mẫu số liệu ghép nhóm trên là

$$\bar{x} = \frac{1}{90}(18 \cdot 81 + 20 \cdot 83 + 24 \cdot 85 + 15 \cdot 87 + 13 \cdot 89) = \frac{254}{3} \approx 84,67.$$

$[84;86)$

Nhóm chứa mốt của mẫu số liệu là nhóm .

$$u_m = 84; n_m = 24; n_{m-1} = 20; n_{m+1} = 15; u_{m+1} = 86$$

Do đó:

$$M_s = 84 + \frac{(24-20)}{(24-20)+(24-15)} (86-84) \approx 84,62$$

Mốt của mẫu số liệu là

$$x_1; x_2; \dots; x_{90}$$

Gọi là mẫu số liệu được xếp theo thứ tự không giảm.

$$x_1, \dots, x_{18} \in [80; 82); x_{19}, \dots, x_{38} \in [82; 84); x_{39}, \dots, x_{62} \in [84; 86);$$

Ta có:

$$x_{63}, \dots, x_{77} \in [86; 88); x_{78}, \dots, x_{90} \in [88; 90)$$

Cỡ mẫu $n = 90$ là số chẵn nên trung vị là $\frac{x_{45} + x_{46}}{2}$. Do x_{45} và x_{46} thuộc nhóm $[84; 86)$ nên trung vị của mẫu số liệu là

$$M_e = 84 + \frac{\frac{90}{2} - (18 + 20)}{24} \cdot (86 - 84) \approx 84,58.$$

$$Q_1, Q_3$$

b) Gọi lần lượt là tứ phân vị thứ nhất và thứ ba của mẫu số liệu. Theo đề bài, ta có:

$$Q_1$$

Măng cụt loại nhỏ có cân nặng nhỏ hơn

$$[Q_1; Q_3)$$

Măng cụt loại vừa có cân nặng trong

$$Q_3$$

Măng cụt loại to có cân nặng không nhỏ hơn

Tứ phân vị thứ nhất của dãy số liệu $x_1; x_2; \dots; x_{90}$ là $x_{23} \in [82; 84)$. Tứ phân vị thứ ba của dãy số liệu $x_1; x_2; \dots; x_{90}$ là $x_{68} \in [86; 88)$.

$$Q_1 = 82 + \frac{\frac{90}{4} - 18}{20} \cdot (84 - 82) = 82,45$$

Do đó, tứ phân vị thứ nhất là ; tứ phân vị thứ ba là

$$Q_3 = 86 + \frac{\frac{90,3}{4} - (18 + 20 + 24)}{15} \cdot (88 - 86) \approx 86,73$$

$$[80; 82,45$$

Vậy măng cụt loại nhỏ có khối lượng (tính theo gam) thuộc). Măng cụt loại vừa có khối lượng (tính theo gam) thuộc

$$[82,45; 86,73$$

khối lượng (tính theo gam) thuộc). Măng cụt loại to có khối lượng (tính theo gam) thuộc

$$[86,73; 90)$$

Câu 62: Thời gian sử dụng điện thoại trong một ngày của 30 sinh viên được ghi lại ở bảng sau (đơn vị: phút).

85	195	187	198	43	223	280	71	205	277
298	142	162	89	167	122	175	168	148	253
234	187	85	193	224	233	117	81	39	85

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

$$[0; 60)$$

b) Tổng hợp lại dãy số liệu trên vào bảng tần số ghép nhóm với nhóm đầu tiên là . Hãy ước lượng các tứ phân vị của mẫu số liệu ghép nhóm.

Lời giải

a) Sắp xếp mẫu số liệu theo thứ tự không giảm ta được:

39; 43; 71; 81; 85; 85; 85; 89; 117; 122; 142; 148; 162; 167; 168; 175; 187; 187; 193; 195; 198; 205; 223; 224; 233; 234; 280; 298

$$n = 30$$

Cỡ mẫu là số chẵn nên trung vị của mẫu số liệu là

$$Q_2 = \frac{1}{2}(x_{15} + x_{16}) = \frac{1}{2}(168 + 175) = 171,5.$$

$$Q_1 = x_8 = 89$$

Tứ phân vị thứ nhất là

$$Q_3 = x_{23} = 223$$

Tứ phân vị thứ ba là

b) Bảng tần số ghép nhóm của mẫu số liệu như sau:

Thời gian (phút)	$[0; 60)$	$[60; 120)$	$[120; 180)$	$[180; 240)$	$[240; 300)$
Số sinh viên	2	7	7	10	4

$$x_1; x_2; \dots; x_{30}$$

Gọi là mẫu số liệu được xếp theo thứ tự không giảm.

Ta có:

$$x_1, x_2 \in [0; 60); x_3, \dots, x_9 \in [60; 120); x_{10}, \dots, x_{16} \in [120; 180)$$

;

$$x_{17}, \dots, x_{26} \in [180; 240); x_{27}, \dots, x_{30} \in [240; 300).$$

$$\frac{1}{2}(x_{15} + x_{16})$$

$$x_{15} \quad x_{16}$$

Tứ phân vị thứ hai của mẫu số liệu ghép nhóm là . Do và thuộc nhóm $[120; 180)$

nên tứ phân vị thứ hai của mẫu số liệu ghép nhóm là

$$Q_2 = 120 + \frac{\frac{30}{2} - (2 + 7)}{7} \cdot (180 - 120) = \frac{1200}{7} \approx 171,43.$$

$$x_8 \in [60; 120)$$

Tứ phân vị thứ nhất của mẫu là .

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

$$Q_1 = 60 + \frac{\frac{30}{4} - 2}{7} \cdot (120 - 60) = \frac{750}{7} \approx 107,14$$

$$x_{23} \in [180; 240)$$

Tứ phân vị thứ ba của mẫu số liệu là .

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

$$Q_3 = 180 + \frac{\frac{3 \cdot 30}{4} - (2 + 7 + 7)}{10} \cdot (240 - 180) = 219$$

Câu 63: Một nhóm gồm 45 học sinh làm một bài kiểm tra trắc nghiệm gồm 40 câu hỏi. Số câu trả lời đúng của mỗi bạn được ghi lại ở bảng sau:

24	35	37	24	30	23	21	39	28	20	32	37	17	40	34
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

27	34	30	21	26	26	38	37	16	35	19	20	22	25	38
34	29	39	40	36	18	31	24	36	33	24	24	36	26	37

a) Tìm các 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:

Số câu trả lời đúng	[16; 20]	[21; 25]	[26; 30]	[31; 35]	[36; 40]
Số học sinh	?	?	?	?	?

c) Hãy ước lượng các tứ phân vị của mẫu số liệu ghép nhóm trên.

Lời giải

a) Mẫu số liệu đã cho được sắp xếp theo thứ tự không giảm như sau:

16; 17; 18; 19; 20; 20; 21; 21; 22; 23; 24; 24; 24; 24; 24;
 25; 26; 26; 26; 27; 28; 29; 30; 30; 31; 32; 33; 34; 34; 34;
 35; 35; 36; 36; 36; 37; 37; 37; 37; 38; 38; 39; 39; 40; 40

$$n = 45$$

$$Q_2 = x_{23} = 30$$

Cỡ mẫu là $n = 45$ là số lẻ nên trung vị của mẫu số liệu là

$$Q_1 = \frac{1}{2}(x_{11} + x_{12}) = \frac{1}{2}(24 + 24) = 24$$

Tứ phân vị thứ nhất là

$$Q_3 = \frac{1}{2}(x_{34} + x_{35}) = \frac{1}{2}(36 + 36) = 36$$

Tứ phân vị thứ ba là

b) Bảng tần số ghép nhóm của mẫu số liệu trên như sau

Số câu trả lời đúng	[16; 20]	[21; 25]	[26; 30]	[31; 35]	[36; 40]
Số học sinh	6	10	8	8	13

c) Do số câu trả lời đúng của học sinh là số nguyên nên ta hiệu chỉnh lại bảng số liệu như sau:

Số câu trả lời đúng	[15, 5; 20, 5)	[20, 5; 25, 5)	[25, 5; 30, 5)	[30, 5; 35, 5)	[35, 5; 40, 5)
Số học sinh	6	10	8	8	13

$$x_1, x_2, \dots, x_{45}$$

Gọi $x_1, x_2, \dots, x_{45}$ là số câu trả lời đúng của 45 học sinh được xếp theo thứ tự không giảm.

Ta có: $x_1, \dots, x_6 \in [15, 5; 20, 5); x_7, \dots, x_{16} \in [20, 5; 25, 5); x_{17}, \dots, x_{24} \in [25, 5; 30, 5);$
 $x_{25}, \dots, x_{32} \in [30, 5; 35, 5); x_{33}, \dots, x_{45} \in [35, 5; 40, 5)$

Tứ phân vị thứ hai của mẫu số liệu $x_1; x_2; \dots; x_{45}$ là $x_{23} \in [25, 5; 30, 5)$. Do đó, tứ phân vị thứ hai của mẫu số liệu ghép nhóm là

$$Q_2 = 25,5 + \frac{\frac{45}{2} - (6+10)}{8} \cdot (30,5 - 25,5) = 29,5625.$$

Tứ phân vị thứ nhất của mẫu số liệu $x_1; x_2; \dots; x_{45}$ là $\frac{1}{2}(x_{11} + x_{12})$.

Do x_{11} và $x_{12} \in [20, 5; 25, 5)$ thuộc nhóm nên tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là

$$Q_1 = 20,5 + \frac{\frac{45}{4} - 6}{10} \cdot (25,5 - 20,5) = 23,125.$$

Tứ phân vị thứ ba của mẫu số liệu $x_1; x_2; \dots; x_{45}$ là $\frac{1}{2}(x_{34} + x_{35})$.

Do x_{34} và $x_{35} \in [35, 5; 40, 5)$ thuộc nhóm nên tứ phân vị thứ ba của mẫu số liệu ghép nhóm là

$$Q_3 = 35,5 + \frac{\frac{3 \cdot 45}{4} - (6+10+8+8)}{13} \cdot (40,5 - 35,5) \approx 36,173.$$

Câu 64: Một trang báo điện tử thống kê thời gian người sử dụng đọc thông tin trên trang trong mỗi lần truy cập ở bảng sau:

Thời gian đọc (phút)	$[0; 2)$	$[2; 4)$	$[4; 6)$	$[6; 8)$	$[8; 10)$
Số lượt truy cập	45	34	23	18	5

Hãy ước lượng các tứ phân vị của mẫu số liệu ghép nhóm trên.

Lời giải

$$Q_1 = \frac{25}{18}; Q_2 = \frac{103}{34}; Q_3 = \frac{243}{46}$$

Câu 65: Người ta thống kê tốc độ của một số xe ô tô di chuyển qua một trạm kiểm soát trên đường cao tốc trong một khoảng thời gian ở bảng sau:

Tốc độ (km/h)	[75;80)	[80;85)	[85;90)	[90;95)	[95;100)
Số xe	5	12	18	24	19

Hãy ước lượng các tứ phân vị của mẫu số liệu ghép nhóm trên.

Lời giải

$$Q_1 = \frac{3085}{36}; Q_2 = \frac{545}{6}; Q_3 = \frac{4555}{48}$$

Câu 66: Thâm niên công tác của các công nhân hai nhà máy A và B .

Thâm niên công tác (năm)	[0;5)	[5;10)	[10;15)	[15;20)	[20;25)
Số công nhân nhà máy A	35	13	12	12	8
Số công nhân nhà máy B	14	26	24	11	5

a) Hãy so sánh thâm niên công tác của nhân viên hai nhà máy theo số trung bình và trung vị.

b) Hãy ước lượng tứ phân vị thứ nhất và thứ ba của hai mẫu số liệu ghép nhóm trên.

Lời giải

a) So sánh theo số trung bình: $\bar{x}_A = 9,0625; \bar{x}_B = 10,4375$, suy ra $\bar{x}_A < \bar{x}_B$.

$$M_c(A) = \frac{90}{13}; M_c(B) = 10, M_c(A) < M_c(B)$$

So sánh theo trung vị:

$$Q_1(A) = \frac{20}{7}, Q_3(A) = 15; Q_1(B) = \frac{80}{13}, Q_3(B) = \frac{85}{6}$$

b)

Câu 67: Thầy giáo thống kê lại số lần kéo xà đơn của các học sinh nam khối 11 ở bảng sau:

Số lần	[6;10]	[11;15]	[16;20]	[21;25]	[26;30]
Số học sinh	35	54	32	17	5

a) Hãy ước lượng số trung bình, mốt và trung vị của mẫu số liệu ghép nhóm trên.

$$25\%$$

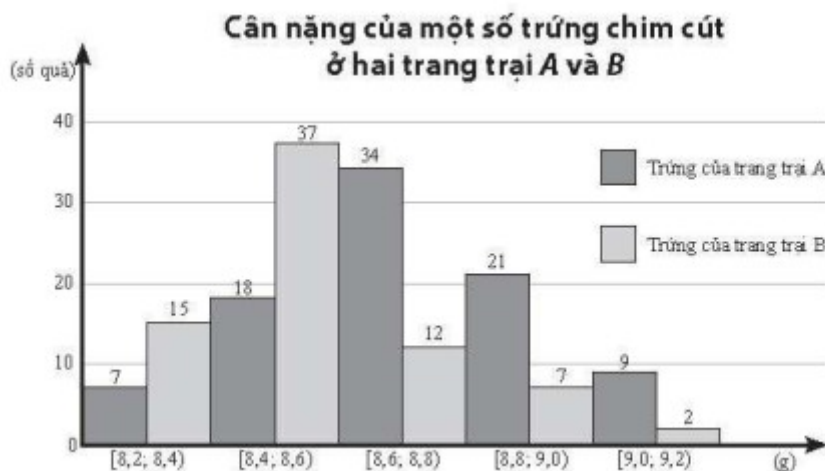
b) Thầy giáo dự định chọn học sinh có số lần kéo thấp nhất để bồi dưỡng thể lực thêm. Thầy giáo nên chọn học sinh có thành tích kéo xà đơn dưới bao nhiêu lần để bồi dưỡng thể lực?

Lời giải

$$a) \bar{x} = \frac{2089}{143}; M_o = \frac{1051}{82}; M_c = \frac{1499}{108}$$

- b) $Q_1 = \frac{797}{72} \approx 11,07$. Thầy giáo nên chọn các bạn có thành tích kéo xà dưới 12 lần để bồi dưỡng thể lực thêm.

Câu 68: Kết quả kiểm tra cân nặng của một số quả trứng chim cút được lựa chọn ngẫu nhiên ở hai trang trại chăn nuôi A và B được biểu diễn ở biểu đồ sau (đơn vị: g).



- a) Hãy so sánh cân nặng của trứng chim cút của hai trang trại A và B theo số trung bình và trung vị.
- b) Hãy ước lượng tứ phân vị thứ nhất và tứ phân vị thứ ba của cân nặng trứng chim cút của trang trại A .

Lời giải

- a) Từ biểu đồ đã cho, ta lập được bảng số liệu ghép nhóm sau:

Cân nặng (gam)	$[8,2; 8,4)$	$[8,4; 8,6)$	$[8,6; 8,8)$	$[8,8; 9,0)$	$[9,0; 9,2)$
Số quả trứng của trang trại A	7	18	34	21	9
Số quả trứng của trang trại B	15	37	12	7	2

Từ đó, ta có bảng thống kê số quả trứng chim cút của hai trang trại theo giá trị đại diện như sau:

Cân nặng đại diện (gam)	8,3	8,5	8,7	8,9	9,1
Số quả trứng của trang trại A	7	18	34	21	9
Số quả trứng của trang trại B	15	37	12	7	2

Đối với trang trại A :

Cỡ mẫu $n_A = 89$.

Cân nặng trung bình của mỗi quả trứng của mẫu số liệu ghép nhóm là

$$\bar{x}_A = \frac{8,3.7 + 8,5.18 + 8,7.34 + 8,9.21 + 9,1.9}{89} \approx 8,72 \text{ (g)}.$$

Trung vị của dãy số liệu là số liệu thứ 45 theo thứ tự không giảm của dãy và thuộc nhóm $[8,6;8,8)$

$$M_e(A) = 8,6 + \frac{\frac{89}{2} - 25}{34} (8,8 - 8,6) \approx 8,71.$$

Trung vị của mẫu số liệu ghép nhóm là

Đối với trang trại B :

$$n_B = 73$$

Cỡ mẫu

Cân nặng trung bình của mỗi quả trứng của mẫu số liệu ghép nhóm là

$$\bar{x}_B = \frac{8,3.15 + 8,5.37 + 8,7.12 + 8,9.7 + 9,1.2}{73} \approx 8,55.$$

Trung vị của dãy số liệu là số liệu thứ 37 theo thứ tự không giảm của dãy và thuộc nhóm $[8,4;8,6)$

$$M_e(B) = 8,4 + \frac{\frac{73}{2} - 15}{37} \cdot (8,6 - 8,4) \approx 8,52.$$

Trung vị của mẫu số liệu ghép nhóm là

$$\bar{x}_A > \bar{x}_B \quad M_e(A) > M_e(B)$$

Ta thấy và . Vậy khi so sánh theo số trung bình hay theo trung vị, cân nặng của trứng chim cút của trang trại A đều lớn hơn cân nặng của trứng chim cút của trang trại B .

b) Đối với trang trại A :

Tứ phân vị thứ nhất của mẫu số liệu là trung bình cộng của số liệu thứ 22 và 23 theo thứ tự $[8,4;8,6)$

không giảm của dãy và thuộc nhóm , nên tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là

$$Q_1(A) = 8,4 + \frac{\frac{1,89}{4} - 7}{18} \cdot (8,6 - 8,4) \approx 8,57.$$

Tứ phân vị thứ ba của mẫu số liệu ghép nhóm là trung bình cộng của số liệu thứ 67 và 68 theo $[8,8;9,0)$

thứ tự không giảm của dãy và thuộc nhóm , nên tứ phân vị thứ ba của mẫu số liệu

$$Q_3(A) = 8,8 + \frac{\frac{3,89}{4} - 59}{21} \cdot (9,0 - 8,8) \approx 8,87$$

ghép nhóm là .

Câu 69: Một công ty bảo hiểm thống kê lại độ tuổi các khách hàng mua bảo hiểm xe ô tô ở bảng sau:

Độ tuổi	[25; 30)	[30; 35)	[35; 40)	[40; 45)	[45; 50)	[50; 55)
Số khách hàng	25	38	62	42	37	29

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

Lời giải

$$\bar{x} \approx 39,97; M_o = \frac{415}{11}; Q_1 = \frac{275}{8}, Q_2 = M_e = \frac{4875}{124}, Q_3 = \frac{6815}{148}$$

Câu 70: Các bạn học sinh một lớp thống kê số túi nhựa mà gia đình bạn đó sử dụng trong một tuần. Kết quả được tổng hợp lại ở bảng sau:

Số túi	[5; 9]	[10; 14]	[15; 19]	[20; 24]	[25; 29]
Số gia đình	8	15	12	7	2

a) Hãy ước lượng số trung bình và một của mẫu số liệu trên.

25%

b) Cô giáo dự định trao danh hiệu "Gia đình xanh" cho gia đình các bạn sử dụng ít túi nhựa nhất. Cô giáo nên trao danh hiệu cho các gia đình dùng không quá bao nhiêu túi nhựa?

Lời giải

a) Số trung bình của mẫu số liệu ghép nhóm là khoảng 14,73.

Một của mẫu số liệu là 13.

$$Q_1 = \frac{21}{2} = 10,5$$

b) Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là .

Do đó, cô giáo nên trao danh hiệu cho gia đình các bạn dùng không quá 10 túi nhựa.

Câu 71: Bảng sau thống kê doanh số bán hàng của các nhân viên một trung tâm thương mại trong một ngày.

Doanh số (triệu đồng)	[20; 30)	[30; 40)	[40; 50)	[50; 60)	[60; 70)
----------------------------	----------	----------	----------	----------	----------

Số nhân viên	4	8	12	7	5
---------------------	---	---	----	---	---

a) Hãy ước lượng số trung bình, một và trung vị của mẫu số liệu ghép nhóm trên.

25%

b) Trung tâm thương mại dự định sẽ thưởng cho số nhân viên có doanh số bán hàng cao nhất. Theo mẫu số liệu trên, trung tâm thương mại nên khen thưởng các nhân viên có doanh số bán hàng ít nhất là bao nhiêu?

Lời giải

$$\bar{x} = \frac{815}{18}; M_o = \frac{400}{9}; M_c = 45$$

a)

$$Q_3 = \frac{380}{7} \approx 54,29$$

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

Do đó, trung tâm thương mại nên khen thưởng các nhân viên có doanh số bán hàng một ngày ít nhất là 54,29 triệu đồng.

Câu 72: Một cửa hàng sách thống kê số truyện thiếu nhi bán được trong hai tháng ở bảng sau:

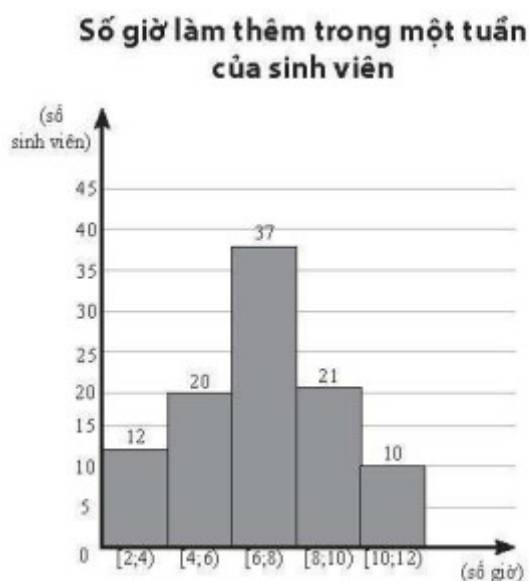
Số sách	[14; 20]	[21; 27]	[28; 34]	[35; 41]	[42; 48]
Số ngày	5	7	25	15	9

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

Lời giải

$$\bar{x} = \frac{2003}{61}; M_o = 32; Q_1 = 28,41; Q_2 = M_c = 32,68; Q_3 = \frac{4967}{90}$$

Câu 73: Kết quả điều tra về số giờ làm thêm trong một tuần của 100 sinh viên được cho ở biểu đồ bên. Hãy ước lượng số trung bình, một và các tứ phân vị của số liệu đó.



Lời giải

Từ mẫu số liệu ghép nhóm, ta có bảng thống kê số giờ làm thêm trong một tuần của 100 sinh viên như sau:

Số giờ làm thêm	$[2; 4)$	$[4; 6)$	$[6; 8)$	$[8; 10)$	$[10; 12)$
Số giờ làm thêm đại diện	3	5	7	9	11
Số sinh viên	12	20	37	21	10

$$n = 100$$

Cỡ mẫu .

Số trung bình của mẫu số liệu ghép nhóm trên là

$$\bar{x} = \frac{3.12 + 5.20 + 7.37 + 9.21 + 11.10}{100} = 6,94.$$

$$[6; 8)$$

Nhóm chứa một của mẫu số liệu trên là nhóm .

$$u_m = 6; n_m = 37; n_{m-1} = 20; n_{m+1} = 21; u_{m+1} = 8$$

Do đó: .

$$M_0 = 6 + \frac{37 - 20}{(37 - 20) + (37 - 21)} \cdot (8 - 6) \approx 7,03.$$

Vậy một của mẫu số liệu ghép nhóm là

$$x_1; x_2; \dots; x_{100}$$

Gọi là mẫu số liệu được xếp theo thứ tự không giảm.

$$x_1; x_2; \dots; x_{100} \quad \frac{1}{2}(x_{50} + x_{51})$$

Tứ phân vị thứ hai của dãy số liệu là .

$$x_{50} \quad x_{51} \quad [6; 8) \quad Q_2 = 6 + \frac{\frac{2.100}{4} - 32}{37} \cdot (8 - 6) \approx 6,97$$

Do và thuộc nhóm nên .

$$x_1; x_2; \dots; x_{100} \quad \frac{1}{2}(x_{25} + x_{26})$$

Tứ phân vị thứ nhất của dãy số liệu là .

$$x_{25} \quad x_{26} \quad [4; 6) \quad Q_1 = 4 + \frac{\frac{1.100}{4} - 12}{20} \cdot (6 - 4) = 5,3$$

Do và thuộc nhóm nên .

Tứ phân vị thứ ba của dãy số liệu $x_1; x_2; \dots; x_{100}$ là $\frac{1}{2}(x_{75} + x_{76})$.

Do x_{75} và x_{76} thuộc nhóm $[8; 10)$ nên $Q_3 = 8 + \frac{\frac{3 \cdot 100}{4} - 69}{21} \cdot (10 - 8) \approx 8,57$.

Câu 74: Độ bão hoà oxygen trong máu (còn được gọi là chỉ số SpO_2) biểu thị cho tỉ lệ hemoglobin có oxygen trên tổng lượng hemoglobin trong máu. Chỉ số SpO_2 (đơn vị đo là %) từ 97- 99 là oxygen trong máu tốt, 94- 96 là oxygen trong máu trung bình, 90- 93 là oxygen trong máu thấp, dưới 90 là trường hợp cấp cứu trên lâm sàng (Theo: Vinmec.com). Đo chỉ số SpO_2 ở một số bệnh nhân Covid-19 người ta thu được kết quả sau:

SpO_2 (%)	90-93	94-96	97-99
Số bệnh nhân	12	31	7

- Cho biết các nhóm số liệu và tần số tương ứng.
- Tính số trung bình, trung vị và giải thích ý nghĩa của các giá trị thu được.

Lời giải

- Có 3 nhóm số liệu gồm 90-93, 94-96, 97-99 với tần số tương ứng là 12, 31, 7.
- Trước hết, ta hiệu chỉnh các nhóm số liệu và thu được bảng thống kê sau:

SpO_2 (%)	$[89,5; 93,5)$	$[93,5; 96,5)$	$[96,5; 99,5)$
Số bệnh nhân	12	31	7

Các giá trị đại diện cho các nhóm số liệu tương ứng là

$$\frac{(89,5 + 93,5)}{2} = 91,5; \quad \frac{93,5 + 96,5}{2} = 95; \quad \frac{96,5 + 99,5}{2} = 98$$

Cỡ mẫu $n = 12 + 31 + 7 = 50$. Do đó số trung bình là: $\bar{x} = \frac{1}{50}(12 \cdot 91,5 + 31 \cdot 95 + 7 \cdot 98) = 94,58$

Do có $\frac{n}{2} = 25$ giá trị nhỏ hơn trung vị nên trung vị thuộc nhóm $[93,5; 96,5)$.

Ta có $a_p = 93,5$; $a_{p+1} = 96,5$; $m_1 + \dots + m_{p-1} = 12$; $m_p = 31$.

Do đó, trung vị là $M_c = 93,5 + \frac{25 - 12}{31} \cdot 3 \approx 94,76$.

Như vậy, chỉ số SpO_2 trung bình của 50 bệnh nhân là 94,58; có 25 bệnh nhân có chỉ số SpO_2 nhỏ hơn 94,76 và 25 bệnh nhân có chỉ số SpO_2 lớn hơn 94,76.

Nhận xét. Nếu chỉ tính số trung bình của mẫu số liệu ghép nhóm thì ta không cần hiệu chỉnh

nhóm rời rạc $k_1 - k_2$, chọn giá trị đại diện là $\frac{k_1 + k_2}{2}$.

Câu 75: Mức thưởng tết (triệu đồng) mà các công nhân một nhà máy nhận được như sau:

Mức thưởng tết	[5;10)	[10;15)	[15;20)	[20;25)
Số công nhân	13	35	47	25

Tìm một của mẫu số liệu ghép nhóm. Cho biết ý nghĩa của giá trị thu được.

Lời giải

Cỡ mẫu: $n = 13 + 35 + 47 + 25 = 120$. Số công nhân có mức thưởng tết từ 15 đến dưới 20 triệu đồng là nhiều nhất nên nhóm chứa một là nhóm $[15;20)$.

Ta có: $a_j = 15$; $m_j = 47$; $m_{j-1} = 35$; $m_{j+1} = 25$; $h = 5$.

$$M_o = 15 + \frac{(47 - 35)}{(47 - 35) + (47 - 25)} \cdot 5 \approx 16,76$$

Do đó, một của mẫu số liệu là

Ý nghĩa. Số công nhân được mức thưởng tết khoảng 16,76 triệu đồng là cao nhất.

Câu 76: Số điểm một cầu thủ bóng rổ ghi được trong 20 trận đấu được cho ở bảng sau:

Điểm số	[6;10]	[11;15]	[16;20]	[21;25]
Số trận	3	9	2	6

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 từ bảng tần số ghép nhóm trên.

Lời giải

a) Tứ phân vị thứ nhất là: 11.

Tứ phân vị thứ hai là: 14.

Tứ phân vị thứ ba là: 21,5.

b)

c) Vì số trận là số nguyên nên ta hiệu chỉnh lại như sau:

Điểm số	[5,5;10,5)	[10,5;15,5)	[15,5;20,5)	[20,5;25,5)
Số trận	3	9	2	6

Gọi $x_1; x_2; x_3; \dots; x_{20}$ lần lượt là số trận theo thứ tự không gian.

Do $x_1, \dots, x_3 \in \text{I}; x_4, \dots, x_{12} \in \text{II}; x_{13}, x_{14} \in \text{III}; x_{15}, \dots, x_{20} \in \text{IV}$

$$\frac{1}{2}(x_{10} + x_{11}) \quad [10,5;15,5)$$

Tứ phân vị thứ hai của dãy số liệu là thuộc nhóm nên tứ phân vị thứ

$$Q_2 = 10,5 + \frac{\frac{20}{2} - 3}{9} (15,5 - 10,5) = 14,4$$

hai của mẫu số liệu là .

$$\frac{1}{2}(x_5 + x_6) \quad [10,5;15,5)$$

Tứ phân vị thứ nhất của dãy số liệu là thuộc nhóm nên tứ phân vị thứ

$$Q_1 = 10,5 + \frac{\frac{20}{4} - 3}{9} (15,5 - 10,5) = 11,6$$

nhất của mẫu số liệu là .

$$\frac{1}{2}(x_{15} + x_{16}) \quad [20,5;25,5)$$

Tứ phân vị thứ ba của dãy số liệu là thuộc nhóm nên tứ phân vị thứ

$$Q_3 = 20,5 + \frac{\frac{3 \cdot 20}{4} - 14}{6} (25,5 - 20,5) = 21,3$$

ba của mẫu số liệu là .

Câu 77: Cân nặng của một số lợn con mới sinh thuộc hai giống A và B được cho ở biểu đồ dưới đây (đơn vị: kg).



a) Hãy so sánh cân nặng của lợn con mới sinh giống A và giống B theo số trung bình và trung vị.

b) Hãy ước lượng tứ phân vị thứ nhất và thứ ba của cân nặng lợn con mới sinh giống A và của cân nặng lợn con mới sinh giống B.

Lời giải

Cân nặng của lợn con giống A và giống B được thống kê như bảng sau:

Cân nặng (kg)	[1,0;1,1)	[1,1;1,2)	[1,2;1,3)	[1,3;1,4)
Giá trị đại diện	1,05	1,15	1,25	1,35
Số con giống A	8	28	32	17
Số con giống B	13	14	24	14

a) Số cân nặng trung bình của lợn con giống A là:

$$(1,05 \cdot 8 + 1,15 \cdot 28 + 1,25 \cdot 32 + 1,35 \cdot 17) : 85 = 1,22 \text{ (kg)}$$

Số cân nặng trung bình của lợn con giống B là:

$$(1,05 \cdot 13 + 1,15 \cdot 14 + 1,25 \cdot 24 + 1,35 \cdot 14) : 65 = 1,21 \text{ (kg)}$$

Vậy cân nặng trung bình của lợn con giống A lớn hơn giống B

Gọi $x_1, x_2, x_3, \dots, x_{85}$ lần lượt là số lợn con giống A theo thứ tự không gian

$$x_1, \dots, x_8 \in [1,0;1,1); x_9, \dots, x_{36} \in [1,1;1,2); x_{37}, \dots, x_{68} \in [1,2;1,3); x_{69}, \dots, x_{85} \in [1,3;1,4)$$

Do

$$[1,2;1,3)$$

Trung vị của mẫu số liệu lợn con giống A thuộc nhóm

$$M_{.4} = 1,2 + \frac{\frac{85}{2} - 36}{32} \cdot (1,3 - 1,2) = 1,22$$

Gọi $y_1, y_2, y_3, \dots, y_{65}$ lần lượt là số lợn con giống B theo thứ tự không gian.

$$y_1, \dots, y_{13} \in [1,0;1,1); y_{14}, \dots, y_{27} \in [1,1;1,2); y_{28}, \dots, y_{51} \in [1,2;1,3); y_{52}, \dots, y_{65} \in [1,3;1,4)$$

Do

;

$$[1, 2; 1, 3)$$

Trung vị của mẫu số liệu lợn con giống B thuộc nhóm

$$M_B = 1, 2 + \frac{\frac{65}{2} - 27}{24} \cdot (1, 3 - 1, 2) = 1, 223$$

B.

Vậy cân nặng trung bình của lợn con giống A nhỏ hơn giống

b) Tứ phân vị thứ nhất của dãy số liệu giống A là $\frac{1}{2}(x_{21} + x_{22})$ thuộc nhóm $[1, 1; 1, 2)$ nên tứ

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

phân vị thứ nhất của mẫu số liệu là

Tứ phân vị thứ ba của dãy số liệu giống A là $\frac{1}{2}(x_{63} + x_{64})$ thuộc nhóm $[1, 2; 1, 3)$ nên tứ phân

$$Q_{3,A} = 1, 2 + \frac{\frac{3,85}{4} - 36}{32} (1, 3 - 1, 2) = 1, 29$$

vị thứ ba của mẫu số liệu là

Tứ phân vị thứ nhất của dãy số liệu giống B là $\frac{1}{2}(y_{16} + y_{17})$ thuộc nhóm $[1, 1; 1, 2)$ nên tứ phân

$$Q_{1,B} = 1, 1 + \frac{\frac{65}{4} - 13}{14} (1, 2 - 1, 1) = 1, 12$$

vị thứ nhất của mẫu số liệu là

Tứ phân vị thứ ba của dãy số liệu giống B là $\frac{1}{2}(y_{48} + y_{49})$ thuộc nhóm $[1, 2; 1, 3)$ nên tứ phân

$$Q_{3,B} = 1, 2 + \frac{\frac{3,65}{4} - 27}{24} (1, 3 - 1, 2) = 1, 29$$

vị thứ ba của mẫu số liệu là

Câu 78: Bảng 15 cho ta bảng tần số ghép nhóm số liệu thống kê chiều cao của 40 mẫu cây ở một vườn thực vật (đơn vị: centimét)

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 ghép nhóm trên là bao nhiêu?

Bảng 15

Nhóm	Tần số	Tần số tích lũy
[30; 40)	4	4
[40; 50)	10	14
[50; 60)	14	28
[60; 70)	6	34
[70; 80)	4	38
[80; 90)	2	40
	$n = 40$	

Lời giải

a) Có bảng ghép nhóm bao gồm cả giá trị đại diện là:

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

Trung bình cộng là:

$$\bar{x} = \frac{35 \cdot 4 + 45 \cdot 10 + 55 \cdot 14 + 65 \cdot 6 + 75 \cdot 4 + 85 \cdot 2}{40} = 55,5$$

Trung vị là

Số phần tử của mẫu là $n = 40$. Ta có:

$$\frac{n}{2} = \frac{40}{2} = 20 \Rightarrow$$

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

Xét nhóm 3 là nhóm $[50; 60)$ có $r = 50; d = 10; n_3 = 14$ và nhóm 2 là nhóm $[45; 50)$ có $cf_2 = 14$.

$$M_e = 50 + \left(\frac{20 - 14}{14} \right) \cdot 10 \approx 54,3 \text{ (centimét)}$$

Áp dụng công thức, ta có trung vị của mẫu số liệu là:

Q_1 là:

Số phần tử của mẫu là $n = 40$.

Ta có $\frac{n}{4} = \frac{40}{4} = 10$. Suy ra nhóm 2 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 10.
 Xét nhóm 2 là nhóm $[40; 50)$ có $r = 40; d = 10; n_2 = 10$ và nhóm 1 là nhóm $[30; 40)$ có $cf_1 = 4$

Áp dụng công thức, ta có Q_1 của mẫu số liệu là: $Q_1 = 40 + \left(\frac{10 - 4}{10} \right) \cdot 10 \approx 46$ (centimét)

Q_2 là Có $Q_2 = M_c \approx 54,3$ (centimét).

Q_3 là:

Ta có $\frac{3n}{4} = 30$. Suy ra nhóm 4 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 30. Xét nhóm 4 là nhóm $[60; 70)$ có $r = 60; d = 10; n_4 = 6$ và nhóm 3 là nhóm $[50; 60)$ có $cf_3 = 28$.

Áp dụng công thức, ta có Q_3 của mẫu số liệu là: $Q_3 = 60 + \left(\frac{30 - 28}{6} \right) \cdot 10 = 63,3$ (centimét)

c) Một của mẫu số liệu là:

Có nhóm 3 là nhóm có tần số lớn nhất $M_o = 50 + \left(\frac{14 - 10}{2 \cdot 14 - 10 - 6} \right) \cdot 10 \approx 53,3$.