

BÀI 9. CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM

- CHƯƠNG 3. CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CỦA MẪU SỐ LIỆU GHEP NHÓM
- |FanPage: Nguyễn Bảo Vương

PHẦN A. LÝ THUYẾT VÀ VÍ DỤ MINH HỌA

1. SỐ TRUNG BÌNH CỦA MẪU SỐ LIỆU GHEP NHÓM

Cho mẫu số liệu ghép nhóm

Nhóm	$[a_1; a_2)$...	$[a_i; a_{i+1})$...	$[a_k; a_{k+1})$
Tần số	m_1	...	m_i	...	m_k

Bảng 3.2

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

$$\bar{x} = \frac{m_1 x_1 + \dots + m_k x_k}{n}$$

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

Chú ý. Đối với số liệu rời rạc, người ta thường cho các nhóm dưới dạng $k_1 - k_2$, trong đó $k_1, k_2 \in \mathbb{N}$. Nhóm $k_1 - k_2$ được hiểu là nhóm gồm các giá trị $k_1, k_1 + 1, \dots, k_2$. Khi đó, ta cần hiệu chỉnh mẫu dữ liệu ghép nhóm để đưa về dạng Bảng 3.2 trước khi thực hiện tính toán các số đặc trưng bằng cách hiệu chỉnh nhóm $k_1 - k_2$ với $k_1, k_2 \in \mathbb{N}$ thành nhóm $[k_1 - 0,5; k_2 + 0,5)$. Chẳng hạn, với dữ liệu ghép nhóm điểm thi môn Toán trong Bảng 3.3 sau khi hiệu chỉnh ta được Bảng 3.4.

Điểm thi	1-4	5-7	8-10
Số học sinh	5	20	10

Bảng 3.3

Điểm thi	[0,5; 4,5)	[4,5; 7,5)	[7,5; 10]
Số học sinh	5	20	10

Bảng 3.4

Ví dụ 1. Tìm cân nặng trung bình của học sinh lớp 11D cho trong Bảng 3.5.

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

Bảng 3.5. Cân nặng của học sinh lớp 11D

Giải

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

Cân nặng (kg)	43	48	53	58	63	68
Số học sinh	10	7	16	4	2	3

Tổng số học sinh là $n = 42$. Cân nặng trung bình của học sinh lớp 11D là

$$\bar{x} = \frac{10 \cdot 43 + 7 \cdot 48 + 16 \cdot 53 + 4 \cdot 58 + 2 \cdot 63 + 3 \cdot 68}{42} \approx 51,81(\text{kg}).$$

2. TRUNG VỊ CỦA MẪU SỐ LIỆU GHEP NHÓM

Cho mẫu số liệu ghép nhóm

Nhóm	$[a_1; a_2)$...	$[a_i; a_{i+1})$...	$[a_k; a_{k+1})$
Tần số	m_1	...	m_i	...	m_k

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

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

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

Ví dụ 2. Thời gian (phút) truy cập Internet mỗi buổi tối của một số học sinh được cho trong bảng sau:

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

Tính trung vị của mẫu số liệu ghép nhóm này.

Giải

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

Gọi x_1, \dots, x_{56} là thời gian vào Internet của 56 học sinh và giả sử dãy này đã được sắp xếp theo thứ tự tăng dần. Khi đó, trung vị là $\frac{x_{28} + x_{29}}{2}$. Do 2 giá trị x_{28}, x_{29} thuộc nhóm [15,5;18,5) nên nhóm này chứa trung vị. Do đó, $p = 3; a_3 = 15,5; m_3 = 15; m_1 + m_2 = 3 + 12 = 15; a_4 - a_3 = 3$ và ta có

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

3. TỨ PHÂN VỊ CỦA MẪU SỐ LIỆU GHEP NHÓM

Cho mẫu số liệu ghép nhóm

Nhóm	$[a_1; a_2)$...	$[a_i; a_{i+1})$...	$[a_k; a_{k+1})$
Tần số	m_1	...	m_i	...	m_k

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

$$Q_1 = a_p + \frac{\frac{n}{4} - (m_1 + \dots + m_{p-1})}{m_p} \cdot (a_{p+1} - a_p),$$

trong đó, n là cỡ mẫu, m_p là tần số nhóm p , với $p=1$ ta quy ước $m_1 + \dots + m_{p-1} = 0$.

Để tính tứ phân vị thứ ba Q_3 của mẫu số liệu ghép nhóm, trước hết ta xác định nhóm chứa Q_3 .

Giả sử đó là nhóm thứ $p: [a_p; a_{p+1})$. Khi đó,

$$Q_3 = a_p + \frac{\frac{3n}{4} - (m_1 + \dots + m_{p-1})}{m_p} \cdot (a_{p+1} - a_p),$$

trong đó, n là cỡ mẫu, m_p là tần số nhóm p , với $p=1$ ta quy ước $m_1 + \dots + m_{p-1} = 0$.

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

Ví dụ 3. Tìm tứ phân vị thứ nhất Q_1 và tứ phân vị thứ ba Q_3 của mẫu số liệu ghép nhóm cho trong Ví dụ 2.

Giải

Cỡ mẫu là $n = 56$.

Tứ phân vị thứ nhất Q_1 là $\frac{x_{14} + x_{15}}{2}$. Do x_{14}, x_{15} đều thuộc nhóm [12,5;15,5) nên nhóm này chứa Q_1 . Do đó, $p = 2; a_2 = 12,5; m_2 = 12; m_1 = 3; a_3 - a_2 = 3$ và ta có

$$Q_1 = 12,5 + \frac{\frac{56}{3} - 3}{12} \cdot 3 = 15,25$$

Với tứ phân vị thứ ba Q_3 là $\frac{x_{42} + x_{43}}{2}$. Do x_{42}, x_{43} đều thuộc nhóm $[18,5; 21,5)$ nên nhóm này chứa Q_3 . Do đó, $p = 4; a_4 = 18,5; m_4 = 24; m_1 + m_2 + m_3 = 3 + 12 + 15 = 30; a_5 - a_4 = 3$ và ta có

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

Nhận xét. Ta cũng có thể xác định nhóm chứa tứ phân vị thứ r nhờ tính chất: có khoảng $\left(\frac{r \cdot n}{4}\right)$ giá trị nhỏ hơn tứ phân vị này.

4. MỘT CỦA MẪU SỐ LIỆU GHEP NHÓM

Cho mẫu số liệu ghép nhóm

Nhóm	$[a_1; a_2)$...	$[a_i; a_{i+1})$...	$[a_k; a_{k+1})$
Tần số	m_1	...	m_i	...	m_k

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

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

Bước 2. Một được xác định là: $M_o = a_j + \frac{m_j - m_{j-1}}{(m_j - m_{j-1}) + (m_j - m_{j+1})} \cdot h$

trong đó m_j là tần số của nhóm j (quy ước $m_0 = m_{k+1} = 0$) và h là độ dài của nhóm.

Lưu ý. Người ta chỉ định nghĩa một cho mẫu ghép nhóm có độ dài các nhóm bằng nhau. Một mẫu có thể không có một hoặc có nhiều hơn một một.

Khi tần số của các nhóm số liệu bằng nhau thì mẫu số liệu ghép nhóm không có một.

Ví dụ 4. Bảng số liệu ghép nhóm sau cho biết chiều cao (cm) của 50 học sinh lớp 11A.

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

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

Giải

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

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

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

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

PHẦN B. BÀI TẬP TỰ LUẬN (PHÂN DẠNG)

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

Câu 1. (SGK - KNTT 11 – Tập 1) Khảo sát thời gian tự học của các học sinh trong lớp theo mẫu bên.

- Hãy lập bảng thống kê cho mẫu số liệu ghép nhóm thu được.
- Có thể tính chính xác thời gian tự học trung bình của các học sinh trong lớp không?
- Có cách nào tính gần đúng thời gian tự học trung bình của các học sinh trong lớp dựa trên mẫu số liệu ghép nhóm này không?

Lời giải

a) Giả sử lớp 11A có 30 học sinh và sau khi khảo sát, ta có được bảng thống kê như sau:

Thời gian (giờ)	Dưới 1,5 giờ	$[1,5; 3)$	$[3; 4,5)$	Từ 4,5 giờ trở lên
-----------------	--------------	------------	------------	--------------------

Số học sinh	5	15	8	2
-------------	---	----	---	---

b) Ta không thể tính chính xác thời gian tự học trung bình của các học sinh trong lớp vì không có mẫu số liệu cụ thể về thời gian tự học của từng học sinh.

c) Có thể tính gần đúng thời gian tự học trung bình của các học sinh trong lớp dựa trên mẫu số liệu ghép nhóm bằng cách chọn thời gian đại diện cho mỗi nhóm, sau đó sử dụng tần số tương ứng để tính số trung bình, cụ thể:

- Thời gian tự học dưới 1,5 giờ, ta chọn giá trị đại diện là 0,75 giờ, tần số tương ứng là 5.

- Thời gian tự học từ 1,5 đến dưới 3 giờ, ta chọn giá trị đại diện là $\frac{1,5+3}{2} = 2,25$, tần số tương ứng là 15.

- Thời gian tự học từ 3 đến dưới 4,5 giờ, ta chọn giá trị đại diện là $\frac{3+4,5}{2} = 3,75$, tần số tương ứng là 8.

- Thời gian tự học là từ 4,5 giờ trở lên, ta chọn giá trị đại diện là 5,25, tần số tương ứng là 2.

Số trung bình là $\bar{x} = \frac{5.0,75 + 15.2,25 + 8.3,75 + 2.5,25}{30} = 2,6$.

Vậy thời gian tự học trung bình của học sinh lớp 11A xấp xỉ khoảng 2,6 giờ.

Câu 2. (SGK -KNTT 11 – Tập 1) Tìm hiểu thời gian xem ti vi trong tuần trước (đơn vị: giờ) của một số học sinh thu được kết quả sau:

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

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

Lời giải

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

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

Tổng số học sinh là $n = 8 + 16 + 4 + 2 + 2 = 32$. Thời gian xem ti vi trung bình trong tuần trước của

các học sinh là $\bar{x} = \frac{8.2,5 + 16.7,5 + 4.12,5 + 2.17,5 + 2.22,5}{32} = 8,4375$ (giờ)

Câu 3. (SGK -KNTT 11 – Tập 1) Tìm hiểu thời gian xem ti vi trong tuần trước (đơn vị: giờ) của một số học sinh thu được kết quả sau:

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

a) Có thể tìm được giá trị chính xác cho một của mẫu số liệu gốc về thời gian xem ti vi của học sinh không?

b) Một thuộc nhóm nào là hợp lí nhất? Nên lấy số nào trong nhóm để ước lượng cho một?

Lời giải

a) Không thể tính được giá trị chính xác cho một của mẫu số liệu gốc về thời gian xem ti vi của học sinh, do không có thời gian cụ thể của từng học sinh.

b) Tần số lớn nhất là 16 nên một thuộc nhóm $[5;10)$ là hợp lí nhất. Ta ước lượng một của mẫu số liệu bằng cách xác định số thứ tự của nhóm chứa một là $j = 2; a_j = a_2 = 5; m_2 = 16; m_1 = 8; m_3 = 4$; độ dài của nhóm $h = 5$.

Do đó, một của mẫu số liệu xấp xỉ bằng $5 + \frac{16-8}{(16-8)+(16-4)} .5 = 7$.

Câu 4. (SGK-KNTT 11 – Tập 1) Thời gian (phút) để học sinh hoàn thành một câu hỏi thi được cho như sau:

Thời gian (phút)	[0,5;10,5)	[10,5;20,5)	[20,5;30,5)	[30,5;40,5)	[40,5;50,5)
Số học sinh	2	10	6	4	3

Tìm một của mẫu số liệu ghép nhóm này.

Lời giải

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

Ta có, $j = 2; a_2 = 10,5; m_2 = 10; m_1 = 2; m_3 = 6; h = 20,5 - 10,5 = 10$. Do đó, một của mẫu số liệu

$$\text{ghép nhóm này là } M_o = 10,5 + \frac{10-2}{(10-2)+(10-6)} \cdot 10 = \frac{103}{6} \approx 17,17$$

Câu 5. (SGK-KNTT 11 – Tập 1) Quảng 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) Giá trị nhỏ nhất của mẫu số liệu là 2, giá trị lớn nhất của mẫu số liệu là 32, do đó khoảng biến thiên là $32 - 2 = 30$.

Các nhóm có độ rộng bằng nhau và độ rộng của mỗi nhóm là 5. Để cho thuận tiện, ta chia thành 7 nhóm là các nhóm $[0;5), [5;10), [10;15), [15;20), [20;25), [25;30), [30;35)$. Đếm số giá trị thuộc mỗi nhóm, ta có mẫu số liệu ghép nhóm như sau:

Quảng đường (km)	[0;5)	[5;10)	[10;15)	[15;20)	[20;25)	[25;30)	[30;35)
Số công nhân	5	11	11	9	1	1	2

Giá trị đại diện cho mỗi nhóm là trung bình của hai đầu mút của nhóm. Ta có bảng giá trị đại diện như sau:

Quảng đường (km) (giá trị đại diện)	2,5	7,5	12,5	17,5	22,5	27,5	32,5
Số công nhân	5	11	11	9	1	1	2

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

$$\bar{x}_g = \frac{5 \cdot 2,5 + 11 \cdot 7,5 + 11 \cdot 12,5 + 9 \cdot 17,5 + 1 \cdot 22,5 + 1 \cdot 27,5 + 2 \cdot 32,5}{40} = 12,625$$

Ta có:

$$5 + 3 + 10 + 20 + 25 + 11 + 13 + 7 + 12 + 31 + 19 + 10 + 12 + 17 + 18 + 11 + 32 + 17 + 16 + 2 + 7 + 9 + 7 + 8 + 3 + 5 + 12 + 15 + 18 + 3 + 12 + 14 + 2 + 9 + 6 + 15 + 15 + 7 + 6 + 12 = 476$$

$$\text{Số trung bình của mẫu số liệu không ghép nhóm là } \bar{x} = \frac{476}{40} = 11,9.$$

Giá trị trung bình của mẫu số liệu không ghép nhóm chính xác hơn vì nó là giá trị của mẫu số liệu gốc.

c) Tần số lớn nhất trong bảng tần số của mẫu số liệu ghép nhóm là 11. Do đó, nhóm chứa một của mẫu số liệu ghép nhóm là các nhóm $[5;10)$ và $[10;15)$.

Câu 6. (SGK -KNTT 11 – Tập 1) 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) Tần số lớn nhất là 14 nên nhóm chứa một là nhóm $[3;3,5)$. Ta có, $j=3, a_3=3, m_3=14, m_2=9, m_4=11, h=0,5$. Do đó một của mẫu số liệu ghép nhóm là

$$M_o = 3 + \frac{14-9}{(14-9)+(14-11)} \cdot 0,5 = 3,3125.$$

Ý nghĩa: Tuổi thọ của bình ắc quy ô tô khoảng 3,3125 năm là nhiều nhất hay tuổi thọ chủ yếu của bình ắc quy ô tô khoảng 3,3125 năm.

b) Trong mỗi khoảng tuổi thọ, giá trị đại diện là trung bình cộng của giá trị hai đầu mút nên ta có bảng sau:

Tuổi thọ (năm)	2,25	2,75	3,25	3,75	4,25	4,75
Tần số	4	9	14	11	7	5

Tổng số ắc quy ô tô là 50. Tuổi thọ trung bình của 50 ắc quy ô tô này là

$$\bar{x} = \frac{4 \cdot 2,25 + 9 \cdot 2,75 + 14 \cdot 3,25 + 11 \cdot 3,75 + 7 \cdot 4,25 + 5 \cdot 4,75}{50} = 3,48$$

Câu 7. Mức thưởng tết (triệu đồng) mà các công nhân một nhà máy nhận được như sau:

Mức thưởng	$[5;10)$	$[10;15)$	$[15;20)$	$[20;25]$
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, $a_j = 15; m_j = 47; m_{j-1} = 35; m_{j+1} = 25; h = 5$. Do đó, một của mẫu số liệu là

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

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

Câu 8. Quãng đường (km) các cầu thủ (không tính thủ môn) chạy trong một trận bóng đá tại giải ngoại hạng Anh được cho trong bảng thống kê sau:

Quãng đường	$[2;4)$	$[4;6)$	$[6;8)$	$[8;10)$	$[10;12)$
Số cầu thủ	2	5	6	9	3

Tính quãng đường trung bình một cầu thủ chạy trong trận đấu này.

Lời giải

ĐS. $7,48 \text{ km}$.

Câu 9. Thống kê số lần đi học muộn trong học kì của các bạn trong lớp, Nam thu được kết quả sau:

Số lần đi muộn	0–2	3–5	6–8	9–11	12–14
Số học sinh	23	8	5	3	1

Tính một của mẫu số liệu và giải thích ý nghĩa của giá trị thu được.

Lời giải

Nhóm chứa một là $[8;10)$. Một là $M_o = 8 + \frac{(9-6)}{(9-6)+(9-3)} \cdot 2 \approx 8,67$.

Số cầu thủ chạy khoảng 8,67 km là nhiều nhất.

Câu 10. Thống kê số lần đi học muộn trong học kì của các bạn trong lớp, Nam thu được kết quả sau:

Số lần đi muộn	0-2	3-5	6-8	9-11	12-14
Số học sinh	23	8	5	3	1

Trung bình mỗi học sinh trong lớp đi muộn bao nhiêu buổi trong học kì?

Lời giải

ĐS. Trung bình trong học kì mỗi học sinh đi muộn 3,325 buổi.

Câu 11. Nồng độ cồn trong hơi thở (đơn vị tính là miligam/1 lít khí thở) của 20 lái xe ô tô vi phạm được cho như sau:

0,09 0,18 0,47 1,20 0,28 0,45 0,72 0,15 0,75 0,36

0,21 0,15 0,23 0,30 0,41 0,13 0,05 0,38 0,42 0,79 .

Theo quy định, mức phạt nồng độ cồn đối với lái xe ô tô như sau:

Mức 1. Nồng độ cồn trong hơi thở chưa vượt quá 0,25 phạt từ 6 đến 8 triệu đồng;

Mức 2. Nồng độ cồn trong hơi thở từ trên 0,25 đến 0,4 phạt từ 16 đến 18 triệu đồng;

Mức 3. Nồng độ cồn trong hơi thở vượt quá 0,4 phạt từ 30 đến 40 triệu đồng.

a) Lập bảng thống kê biểu diễn số lượng lái xe vi phạm theo mức tiền bị phạt.

b) Trung bình mỗi lái xe bị phạt bao nhiêu tiền? Tổng số tiền phạt của 20 lái xe khoảng bao nhiêu?

Lời giải

a) Bảng thống kê:

Số tiền phạt (triệu đồng)	6-8	16-18	30-40
Số người vi phạm	8	4	8

b) Cỡ mẫu $n = 20$. Số tiền trung bình một người bị phạt là $\bar{x} = \frac{8 \cdot 7 + 4 \cdot 17 + 8 \cdot 35}{20} = 20,2$ (triệu đồng).

Tổng số tiền 20 lái xe bị phạt là $20,2 \cdot 20 = 404$ (triệu đồng).

Câu 12. Trong các mẫu số liệu cho trong bài tập 3.23 và 3.24, ta có thể tìm một cho mẫu số liệu nào? Tìm một của mẫu số liệu đó và giải thích ý nghĩa của giá trị tìm được.

Lời giải

Các nhóm số liệu trong bài tập 3.23 không có độ dài bằng nhau nên người ta không định nghĩa một. Hiệu chỉnh mẫu số liệu trong bài 3.24 như trên ta thấy nhóm chứa một là nhóm $[3,5;6,5)$, do

đó một là $M_o = 3,5 + \frac{(18-5)}{(18-5)+(18-13)} \cdot 3 \approx 5,67$

Số học sinh đăng kí khoảng 5,67 nguyện vọng là nhiều nhất.

Câu 13. 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.

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

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

Nhóm chứa một của mẫu số liệu trên là nhóm [29,5; 34,5).

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

Mốt của mẫu số liệu ghép nhóm là $M_o = 29,5 + \frac{10 - 6}{(10 - 6) + (10 - 6)} \cdot 5 = 32$.

Câu 14. Kết quả khảo sát cân nặng của 20 quả cam Canh ở mỗi lô hàng 1 và lô hàng 2 được cho ở bảng sau:

Cân nặng (g)	[100;110)	[110;120)	[120;130)	[130;140)	[140;150)
Số quả cam Canh ở lô hàng 1	1	4	5	4	6
Số quả cam Canh ở lô hàng 2	2	3	6	4	5

a) Hãy ước lượng cân nặng trung bình của mỗi quả cam Canh ở lô hàng 1 và lô hàng 2.

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

Giải

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

Cân nặng đại diện (g)	105	115	125	135	145
Số quả cam Canh ở lô hàng 1	1	4	5	4	6
Số quả cam Canh ở lô hàng 2	2	3	6	4	5

a) Cân nặng trung bình của mỗi quả cam Canh ở lô hàng 1 là

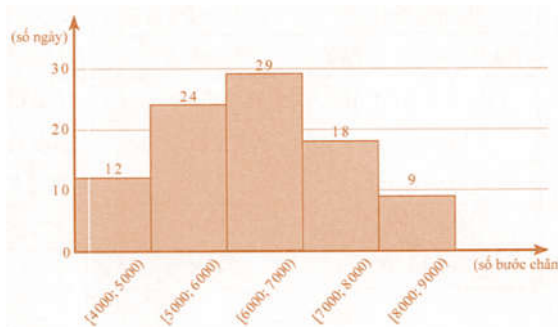
$$(1.105 + 4.115 + 5.125 + 4.135 + 6.145) : 20 = 130(g).$$

Cân nặng trung bình của mỗi quả cam Canh ở lô hàng 2 là

$$(2.105 + 3.115 + 6.125 + 4.135 + 5.145) : 20 = 128,5(g).$$

b) Nếu so sánh theo số trung bình thì số cam Canh ở lô hàng 1 nặng hơn số cam Canh ở lô hàng 2.

Câu 15. 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.

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).

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

Một của mẫu số liệu ghép nhóm là

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

Câu 16. 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.

Giải

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

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

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

$$\bar{x} = \frac{7.1750 + 6.3250 + 3.4750 + 3.6250}{19} \approx 3407,89.$$

Nhóm chứa một của mẫu số liệu trên là nhóm $[1000; 2500)$.

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

Mốt của mẫu số liệu ghép nhóm là

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

Câu 17. 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 một 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.

Mốt của mẫu số liệu ghép nhóm là $\frac{145}{26} \approx 5,58$.

Câu 18. 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.

b) Lập bảng tần số ghép nhóm của mẫu số liệu trên với nhóm đầu tiên là $[5,40;5,45)$ và ước lượng số trung bình của mẫu số liệu ghép nhóm trên.

Lời giải

a) Cân nặng trung bình của mỗi chi tiết máy là 5,54 g.

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 19. 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.

b) Hãy lập bảng tần số ghép nhóm của mẫu số liệu trên với nhóm đầu tiên là $[4,5;7,5)$.

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.

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

Câu 20. 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

a) Hãy lập bảng tần số ghép nhóm của mẫu số liệu trên với nhóm đầu tiên là $[4,6;4,7)$.

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 ngen	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 21. Bảng sau thống kê chiều cao (đơn vị: cm) của một số cây giống sau khi nảy mầm được 2 tuần.

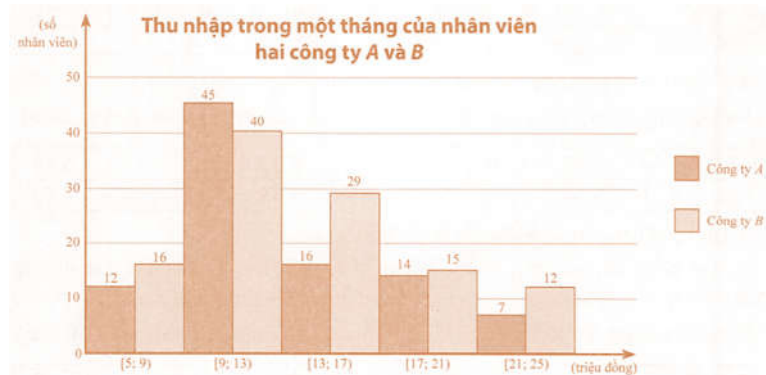
Chiều cao (cm)	$[6,2;6,7)$	$[6,7;7,2)$	$[7,2;7,7)$	$[7,7;8,2)$	$[8,2;8,7)$
Số cây	10	21	28	12	9

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 22. 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

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

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

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

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

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

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

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

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

Lời giải

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

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

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

Câu 24. Kết quả khảo sát cân nặng của 25 quả cam ở mỗi lô hàng A và B được cho ở bảng sau:

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

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

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

Lời giải

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

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

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

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

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

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

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

Câu 25. Cân nặng của 28 học sinh nam lớp 11 được cho như sau:

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

49,7 45,1 56,2 63,2 46,1 49,6 59,1 55,3 55,8 45,5 46,8 54 49,2 52,6 a) Hãy chia mẫu dữ liệu trên thành 5 nhóm, lập bảng tần số ghép nhóm và xác định giá trị đại diện cho mỗi nhóm.

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

Lời giải

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

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

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

Khi đó ta có bảng tần số ghép nhóm sau:

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

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

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

Câu 26. 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 $[18; 22)$.

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

Một 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 27. Hãy sử dụng dữ liệu ở để 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:

$$M_0 = 40 + \frac{10-6}{(10-6)+(10-7)} \cdot 10 = 45,7$$

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:

$$M_0 = 30 + \frac{9-3}{(9-3)+(9-6)} \cdot 10 = 36,7$$

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 28. 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

$$M_0 = 5,5 + \frac{13-5}{(13-5)+(13-7)} \cdot 3 = \frac{101}{14} \approx 7,2$$

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

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

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

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

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

Số lần	?	?	?	?	?
--------	---	---	---	---	---

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

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

Lời giải

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

b) Bảng tần số ghép nhóm

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

c)

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

Cự li trung bình mỗi lần ném xấp xỉ bằng $\frac{69,6 \cdot 4 + 70,4 \cdot 2 + 71,2 \cdot 9 + 72 \cdot 10 + 72,8 \cdot 5}{30} = 71,5 \text{ (m)}$

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

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

Mốt của mẫu số liệu ghép nhóm là:

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

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

Câu 30. Cho bảng tần số ghép nhóm số liệu thống kê chiều cao của 40 mẫu cây ở một vườn thực vật (đơn vị: centimet).

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

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

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

Lời giải

a)

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

Giá trị trung bình của mẫu số liệu là

$$\bar{x} = \frac{35 \cdot 4 + 45 \cdot 10 + 55 \cdot 14 + 65 \cdot 6 + 75 \cdot 4 + 85 \cdot 2}{40} \approx 63,43$$

b) Nhóm chứa một: [50;60)

Mốt của mẫu số liệu:

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

Câu 31. Mẫu số liệu sau ghi lại cân nặng của 30 bạn học sinh (đơn vị: kilogam)

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

- a) Lập bảng tần số ghép nhóm cho mẫu số liệu trên có tám nhóm ứng với tám nửa khoảng $[15; 20), [20; 25), [25; 30), [30; 35), [35; 40), [40; 45), [45; 50), [50; 55)$.
- b) Xác định số trung bình cộng của mẫu số liệu ghép nhóm trên.
- c) Một của mẫu số liệu ghép nhóm trên là bao nhiêu

Lời giải

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

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

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

- c) Một của mẫu số liệu

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

Câu 32. An tìm hiểu hàm lượng chất béo (đơn vị: g) có trong 100 g mỗi loại thực phẩm. Sau khi thu thập dữ liệu về 60 loại thực phẩm, An lập được bảng thống kê

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

- a) Xác định giá trị trung bình của mẫu số liệu.
- b) Tìm một của mẫu số liệu.

Lời giải

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

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

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

Câu 33. 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) Bảng tần số ghép nhóm

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

- c) Hiệu chỉnh lại bảng số liệu ta có:

Số xe	[5,5;10,5)	[10,5;15,5)	[15,5;20,5)	[20,5;25,5)	[25,5;30,5)
Giá trị đại diện	8	13	18	23	28
Số lần	5	9	3	9	4

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

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

Câu 34. 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:

$$(18.3 + 23.6 + 28.15 + 33.27 + 38.22 + 43.14 + 48.5) : 92 = 34,6$$

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

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

Một của mẫu số liệu ghép nhóm là:

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

Câu 35. 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 một 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:

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

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

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

Mốt của mẫu số liệu trên là:

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

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

Câu 36. (SGK -KNTT 11 – Tập 1) Cho mẫu số liệu ghép nhóm về chiều cao của 21 cây na giống.

Chiều cao (cm)	[0;5)	[5;10)	[10;15)	[15;20)
Số cây	3	8	7	3

Gọi x_1, x_2, \dots, x_{21} là chiều cao của các cây giống, đã được sắp xếp theo thứ tự tăng dần. Khi đó, x_1, \dots, x_3 thuộc $[0;5)$, x_4, \dots, x_{11} thuộc $[5;10)$,... Hỏi trung vị thuộc nhóm nào?

Lời giải

Ta có: cỡ mẫu $n = 21$, là số lẻ nên trung vị là giá trị chính giữa của mẫu số liệu và là giá trị ở vị trí thứ 11 của mẫu số liệu. Mà x_{11} thuộc $[5;10)$ nên trung vị của mẫu số liệu thuộc nhóm $[5;10)$.

Câu 37. (SGK -KNTT 11 – Tập 1) Ghi lại tốc độ bóng trong 200 lần giao bóng của một vận động viên môn quần vợt cho kết quả như bảng bên.

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

Tính trung vị của mẫu số liệu ghép nhóm này.

Lời giải

Cỡ mẫu là $n = 200$.

Gọi x_1, x_2, \dots, x_{200} là tốc độ giao bóng của vận động viên trong 200 lần giao bóng và giả sử dãy này đã được sắp xếp theo thứ tự tăng dần. Khi đó, trung vị là $\frac{x_{100} + x_{101}}{2}$. Do 2 giá trị x_{100}, x_{101} thuộc nhóm $[165;170)$ (vì $18+28+35+43=124$) nên nhóm này chứa trung vị. Do đó, $p = 4; a_4 = 165; m_4 = 43; m_1 + m_2 + m_3 = 18+28+35=81; a_5 - a_4 = 170-165=5$ và ta có

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

Câu 38. (SGK -KNTT 11 – Tập 1) Cho mẫu số liệu ghép nhóm về chiều cao của 21 cây na giống.

Chiều cao (cm)	[0;5)	[5;10)	[10;15)	[15;20)
Số cây	3	8	7	3

hãy cho biết tứ phân vị thứ nhất Q_1 và tứ phân vị thứ ba Q_3 thuộc nhóm nào.

Lời giải

Vì $n = 21$ nên tứ phân vị thứ nhất là trung vị của dãy gồm 10 số liệu đầu tiên và chính là trung bình cộng của giá trị ở vị trí thứ 5 và thứ 6, do đó $Q_1 = \frac{x_5 + x_6}{2}$,

Tứ phân vị thứ ba là trung vị của dãy gồm 10 số liệu nằm bên phải trung vị là dãy $x_{12}, x_{13}, \dots, x_{21}$ nên $Q_3 = \frac{x_{16} + x_{17}}{2}$. Ta có: $3 + 8 + 7 = 18$, do đó x_{16}, x_{17} thuộc nhóm $[10;15)$ nên tứ phân vị thứ ba Q_3 thuộc nhóm $[10;15)$.

Câu 39. (SGK -KNTT 11 – Tập 1) Ghi lại tốc độ bóng trong 200 lần giao bóng của một vận động viên môn quần vợt cho kết quả như bảng bên.

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

Tìm tứ phân vị thứ nhất và tứ phân vị thứ ba cho mẫu số liệu ghép nhóm

Lời giải

Cỡ mẫu là $n = 200$.

Tứ phân vị thứ nhất Q_1 là $\frac{x_{50} + x_{51}}{2}$. Do x_{50}, x_{51} đều thuộc nhóm $[160;165)$ nên nhóm này chứa Q_1 . Do đó, $p = 3$; $a_3 = 160$; $m_3 = 35$; $m_1 + m_2 = 18 + 28 = 46$; $a_4 - a_3 = 165 - 160 = 5$ và ta có

$$Q_1 = 160 + \frac{\frac{200}{4} - 46}{35} \cdot 5 \approx 160,57$$

Tứ phân vị thứ ba Q_3 là $\frac{x_{150} + x_{151}}{2}$. Do x_{150}, x_{151} đều thuộc nhóm $[170;175)$ nên nhóm này chứa Q_3 . Do đó, $p = 5$; $a_5 = 170$; $m_5 = 41$; $m_1 + m_2 + m_3 + m_4 = 18 + 28 + 35 + 43 = 124$; $a_6 - a_5 = 175 - 170 = 5$ và ta có

$$Q_3 = 170 + \frac{\frac{3 \cdot 200}{4} - 124}{41} \cdot 5 \approx 173,17$$

Câu 40. (SGK -KNTT 11 – Tập 1) Một cửa hàng đã ghi lại số tiền bán xăng cho 35 khách hàng đi xe máy. Mẫu số liệu gốc có dạng: x_1, x_2, \dots, x_{35} trong đó x_i là số tiền bán xăng cho khách hàng thứ i . Vì một lý do nào đó, cửa hàng chỉ có mẫu số liệu ghép nhóm dạng sau:

Số tiền (nghìn đồng)	$[0;30)$	$[30;60)$	$[60;90)$	$[90;120)$
Số khách hàng	3	15	10	7

Bảng 3.1. Số tiền khách hàng mua xăng

Hãy tính các số đặc trưng cho mẫu số liệu trong Bảng 3.1 và giải thích ý nghĩa của các giá trị thu được.

Lời giải

Ta có:

Số tiền (nghìn đồng)	$[0;30)$	$[30;60)$	$[60;90)$	$[90;120)$
Số khách hàng	3	15	10	7

Bảng 3.1. Số tiền khách hàng mua xăng

+) Số trung bình

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

Số tiền (nghìn đồng)	15	45	75	105
Số khách hàng	3	15	10	7

Tổng số khách hàng là $n = 35$. Số trung bình là $\bar{x} = \frac{3.15 + 15.45 + 10.75 + 7.105}{35} = 63$

Số trung bình của mẫu số liệu ghép nhóm xấp xỉ cho số trung bình của mẫu số liệu gốc. Từ đó, ta thấy số tiền bán xăng trung bình của 35 khách hàng xấp xỉ 63 nghìn đồng và có thể dùng làm đại diện cho mẫu số liệu.

+) Số trung vị, tứ phân vị

Cỡ mẫu là $n = 35$.

- Gọi x_1, x_2, \dots, x_{35} là số tiền xăng của 35 khách hàng và giả sử dãy này đã được sắp xếp theo thứ tự tăng dần. Khi đó, trung vị là x_{18} . Do x_{18} thuộc nhóm $[30; 60)$ nên nhóm này chứa trung vị. Do đó, $p = 2; a_2 = 30; m_2 = 15; m_1 = 3; a_3 - a_2 = 60 - 30 = 30$ và ta có

$$M_e = 30 + \frac{\frac{35}{2} - 3}{15} \cdot 30 = 59$$

Trung vị của mẫu số liệu ghép nhóm xấp xỉ cho trung vị của mẫu số liệu gốc, nó chia mẫu số liệu gốc thành hai phần, mỗi phần chứa 50 % giá trị. Từ đó ta thấy trung vị của mẫu số liệu gốc xấp xỉ bằng 59, giá trị này là ngưỡng để chia mẫu số liệu gốc thành 2 phần.

- Tứ phân vị thứ nhất Q_1 là x_9 . Do x_9 thuộc nhóm $[30; 60)$ nên nhóm này chứa Q_1 . Do đó,

$$p = 2; a_2 = 30; m_2 = 15; m_1 = 3; a_3 - a_2 = 60 - 30 = 30 \text{ và ta có } Q_1 = 30 + \frac{\frac{35}{4} - 3}{15} \cdot 30 = 41,5$$

- Tứ phân vị thứ ba Q_3 là x_{27} . Do x_{27} thuộc nhóm $[60; 90)$ nên nhóm này chứa Q_3 . Do đó, $p = 3; a_3 = 60; m_3 = 10; m_1 + m_2 = 3 + 15 = 18; a_4 - a_3 = 90 - 60 = 30$ và ta có

$$Q_3 = 60 + \frac{\frac{3.35}{4} - 18}{10} \cdot 30 = 84,75$$

- Tứ phân vị thứ hai $Q_2 = M_e = 59$.

Do đó, các tứ phân vị của mẫu số liệu gốc xấp xỉ là $Q_1 = 41,5; Q_2 = 59$ và $Q_3 = 84,75$, mỗi giá trị này là các ngưỡng để chia mẫu số liệu thành 4 phần, mỗi phần chứa 25% giá trị.

+) Mốt

Tần số lớn nhất là 15 nên nhóm chứa mốt là nhóm $[30; 60)$. Ta có,

$$j = 2; a_2 = 30; m_2 = 15; m_1 = 3; m_3 = 10, h = 30. \text{ Do đó } M_o = 30 + \frac{15 - 3}{(15 - 3) + (15 - 10)} \cdot 30 \approx 51,18.$$

Do đó, mốt của mẫu số liệu gốc xấp xỉ bằng 51,18. Vậy số khách hàng mua xăng với giá tiền khoảng 51,18 nghìn đồng là nhiều nhất.

Câu 41. (SGK -KNTT 11 – Tập 1) 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) Hiệu chỉnh để thu được mẫu số liệu ghép nhóm dạng Bảng 3.2 ta được mẫu số liệu ghép nhóm như sau:

Điểm	$[0; 9,5)$	$[9,5; 19,5)$	$[19,5; 29,5)$	$[29,5; 39,5)$	$[39,5; 49,5)$
------	------------	---------------	----------------	----------------	----------------

Số thí sinh	1	2	4	6	15
Điểm	[49,5;59,5)	[59,5;69,5)	[69,5;79,5)	[79,5;89,5)	[89,5;99,5)
Số thí sinh	12	10	6	3	1

b) Cỡ mẫu là $n = 60$.

Gọi x_1, x_2, \dots, x_{60} là điểm thi môn Toán của 60 thí sinh và giả sử dãy này đã được sắp xếp theo thứ tự tăng dần. Khi đó, trung vị là $\frac{x_{30} + x_{31}}{2}$. Do hai giá trị x_{30}, x_{31} thuộc nhóm [49,5;59,5) nên nhóm này chứa trung vị. Do đó, $p = 6; a_6 = 49,5; m_6 = 12; m_1 + m_2 + m_3 + m_4 + m_5 = 1 + 2 + 4 + 6 + 15 = 28; a_7 - a_6 = 59,5 - 49,5 = 10$

$$\text{và ta có } M_e = 49,5 + \frac{\frac{60}{2} - 28}{12} \cdot 10 \approx 51,17$$

Tứ phân vị thứ nhất Q_1 là $\frac{x_{15} + x_{16}}{2}$. Do x_{15} và x_{16} đều thuộc nhóm [39,5;49,5) nên nhóm này chứa Q_1 . Do đó, $p = 5; a_5 = 39,5; m_5 = 15; m_1 + m_2 + m_3 + m_4 = 13; a_6 - a_5 = 10$ và ta có

$$Q_1 = 39,5 + \frac{\frac{60}{4} - 13}{15} \cdot 10 \approx 40,83$$

Tứ phân vị thứ ba Q_3 là $\frac{x_{45} + x_{46}}{2}$. Do x_{45} và x_{46} đều thuộc nhóm [59,5;69,5) nên nhóm này chứa Q_3 . Do đó, $p = 7; a_7 = 59,5; m_7 = 10; m_1 + m_2 + m_3 + m_4 + m_5 + m_6 = 40; a_6 - a_5 = 10$ và ta có

$$Q_3 = 59,5 + \frac{\frac{3 \cdot 60}{4} - 40}{10} \cdot 10 = 64,5$$

Tứ phân vị thứ hai $Q_2 = M_e \approx 51,17$.

Vậy các tứ phân vị của mẫu số liệu là $Q_1 \approx 40,83; Q_2 \approx 51,17$ và $Q_3 = 64,5$. Các giá trị này các là ngưỡng để phân điểm của 60 học sinh thành 4 phần để xếp loại học sinh.

Câu 42. (SGK -KNTT 11 – Tập 1) Phòng vấn một số học sinh khối 11 về thời gian (giờ) ngủ của một buổi tối, thu được bảng số liệu ở bên.

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

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?

Lời giải

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

Thời gian	Số học sinh nam	Số học sinh nữ
4,5	6	4
5,5	10	8

6,5	13	10
7,5	9	8
8,5	7	8

Tổng số các bạn nam là $n_1 = 6 + 10 + 13 + 9 + 7 = 45$.

Thời gian ngủ trung bình của các bạn học sinh nam là

$$\bar{x}_1 = \frac{6.4,5 + 10.5,5 + 13.6,5 + 9.7,5 + 7.8,5}{45} \approx 6,52$$

Tổng số các bạn nữ là $n_2 = 4 + 8 + 10 + 11 + 8 = 41$.

Thời gian ngủ trung bình của các bạn học sinh nữ là

$$\bar{x}_2 = \frac{4.4,5 + 8.5,5 + 10.6,5 + 11.7,5 + 8.8,5}{41} \approx 6,77.$$

Vì $6,52 < 6,77$ nên thời gian ngủ trung bình của các học sinh nam ít hơn các học sinh nữ.

b) Ta có:

Thời gian	Số học sinh nam	Số học sinh nữ	Số học sinh khối 11
[4;5)	6	4	10
[5;6)	10	8	18
[6;7)	13	10	23
[7;8)	9	11	20
[8;9)	7	8	15

Tổng số học sinh khối 11 được khảo sát là $n = 45 + 41 = 86$.

Gọi $x_1, x_2, x_3, \dots, x_{86}$ là thời gian ngủ của các học sinh khối 11 được khảo sát và giả sử dãy này đã sắp xếp theo thứ tự tăng dần. Khi đó trung vị của mẫu số liệu là $\frac{x_{43} + x_{44}}{2}$.

Do đó, tứ phân vị thứ ba $Q_3 = x_{65}$. Vì x_{65} thuộc nhóm [7;8) nên nhóm này chứa Q_3 . Do đó, $p = 4; a_4 = 7; m_4 = 20; m_1 + m_2 + m_3 = 10 + 18 + 23 = 51; a_5 - a_4 = 8 - 7 = 1$ và ta có

$$Q_3 = 7 + \frac{\frac{3.86}{20} - 51}{20} \cdot 1 = 7,675$$

Tứ phân vị thứ ba Q_3 chia mẫu số liệu thành 2 phần, phần dưới chiếm 75% số liệu của mẫu và phần trên chiếm 25% số liệu của mẫu.

Vậy 75% học sinh khối 11 ngủ ít nhất 7,675 giờ.

Câu 43. 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

a) Cho biết các nhóm số liệu và tần số tương ứng.

b) Tính số trung bình, trung vị và giải thích ý nghĩa của các giá trị thu được.

Lời giải

a) Có 3 nhóm số liệu gồm 90 - 93, 94 - 96, 97 - 99 với tần số tương ứng là 12, 31, 7.

b) 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 $(89,5 + 93,5) / 2 = 91,5; (93,5 + 96,5) / 2 = 95$ và $(96,5 + 99,5) / 2 = 98$.

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 $n/2 = 25$ 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

$$M_e = 93,5 + \frac{25 - 12}{31} \cdot 3 \approx 94,76.$$

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 44. Quãng đường (km) các cầu thủ (không tính thủ môn) chạy trong một trận bóng đá tại giải ngoại hạng Anh được cho trong bảng thống kê sau:

Quãng đường	[2;4)	[4;6)	[6;8)	[8;10)	[10;12)
Số cầu thủ	2	5	6	9	3

Tìm trung vị của mẫu số liệu và giải thích ý nghĩa của giá trị thu được.

Lời giải

Cỡ mẫu $n = 2 + 5 + 6 + 9 + 3 = 25$. Nhóm chứa trung vị là $[6;8)$. Trung vị là

$$M_e = 6 + \frac{\frac{25}{2} - (2 + 5)}{6} \cdot (8 - 6) \approx 7,83$$

Có 50% số cầu thủ chạy nhiều hơn 7,83 km và có 50% số cầu thủ chạy ít hơn 7,83 km.

Câu 45. 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, mốt 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

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

Nhóm chứa mốt của mẫu số liệu ghép nhóm trên là nhóm $[6;9)$.

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

$$\text{Mốt của mẫu số liệu là } M_o = 6 + \frac{(57 - 24)}{(57 - 24) + (57 - 42)} \cdot (9 - 6) = 8,0625.$$

Gọi $x_1; x_2; \dots; x_{160}$ là mẫu số liệu được xếp theo thứ tự không giảm.

Ta có: $x_1, \dots, x_{24} \in [3;6); x_{25}, \dots, x_{81} \in [6;9); x_{82}, \dots, x_{123} \in [9;12); x_{124}, \dots, x_{152} \in [12;15); x_{153}, \dots, x_{160} \in [15;18)$.

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 $[6;9)$ nên trung vị của mẫu số liệu là $M_e = 6 + \frac{\frac{160}{2} - 24}{57} \cdot (9 - 6) \approx 8,95$.

b) 25% các hộ gia đình có lượng nước tiêu thụ cao nhất có lượng nước tiêu thụ không nhỏ hơn Q_3 , với Q_3 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 $11,79 m^3$ nước trở lên.

Câu 46. 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, một 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, loại vừa và loại to. Các loại này lần lượt chiếm khoảng 25%, 50% và 25% 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

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

Nhóm chứa một của mẫu số liệu là nhóm $[84;86)$.

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

Mốt của mẫu số liệu là $M_o = 84 + \frac{(24 - 20)}{(24 - 20) + (24 - 15)} \cdot (86 - 84) \approx 84,62$.

Gọi $x_1; x_2; \dots; x_{90}$ là mẫu số liệu được xếp theo thứ tự không giảm.

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

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

Cỡ mẫu $n = 90$ là số chẵn nên trung vị là $M_e = \frac{1}{2}(x_{45} + x_{46})$.

Do x_{45} và x_{46} thuộc nhóm $[84;86)$ nên trung vị của mẫu số liệu là

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

b) Gọi Q_1, Q_3 lần lượt là tứ phân vị thứ nhất và thứ ba của mẫu số liệu. Theo đề bài, ta có:

Măng cụt loại nhỏ có cân nặng nhỏ hơn Q_1 .

Măng cụt loại vừa có cân nặng trong $[Q_1; Q_3)$.

Măng cụt loại to có cân nặng không nhỏ hơn Q_3 .

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 $x_{68} \in [86; 88)$.

Do đó, tứ phân vị thứ nhất là $Q_1 = 82 + \frac{\frac{90}{4} - 18}{20} \cdot (84 - 82) = 82,45$;

tứ phân vị thứ ba là $Q_3 = 86 + \frac{\frac{90 \cdot 3}{4} - (18 + 20 + 24)}{15} \cdot (88 - 86) \approx 86,73$.

Vậy măng cụt loại nhỏ có khối lượng (tính theo gam) thuộc $[80; 82,45)$.

Măng cụt loại vừa có khối lượng (tính theo gam) thuộc $[82,45; 86,73)$.

Măng cụt loại to có khối lượng (tính theo gam) thuộc $[86,73; 90)$.

Câu 47. 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.

b) Tổng hợp lại dãy số liệu trên vào bảng tần số ghép nhóm với nhóm đầu tiên là $[0; 60)$. 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; 253; 277; 280; 298.

Cỡ mẫu $n = 30$ là số chẵn nên trung vị của mẫu số liệu là

$Q_2 = \frac{1}{2}(x_{15} + x_{16}) = \frac{1}{2}(168 + 175) = 171,5$. Tứ phân vị thứ nhất là $Q_1 = x_8 = 89$.

Tứ phân vị thứ ba là $Q_3 = x_{23} = 223$.

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

Gọi $x_1; x_2; \dots; x_{30}$ là mẫu số liệu được xếp theo thứ tự không giảm.

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

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

Tứ phân vị thứ hai của mẫu số liệu ghép nhóm là $\frac{1}{2}(x_{15} + x_{16})$. Do x_{15} và x_{16} thuộc nhóm $[120; 180)$ nên tứ phân vị thứ hai của mẫu số liệu ghép nhóm là

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

Tứ phân vị thứ nhất của mẫu là $x_8 \in [60; 120)$. Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là

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

Tứ phân vị thứ ba của mẫu số liệu là $x_{23} \in [180; 240)$. Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm

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

Câu 48. 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

Cỡ mẫu là $n = 45$ là số lẻ nên trung vị của mẫu số liệu là $Q_2 = x_{23} = 30$.

$$\text{Tứ phân vị thứ nhất là } Q_1 = \frac{1}{2}(x_{11} + x_{12}) = \frac{1}{2}(24 + 24) = 24.$$

$$\text{Tứ phân vị thứ ba là } Q_3 = \frac{1}{2}(x_{34} + x_{35}) = \frac{1}{2}(36 + 36) = 36.$$

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

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} thuộc nhóm $[20, 5; 25, 5)$ nên tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là

$$Q_1 = 20,5 + \frac{\frac{45}{2} - 6}{10} \cdot (25,5 - 20,5) = 23,125.$$

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} thuộc nhóm $[35,5; 40,5)$ nên tứ phân vị thứ ba của mẫu số liệu ghép nhóm là

$$Q_3 = 35,5 + \frac{\frac{3.45}{4} - (6+10+8+8)}{13} \cdot (40,5 - 35,5) \approx 36,173.$$

Câu 49. 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 50. 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 51. 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 $\bar{x}_A < \bar{x}_B$.

So sánh theo trung vị: $M_e(A) = \frac{90}{13}; M_e(B) = 10, M_e(A) < M_e(B)$.

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

Câu 52. 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

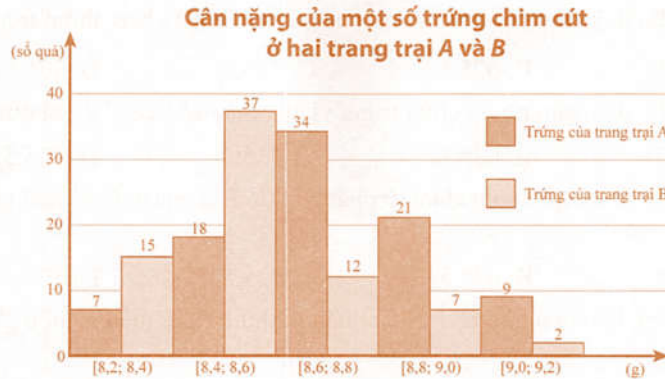
- a) Hãy ước lượng số trung bình, mốt và trung vị của mẫu số liệu ghép nhóm trên.
b) Thầy giáo dự định chọn 25% học sinh có số lần kéo thấp nhất để bồi dưỡng thể lực thêm. Thầy giáo nên chọn học sinh có thành tích kéo xà đơn dưới bao nhiêu lần để bồi dưỡng thể lực?

Lời giải

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

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 53. 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 trại A	7	18	34	21	9
Số quả trứng của 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

$$\bar{x}_A = \frac{8,3 \cdot 7 + 8,5 \cdot 18 + 8,7 \cdot 34 + 8,9 \cdot 21 + 9,1 \cdot 9}{89} \approx 8,72(\text{g}).$$

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

$$\text{Trung vị của mẫu số liệu ghép nhóm là } M_e(A) = 8,6 + \frac{\frac{89}{2} - 25}{34} \cdot (8,8 - 8,6) \approx 8,71.$$

- Đối với trang trại B :

Cỡ mẫu $n_B = 73$.

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

$$\bar{x}_B = \frac{8,3 \cdot 15 + 8,5 \cdot 37 + 8,7 \cdot 12 + 8,9 \cdot 7 + 9,1 \cdot 2}{73} \approx 8,55.$$

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

Trung vị của mẫu số liệu ghép nhóm là $M_e(B) = 8,4 + \frac{\frac{73}{2} - 15}{37} \cdot (8,6 - 8,4) \approx 8,52$.

Ta thấy $\bar{x}_A > \bar{x}_B$ và $M_e(A) > M_e(B)$. 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 không giảm của dãy và thuộc nhóm $[8,4;8,6)$, 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 thứ tự không giảm của dãy và thuộc nhóm $[8,8;9,0)$, nên tứ phân vị thứ ba của mẫu số liệu ghép nhóm là

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

Câu 54. 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, mốt 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 55. 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 mốt của mẫu số liệu trên.

b) Cô giáo dự định trao danh hiệu "Gia đình xanh" cho 25% 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.

Mốt của mẫu số liệu là 13.

b) Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là $Q_1 = \frac{21}{2} = 10,5$.

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

Câu 56. 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, mốt và trung vị của mẫu số liệu ghép nhóm trên.

b) Trung tâm thương mại dự định sẽ thưởng cho 25% 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

$$a) \bar{x} = \frac{815}{18}; M_o = \frac{400}{9}; M_e = 45.$$

b) Tứ phân vị thứ ba của mẫu số liệu ghép nhóm là $Q_3 = \frac{380}{7} \approx 54,29$.

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 57. 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

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

Câu 58. 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

Cỡ mẫu $n = 100$.

Số trung bình của mẫu số liệu ghép nhóm trên là

$$\bar{x} = \frac{3 \cdot 12 + 5 \cdot 20 + 7 \cdot 37 + 9 \cdot 21 + 11 \cdot 10}{100} = 6,94.$$

Nhóm chứa một của mẫu số liệu trên là nhóm $[6; 8)$.

Do đó: $u_m = 6; n_m = 37, n_{m-1} = 20; n_{m+1} = 21; u_{m+1} = 8$.

Vậy một của mẫu số liệu ghép nhóm là

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

Gọi $x_1; x_2; \dots; x_{100}$ là mẫu số liệu được xếp theo thứ tự không giảm.

Tứ phân vị thứ hai của dãy số liệu $x_1; x_2; \dots; x_{100}$ là $\frac{1}{2}(x_{50} + x_{51})$.

$$\text{Do } x_{50} \text{ và } x_{51} \text{ thuộc nhóm } [6; 8) \text{ nên } Q_2 = 6 + \frac{2 \cdot 100}{37} - 32 \cdot (8 - 6) \approx 6,97.$$

Tứ phân vị thứ nhất của dãy số liệu $x_1; x_2; \dots; x_{100}$ là $\frac{1}{2}(x_{25} + x_{26})$. Do x_{25} và x_{26} thuộc nhóm [4;6) nên

$$Q_1 = 4 + \frac{\frac{1 \cdot 100}{4} - 12}{20} \cdot (6 - 4) = 5,3.$$

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 59. Xác định trung vị và tứ phân vị của mẫu số liệu ghép nhóm được cho ở Bảng 5 (làm tròn các kết quả đến hàng phần mười).

Nhóm	Tần số	Tần số tích lũy
[40;45)	5	5
[45;50)	10	15
[50;55)	7	22
[55;60)	9	31
[60;65)	7	38
[65;70)	4	42
	$n = 42$	

Bảng 5

Giải

Số phần tử của mẫu là $n = 42$.

- Ta có: $\frac{n}{2} = \frac{42}{2} = 21$ mà $15 < 21 < 22$. Suy ra nhóm 3 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 21.

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

Áp dụng công thức, ta có trung vị của mẫu số liệu là: $M_e = 50 + \left(\frac{21-15}{7}\right) \cdot 5 \approx 54,3(km/h)$.

Tứ phân vị thứ hai của mẫu số liệu là: $Q_2 = M_e \approx 54,3(km/h)$.

- Ta có: $\frac{n}{4} = \frac{42}{4} = 10,5$ mà $5 < 10,5 < 15$ nên nhóm 2 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 10,5.

Xét nhóm 2 là nhóm [45;50) có $s = 45, h = 5, n_2 = 10$ và nhóm 1 là nhóm [40;45) có $cf_1 = 5$.

Áp dụng công thức, ta có tứ phân vị thứ nhất của mẫu số liệu là:

$$Q_1 = 45 + \left(\frac{10,5-5}{10}\right) \cdot 5 \approx 47,8(km/h).$$

- Ta có: $\frac{3n}{4} = \frac{3 \cdot 42}{4} = 31,5$ mà $31 < 31,5 < 38$ nên nhóm 5 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 31,5.

Xét nhóm 5 là nhóm [60;65) có $t = 60, l = 5, n_5 = 7$ và nhóm 4 là nhóm [55;60) có $cf_4 = 31$.

Áp dụng công thức, ta có tứ phân vị thứ ba của mẫu số liệu là:

$$Q_3 = 60 + \left(\frac{31,5-31}{7}\right) \cdot 5 \approx 60,4(km/h).$$

Vậy tứ phân vị của mẫu số liệu trên là: $Q_1 \approx 47,8(km/h); Q_2 \approx 54,3(km/h); Q_3 \approx 60,4(km/h)$.

Câu 60. Cho mẫu số liệu ghép nhóm thống kê thời gian sử dụng điện thoại trước khi ngủ (đơn vị: phút) của một người trong 120 ngày như ở Bảng 8. Xác định các số đặc trưng đo xu thế trung tâm cho mẫu số liệu đó (làm tròn các kết quả đến hàng phần mười).

Nhóm	Tần số
[0; 4)	13
[4; 8)	29
[8; 12)	48
[12; 16)	22
[16; 20)	8
	$n = 120$

Bảng 8
Lời giải

Số trung bình cộng là:

$$\bar{x} = \frac{13 \cdot 2 + 29 \cdot 6 + 48 \cdot 10 + 22 \cdot 14 + 8 \cdot 18}{120} \approx 9,4.22$$

Bảng tần số ghép nhóm bao gồm cả tần số tích lũy được cho như ở Bảng 10:

Nhóm	Tần số	Tần số tích lũy
[0; 4)	13	13
[4; 8)	29	42
[8; 12)	48	90
[12; 16)	22	112
[16; 20)	8	120
	$n = 120$	

Bảng 10

Ta có: $\frac{n}{2} = 60, \frac{n}{4} = 30, \frac{3n}{4} = 90$.

Vì $42 < 60 < 90$ nên nhóm 3 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 60.

Suy ra trung vị là: $M_e = 8 + \left(\frac{60 - 42}{48} \right) \cdot 4 = 9,5$.

Tứ phân vị thứ hai là: $Q_2 = M_e = 9,5$.

Vì $13 < 30 < 42$ nên nhóm 2 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 30. Suy ra tứ phân vị thứ nhất là: $Q_1 = 4 + \left(\frac{30 - 13}{29} \right) \cdot 4 \approx 6,3$

Vì $42 < 90 \leq 90$ nên nhóm 3 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 90. Suy ra tứ phân vị thứ ba là: $Q_3 = 8 + \left(\frac{90 - 42}{48} \right) \cdot 4 = 12$

Trong các nhóm, nhóm 3 có tần số lớn nhất. Suy ra một là: $M_0 = 8 + \left(\frac{48 - 29}{2 \cdot 48 - 29 - 22} \right) \cdot 4 \approx 9,7$.

Câu 61. Kết quả khảo sát cân nặng của 1 thùng táo ở một lô hàng cho trong bảng sau:

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

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

Lời giải

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

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

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

Vậy trung vị của mẫu số liệu ghép nhóm là

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

Câu 62. 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

$$x_1; x_2; \dots; x_{240} \text{ là } \frac{1}{2}(x_{133} + x_{134}) \in [15,5;20,5).$$

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 63. 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 64. 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 lí hay không?

Lời giải

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

Tứ phân vị thứ hai của dãy số liệu $x_1 \leq x_2 \leq \dots \leq x_{110}$ là $\frac{1}{2}(x_{55} + x_{56})$. Do $x_{55} \in [10;15)$ và $x_{56} \in [15;20)$. Nên đó tứ phân vị thứ hai của mẫu số liệu ghép nhóm là $Q_2 = 15$

Tứ phân vị thứ nhất của dãy số liệu $x_1; x_2; \dots; x_{110}$ là $\frac{1}{2}(x_{27} + x_{28})$. Do $x_{27} \in [10;15)$ và $x_{28} \in [10;15)$ đó tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là

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

Tứ phân vị thứ ba của dãy số liệu $x_1; x_2; \dots; x_{110}$ là $\frac{1}{2}(x_{82} + x_{83})$. Do $x_{82} \in [15;20)$ và $x_{83} \in [20;25)$ đó tứ phân vị thứ ba của mẫu số liệu ghép nhóm là $Q_3 = 20$.

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

Câu 65. Lương tháng của một số giáo viên THPT được ghi lại như sau (đơn vị: triệu đồng):

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

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

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

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

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

Lời giải

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

b) Bảng tần số ghép nhóm theo mẫu :

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

c) Tứ phân vị thứ nhất là $Q_1 = 8 + \frac{\frac{24}{4} - 3}{6} \cdot (10 - 8) = 9$.

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

Tứ phân vị thứ ba là $Q_3 = 12 + \frac{\frac{3 \cdot 24}{4} - 17}{7} \cdot (14 - 12) \approx 12,29$.

Câu 66. Một học sinh làm bài kiểm tra môn Toán gồm 30 câu trong 1 tháng được thống kê lại như sau:

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

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

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

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

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

Lời giải

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

b) Bảng tần số ghép nhóm theo mẫu :

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

c) Tứ phân vị thứ nhất là $Q_1 = 11 + \frac{\frac{1 \cdot 20}{4} - 4}{8} \cdot (15 - 11) = 11,5$.

Tứ phân vị thứ hai là $Q_2 = 15,5$.

Tứ phân vị thứ ba là $Q_3 = 21 + \frac{\frac{3 \cdot 20}{4} - 14}{6} \cdot (25 - 21) \approx 21,67$.

Câu 67. Trong một cuộc đua Marathon được tổ chức ở thành phố A người ta thống kê lại được như sau

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

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

Lời giải

Ta có $n = 4 + 6 + 10 + 15 + 25 = 60$

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

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

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

Câu 68. Chiều cao của hai loài hoa được một người thống kê theo biểu đồ sau:



- a) Hãy so sánh chiều cao của loài A và loài B theo số trung bình và trung vị.
b) Hãy ước lượng tứ phân vị thứ nhất và thứ ba của chiều cao của loài A và loài B.

Lời giải

a) Bảng giá trị đại diện

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

Chiều trung bình của loài A xấp xỉ bằng

$$\bar{x}_A = \frac{149,5.20 + 249,5.18 + 349,5.14 + 449,5.10}{62} \approx 272,08.$$

Chiều trung bình của loài B xấp xỉ bằng

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

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

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

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

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

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

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

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

Chiều cao trung bình của loài A lớn hơn B .

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

Do đó tứ phân vị thứ nhất của loài A là

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

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

Do đó tứ phân vị thứ ba của loài A là

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

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

Do đó tứ phân vị thứ nhất của loài B là

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

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

Do đó tứ phân vị thứ ba của loài B là

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

Câu 69. 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

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

□ Tìm trung vị

□ Khoảng chứa trung vị [50;60)

Trung vị của mẫu số liệu là

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

□ Tìm tứ phân vị

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

Khi đó:

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

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

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

□ Tìm một

Nhóm chứa một: [50;60)

Một của mẫu số liệu:

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

Câu 70. Mẫu số liệu sau ghi lại cân nặng của 30 bạn học sinh (đơn vị: kilogam)

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

- a) Lập bảng tần số ghép nhóm cho mẫu số liệu trên có tám nhóm ứng với tám nửa khoảng [15;20), [20;25), [25;30), [30;35), [35;40), [40;45), [45;50), [50;55).
b) Xác định số trung bình cộng, 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)

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

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

Trung vị Q_1 và Q_3 :

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

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

Tứ 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 71. An tìm hiểu hàm lượng chất béo (đơn vị: g) có trong 100 g mỗi loại thực phẩm. Sau khi thu thập dữ liệu về 60 loại thực phẩm, An lập được bảng thống kê

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

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

b) Tìm một của mẫu số liệu.

Lời giải

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

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

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

Trung vị thứ nhất và thứ ba:

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

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

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

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

Câu 72. Để chuẩn bị cho đồ án tốt nghiệp, một sinh viên y khoa đã khảo sát huyết áp tối đa của một số bệnh nhân và lập được bảng tần số ghép nhóm sau:

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

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

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

Lời giải

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

a) Giá trị trung bình:

$$\bar{x} = \frac{100 \cdot 6 + 120 \cdot 20 + 140 \cdot 35 + 160 \cdot 45 + 180 \cdot 30 + 200 \cdot 16}{152} = 155,9$$

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

Trung vị thứ nhất và thứ ba:

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

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

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

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

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

Câu 73. Một câu lạc bộ thể dục thể thao đã ghi lại số giờ các thành viên của mình sử dụng cơ sở vật chất của câu lạc bộ để tập luyện trong một tháng. Họ tổ chức dữ liệu thu được vào bảng

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

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

a) Trung vị của mẫu số liệu.

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

Lời giải

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

a) Trung vị của mẫu số liệu

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

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

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