

PHIẾU SỐ 01

ĐIỂM SỐ

Họ tên:

Nhận xét:

(Ước mơ chỉ thành hiện thực khi bạn nỗ lực hành động,

Hãy hành động vì ước mơ của bạn !)

=====^^=====

BÀI 1. CÁC SỐ ĐẶC TRƯNG CỦA MẪU SỐ LIỆU GHÉP NHÓM

I. SỐ TRUNG BÌNH

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

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

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

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

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

Câu 1: Cân nặng của 28 học sinh nam lớp 11 được cho như sau:

Cân nặng	[45;49)	[49;53)	[53;57)	[57;61)	[61;65)
Số học sinh	4	5	7	7	5

a) 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) 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:

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

Câu 2: 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

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

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

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

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

II. MỐT

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

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

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

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 $n_{m+1} = 0$.

Câu 3: 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 $u_m = 18, n_{m-1} = 78, n_m = 120, n_{m+1} = 45, u_{m+1} - u_m = 22 - 18 = 4$.

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 4: 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

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.

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

III. SỐ TRUNG VỊ

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

- Gọi n là cỡ mẫu.

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

Khi đó

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

Câu 5: 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

Cỡ mẫu : $n = 4 + 7 + 12 + 6 + 2 = 31$

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

Do n là số lẻ nên $M_e = x_{16} \in [160;165)$

Suy ra nhóm chứa trung vị là : $[u_m; u_{m+1}) = [160;165)$

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

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

IV. TỨ PHÂN VỊ

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

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

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

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

Khi đó

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

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

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

Khi đó

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

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

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

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

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

Câu 7: Mức lương hàng tháng ở 1 công ty được Công đoàn thu thập theo bảng sau(đơn vị triệu đồng):

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

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

Lời giải

- 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.

BÀI TẬP VỀ NHÀ

TRẮC NGHIỆM

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

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

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

Câu 1: Giá trị đại diện của nhóm $[25;30)$ là

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

Câu 2: Xác định số trung bình cộng của mẫu số liệu ghép nhóm trên.

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

Câu 3: Một của mẫu số liệu ghép nhóm trên là bao nhiêu

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

Câu 4: Số lượng huy chương vàng tại Sea Games 31 được thống kê

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

Xác định trung vị của mẫu số liệu ghép nhóm trên

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

Lời giải

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

Câu 5: Đo cân nặng của 1 lớp gồm 40 học sinh lớp 12B

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

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

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

Lời giải

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

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

Câu 6: Đo cân nặng của 1 lớp gồm 40 học sinh lớp 12B

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

A. [40;45] .

B. [45;50] .

C. [50;55] .

D.

[55;60]

Lời giải

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

TỰ LUẬN.

Câu 7: Cho bảng tần số ghép nhóm số liệu thống kê chiều cao của 40 mẫu cây ở một vườn thực vật (đơn vị: centimét).

Chiều cao(m)	[30;40)	[40;50)	[50;60)	[60;70)	[70;80)	[80;90)
Số cây	4	10	14	6	4	2

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) Ta có bảng tần số ghép nhóm của mẫu số liệu:

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

Giá trị trung bình của mẫu số liệu là

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

b) Nhóm chứa mốt: [50;60)

Mốt của mẫu số liệu:

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

Câu 8: 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 tìm tứ phân vị của mẫu số liệu ghép nhóm trên.

Lời giải

Ta có $n = 4 + 6 + 10 + 15 + 25 = 60$

Tứ phân vị thứ hai là $Q_2 = M_e = 180 + \frac{\frac{60}{2} - 20}{15} \cdot (200 - 180) = \frac{580}{3} \approx 193,33$.

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 $Q_3 = 200 + \frac{3 \cdot \frac{60}{4} - 35}{25} \cdot (220 - 200) = 208$.