



CÁC SỐ ĐẶC TRƯNG CỦA MẪU SỐ LIỆU KHÔNG GHÉP NHÓM

BÀI 13. CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM



LÝ THUYẾT.

1. SỐ TRUNG BÌNH VÀ TRUNG VỊ

a. Số trung bình

Số trung bình (số trung bình cộng) của mẫu số liệu $x_1, x_2, \dots, x_n$, kí hiệu là $\bar{x}$, được tính bằng công thức:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Chú ý. Trong trường hợp mẫu số liệu cho dưới dạng bảng tần số thì số trung bình được tính theo công thức:

$$\bar{x} = \frac{m_1 x_1 + m_2 x_2 + \dots + m_k x_k}{n}$$

Trong đó m_k là tần số của giá trị x_k và $n = m_1 + m_2 + \dots + m_k$.

Ý nghĩa. Số trung bình là giá trị trung bình cộng của các số trong mẫu số liệu, nó cho biết vị trí trung tâm của mẫu số liệu và có thể dùng để đại diện cho mẫu số liệu.

b. Trung vị

Để tìm trung vị của một mẫu số liệu, ta thực hiện như sau:

- Sắp xếp các giá trị trong mẫu số liệu theo thứ tự không giảm.
- Nếu số giá trị của mẫu số liệu là số lẻ thì giá trị chính giữa của mẫu là trung vị. Nếu là số chẵn thì trung vị là trung bình cộng của hai giá trị chính giữa của mẫu.

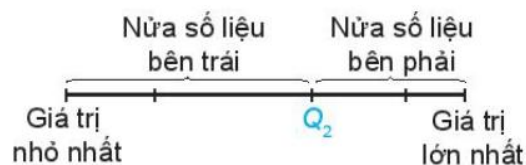
Ý nghĩa. Trung vị là giá trị chia đôi mẫu số liệu, nghĩa là trong mẫu số liệu được sắp xếp theo thứ tự không giảm thì giá trị trung vị ở vị trí chính giữa. Trung vị không bị ảnh hưởng bởi giá trị bất thường trong khi số trung bình bị ảnh hưởng bởi giá trị bất thường.

2. TỨ PHÂN VỊ

Để tìm các tứ phân vị của mẫu số liệu có n giá trị, ta làm như sau:

- Sắp xếp mẫu số liệu theo thứ tự không giảm.
- Tìm trung vị. Giá trị này là Q_2 .
- Tìm trung vị của nửa số liệu bên trái Q_2 (không bao gồm Q_2 nếu n lẻ). Giá trị này là Q_1 .
- Tìm trung vị của nửa số liệu bên phải Q_2 (không bao gồm Q_2 nếu n lẻ). Giá trị này là Q_3 .

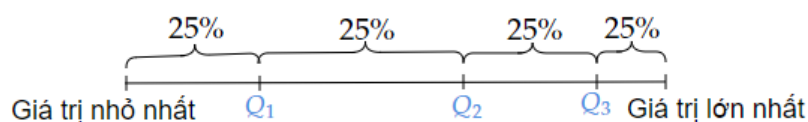
Q_1, Q_2, Q_3 được gọi là các **tứ phân vị** của mẫu số



Hình 5.3b

Chú ý. Q_1 được gọi là tứ phân vị thứ nhất hay tứ phân vị dưới, Q_3 được gọi là tứ phân vị thứ ba hay tứ phân vị trên.

Ý nghĩa. Các điểm Q_1, Q_2, Q_3 chia mẫu số liệu đã sắp xếp theo thứ tự từ nhỏ đến lớn thành bốn phần, mỗi phần đều chứa 25% giá trị (hình 5.3a).



Hình 5.3a. Các tứ phân vị

VÍ DỤ: Hàm lượng Natri (đơn vị miligam, $1mg = 0,001g$) trong 100 g một số loại ngũ cốc được cho như sau:

0	340	70	140	200	180	210	150	100	130
140	180	190	160	290	50	220	180	200	210.

Hãy tìm các tứ phân vị. Các phân vị này cho ta thông tin gì?

Giải

- Sắp xếp các giá trị này theo thứ tự không giảm:

0 50 70 100 130 140 140 150 160 180 180 180 190 200 200 210 210 220 290 340.

Hai giá trị chính giữa

- Vì $n = 20$ là số chẵn nên Q_2 là trung bình cộng của hai giá trị chính giữa:

$$Q_2 = (180 + 180) : 2 = 180.$$

- Ta tìm Q_1 là trung vị của nửa số liệu bên trái Q_2 :

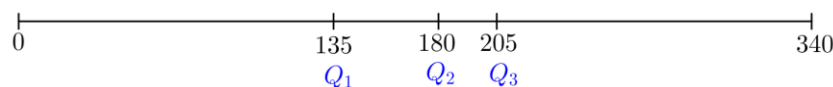
0 50 70 100 130 140 140 150 160 180.

và ta tìm được $Q_1 = (130 + 140) : 2 = 135$.

- Ta tìm Q_3 là trung vị của nửa số liệu bên phải Q_2 :

180 180 190 200 200 210 210 220 290 340.

và tìm được $Q_3 = (200 + 210) : 2 = 205$.



Hình 5.4. Hình ảnh về sự phân bố của mẫu số liệu

Các tứ phân vị cho ta hình ảnh phân bố của mẫu số liệu. Khoảng cách từ Q_1 đến Q_2 là 45 trong khi khoảng cách từ Q_2 đến Q_3 là 25. Điều này cho thấy mẫu số liệu tập trung mật độ cao ở bên phải Q_2 và mật độ thấp ở bên trái Q_2 (H.5.4).

3. MỐT

Mốt của mẫu số liệu là giá trị xuất hiện với tần số lớn nhất.

Ý nghĩa. Có thể dùng một để đo xu thế trung tâm của mẫu số liệu khi mẫu số liệu có nhiều giá trị trùng nhau.



BÀI TẬP SÁCH GIÁO KHOA.

5.7. Tìm số trung bình, trung vị, mốt và tứ phân vị của mỗi mẫu số liệu sau đây:

- a) Số điểm mà năm vận động viên bóng rổ ghi được trong một trận đấu:

9 8 15 8 20

- b) Giá của một số loại giày (đơn vị nghìn đồng):

350 300 650 300 450 500 300 250 .

- c) Số kênh được chiếu của một số hãng truyền hình cáp:

36 38 33 34 32 30 34 35.

5.8. Hãy chọn số đặc trưng đo xu thế trung tâm của mỗi mẫu số liệu sau. Giải thích và tính giá trị của số đặc trưng đó.

a) Số mặt trăng đã biết của các hành tinh:

Hành tinh	Thu ỷ tinh	Kim tinh	Trái Đ ất	Hoả tinh	Mộc tinh	Thổ t i n h	Thiên Vương tinh	Hải Vươn g tinh
Số mặt trăng	0	0	1	2	63	34	27	13

(Theo NASA)

b) Số đường chuyền thành công trong một trận đấu của một số cầu thủ bóng đá:

32 24 20 14 23.

c) Chỉ số IQ của một nhóm học sinh: 60 72 63 83 68 74 90 86 74 80.

d) Các sai số trong phép đo: 10 15 18 15 14 13 42 15 12 14 42.

5.9. Số lượng học sinh giỏi Quốc gia năm học 2018 - 2019 của 10 trường Trung học phổ thông được cho như sau: 0 0 4 0 0 0 10 0 6 0.

a) Tìm số trung bình, mốt, các tứ phân vị của mẫu số liệu trên.

b) Giải thích tạo sao tứ phân vị thứ nhất và trung vị trùng nhau.

Giải

a) Số trung bình là $\frac{0.7 + 4 + 6 + 10}{10} = 2$.

Sắp xếp số liệu theo thứ tự không giảm 0 0 0 0 0 0 0 4 6 10.

Số 0 xuất hiện nhiều nhất nên mốt là 0.

Tứ phân vị $Q_1 = 0$; $Q_2 = 0$; $Q_3 = 4$.

b) Tứ phân vị thứ nhất và trung vị trùng nhau do mẫu có 10 số liệu mà số 0 đã xuất hiện 7 lần.

5.10. Bảng sau đây cho biết số chỗ ngồi của một số sân vận động được sử dụng trong Giải Bóng đá Vô địch Quốc gia Việt Nam năm 2018 (số liệu gần đúng).

Sân vận động	Cẩm phả	Thiên Trường	Hàng Đẫy	Thanh Hoá	Mỹ Đình
Chỗ ngồi	20 120	21 315	23 405	20 120	37 546

(Theo vov.vn)

Các giá trị số trung bình, trung vị, mốt bị ảnh hưởng như thế nào nếu bỏ đi số liệu chỗ ngồi của Sân vận động Quốc gia Mỹ Đình?