CHƯƠNG

**V**

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

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

**LÝ THUYẾT.**

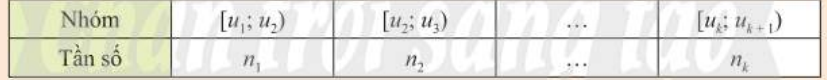
**I ===I**

**1. SỐ LIỆU GHÉP NHÓM**

Một số loại số liệu điều tra có thể nhận rất nhiều những giá trị khác nhau, hoặc khó xác định được giá trị chính xác, ví dụ như chiều cao, cân nặng, tuổi thọ, … Để thuận tiện cho việc lưu trữ và xử lí các loại số liệu này, người ta thường ghép các số liệu gần nhau lại thành nhóm.

Mẫu số liệu ghép hóm thường được trình bày dưới dạng bảng thống kê có dạng như sau :

**Bảng 1 : Bảng tần số ghép nhóm**



Chú ý :

* Bảng trên gồm  nhóm  với , mỗi nhóm gồm một số giá trị được ghép theo một tiêu chí xác định.
* Cỡ mẫu .
* Giá trị chính giữa của mỗi nhóm được dùng làm giá trị đại diện cho nhóm ấy. Ví dụ nhóm  có giá trị đại diện là .
* Hiệu  được gọi là độ dài của nhóm  .

Một số quy tắc ghép nhóm của mẫu số liệu

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

- Sử dụng từ  đến  nhóm. Cỡ mẫu càng lớn thì cần càng nhiều nhóm số liệu.

Các nhóm có cùng độ dài bằng  thoả mãn , trong đó  là khoảng biến thiên,  là số nhóm.

- Giá trị nhỏ nhất của mẫu số thuộc vào nhóm  và càng gần  càng tốt. Giá trị lớn nhất của mẫu thuộc nhóm  và càng gần  càng tốt.

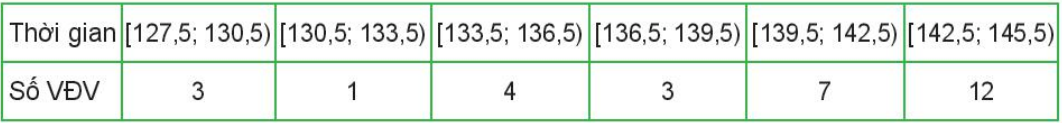
**Ví dụ.** Bảng thống kê sau cho biết thời gian chạy (phút) của 30 vận động viên (VĐV) trong một giải chạy Marathon.



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

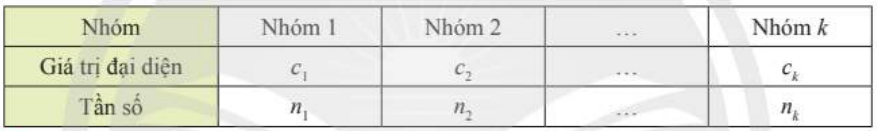
**Lời giải**

Giá trị nhỏ nhất là 129, giá trị lớn nhất là 145 nên khoảng biến thiên là . Tổng độ dài của sáu nhóm là 18. Để cho đối xứng, ta chọn đầu mút trái của nhóm đầu tiên là 27,5 và đầu mút phải của nhóm cuối cùng là 145,5 ta được các nhóm là , . Đếm số giá trị thuộc mỗi nhóm, ta có mẫu số liệu ghép nhóm như sau:



**2. SỐ TRUNG BÌNH**

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



🞜 Số trung bình của mẫu số liệu ghép nhóm, kí hiệu , được tính như sau :

, trong đó .

***Ý nghĩa của số trung bình của mẫu số liệu ghép nhóm***

Số trung bình của mẫu số liệu ghép nhóm là giá trị xấp xỉ cho số trung bình của mẫu số liệu gốc. Nó thường dùng để đo xu thế trung tâm của mẫu số liệu.

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

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

**Lời giải**

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

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

Tổng số học sinh là . Cân nặng trung bình cảu học sinh lớp 11D là



**Ví dụ 2.** Tìm hiểu thời gian xem tivi trong tuần trước (đơn vị: giờ) của một số học sinh thu được kết quả sau:

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

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

**Lời giải**

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Thòi 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 . Thời gian xem tivi trung bình của học sinh là



**3. MỐT**

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

Giả sử nhóm chứa mốt là , khi đó ***mốt của mẫu số liệu ghép nhóm***, kí hiệu là , được xác định bởi công thức



***Chú ý:*** Nếu không có nhóm kể trước của nhóm chứa mốt thì . Nếu không có nhóm kề sau của nhóm chứa mốt thì .

***Ý nghĩa của mốt của mẩu số liệu ghép nhóm***

- Mốt của mẫu số liệu không ghép nhóm là giá trị có khả năng xuất hiện cao nhất khi lấy mẫu. Mốt của mẫu số liệu sau khi ghép nhóm  xấp xi với mốt của mẫu số liệu không ghép nhóm. Các giá trị nằm xung quanh  thường có khả năng xuất hiện cao hơn các giá trị khác.

- Một mẫu số liệu ghép nhóm có thể có nhiều nhóm chứa mốt và nhiều mốt.

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Khoảng chiều cao (cm) |  |  |  |  |  |
| Số học sinh | 7 | 14 | 10 | 10 | 9 |

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

**Lời giải**

Tần số lớn nhất là 14 nên nhóm chứa mốt là nhóm .

Ta có . Do đó .

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

**Ví dụ 2.** Thời gian (phút) để học sinh hoàn thành một câu hỏi thi được cho như sau:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Thời gian (phút) |  |  |  |  |  |
| Số học sinh | 2 | 10 | 6 | 4 | 3 |

Tính mốt của mẫu số liệu ghép nhóm này.

**Lời giải**

Tần số lớn nhất là 10 nên nhóm chứa mốt là nhóm .

Ta có .

Do đó .

**HỆ THỐNG BÀI TẬP TỰ LUẬN.**

**II ===I**

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

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



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



**Lời giải**

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

1. Có năm nhóm là

Dưới 50 nghìn đồng có 5 sinh viên.

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

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

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

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

1. Có bốn nhóm là

Từ C đến dưới C có 7 ngày.

Từ C đến dưới C có 15 ngày.

Từ C đến dưới C có 12 ngày.

Từ C đến dưới C có 6 ngày.

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

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Giá trị |  |  |  |  |  |  |
| Số lượng gia đình | 6 | 15 | 10 | 6 | 9 | 4 |

Đọc và giải thích mẫu số liệu này.

**Lời giải**

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

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

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Giá trị |  |  |  |  |  |
| Số lượng củ khoai | 3 | 6 | 12 | 6 | 3 |

Đọc và giải thích mẫu số liệu này.

**Lời giải**

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

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

1. Số sản phẩm một công nhân làm được trong một ngày được cho như sau



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

**Lời giải**

Khoảng biến thiên là .

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

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



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



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

**Lời giải**

Khoảng biến thiên là .

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

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



1. Bảng thống kê sau cho biết điện năng tiêu thụ của **** hộ ở một khu dân cư trong một tháng như sau (*đơn vị:* ****):

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 50 | 47 | 30 | 65 | 63 | 70 | 38 | 34 | 48 | 53 | 33 | 39 | 32 | 40 | 50 |
| 55 | 50 | 61 | 37 | 37 | 43 | 35 | 65 | 60 | 31 | 33 | 41 | 45 | 55 | 59 |

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

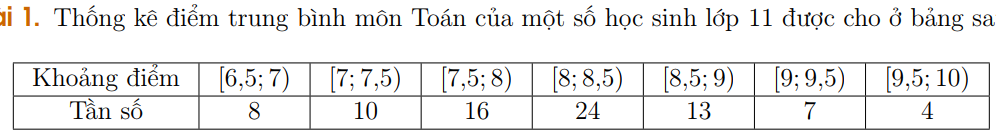
**Lời giải**

Giá trị nhỏ nhất là 30, giá trị lớn nhất là 70 nên khoảng biến thiên là . Tổng độ dài của 8 nhóm là 40 nên ta được các nhóm như sau: .

Đếm số giá trị thuộc mỗi nhóm ta có mẫu số liệu ghép nhóm như sau:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Giá trị |  |  |  |  |  |  |  |  |
| Số lượng | 6 | 5 | 3 | 3 | 4 | 3 | 3 | 3 |

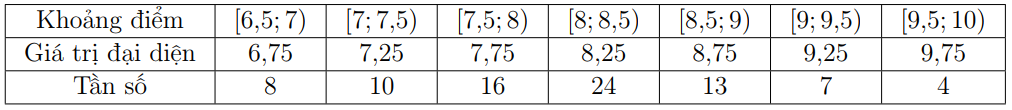
1. Thống kê điểm trung bình môn Toán của một số hoc sinh lớp 11 được cho ở bảng sau:



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

**Lời giải**

Bảng tần số ghép nhóm theo giá trị đại diện là



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

****

**Mốt**

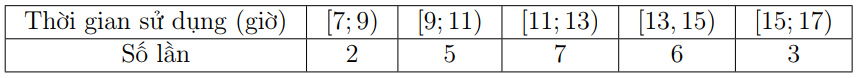
**Mốt  chứa trong nhóm **

**Do đó: **

****

****

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



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

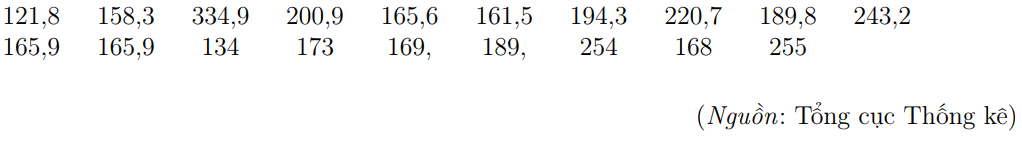
pin.

**Lời giải**

Thời gian sử dụng trung bình:

****

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



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

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

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

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

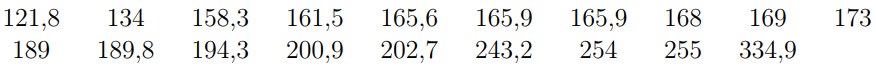
**Lời giải**

**a)**

**Số trung bình: **

**Tứ phân vị:**

Xếp mẫu số liệu không giảm ta được:



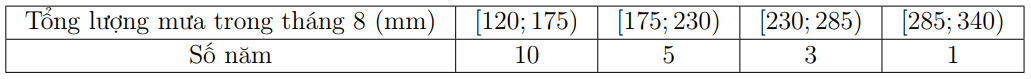
**Từ đó ta có:**

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

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

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

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



Giá trị đại diện của các lớp:



Tần số các lớp: 

Số trung bình: 

**Mốt**

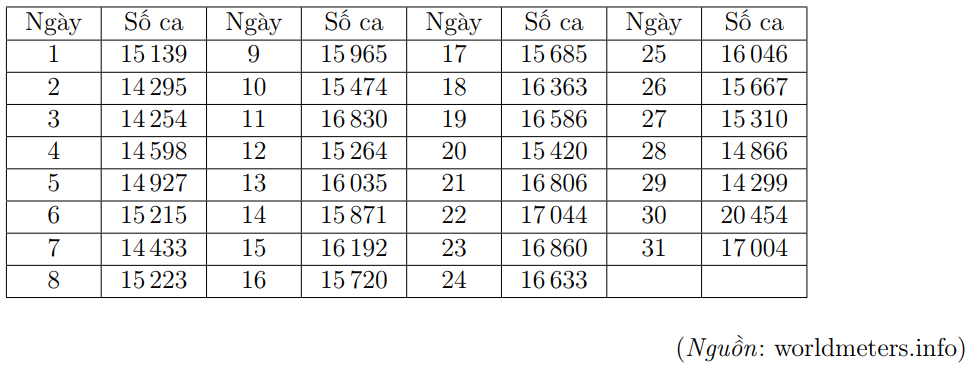
**Mốt  chứa trong nhóm **

**Do đó: **

****

****

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



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

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

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

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

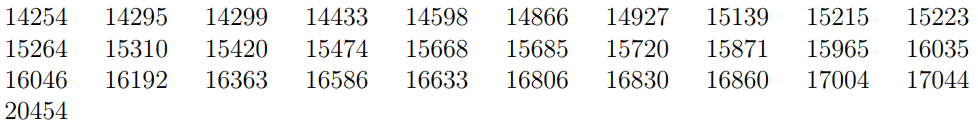
**Lời giải**

**a)**

**Số trung bình: .**

**Tứ phân vị:**

Xếp mẫu số liệu không giảm ta được:



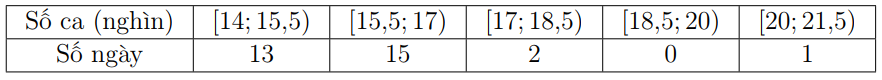
**Từ đó ta có:**

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

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

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

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



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

**Số trung bình: .**

**Mốt**

**Mốt  chứa trong nhóm **

**Do đó: **

****

****