# **CHƯƠNG III. CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CỦA MẪU SỐ LIỆU GHÉP NHÓM**

## **BÀI 8. MẪU SỐ LIỆU GHÉP NHÓM**

### **A. KIẾN THỨC CẦN NHỚ**

**1.** Khi ta không thể thu thập được số liệu chính xác hoặc do yêu cầu của bài toán ta phải biểu diễn mẫu số liệu dưới dạng ghép nhóm để thuận lợi cho việc tổ chức, đọc và phân tích số liệu. Mẫu số liệu ghép nhóm là mẫu số liệu cho dưới dạng bảng tần số của các nhóm số liệu (Bảng 1). Mỗi nhóm số liệu là tập hợp gồm các giá trị của số liệu được ghép nhóm theo một tiêu chí xác định. Nhóm số liệu thường được cho dưới dạng , trong đó  là đầu mút trái,  là đầu mút phải.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  Nhóm  |    |    |   |    |   |
|  Tần số  |    |    |    |    |    |

Bảng 1. Mẫu số liệu ghép nhóm

**2.** Để chuyển mẫu số liệu không ghép nhóm sang mẫu số liệu ghép nhóm ta thực hiện theo các bước sau:

Bước 1: Chia miền giá trị của mẫu số liệu thành một số nhóm theo tiêu chí cho trước.

Bước 2: Đếm số giá trị của mẫu số liệu thuộc mỗi nhóm (tần số) và lập bảng thống kê cho mẫu số liệu ghép nhóm.

Trong các bài tập, ta không nên chia thành quá nhiều hoặc quá ít nhóm. Các nhóm không giao nhau và nên có độ dài bằng nhau, tổng độ dài các nhóm lớn hơn khoảng biến thiên.

### **B. VÍ DỤ**

**Ví dụ 1.** 100 người thực hiện bài trắc nghiệm để đo chỉ số , kết quả thu được như sau:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  Chỉ số   |  Dưới 70  |   |    |   |   |  Từ 145 trở lên  |
|  Số người  |  2  |  15  |  45  |  20  |  15  |  3  |

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

b) Người có chỉ số  từ 85 đến dưới 115 là ở mức trung bình. Xác định tỉ lệ người có  cao hơn mức trung bình.

**Giải**

a) Các nhóm số liệu gồm Dưới 

Từ 145 trở lên với tần số tương ứng là .

b) Số người có chỉ số  cao hơn mức trung bình là . Vậy tỉ lệ người có chỉ số  cao hơn mức trung bình là .

**Ví dụ 2.** Theo Tổ chức Y tế Thế giới (WHO), thiếu máu là tình trạng giảm lượng huyết sắc tố (Hb) dẫn tới sự thiếu cung cấp oxygen cho các mô trong cơ thể. Đối với nam giới trên 15 tuổi, chỉ số Hb (đơn vị tính là g/l) lớn hơn hoặc bằng 130 được xem là không bị thiếu máu, từ 110 đến dưới 130 là thiếu máu mức nhẹ, từ 80 đến dưới 110 là thiếu máu mức vừa, dưới 80 là mức nặng. Đo chỉ số Hb của một số học sinh nam lớp 12 cho kết quả như sau:



Ghép nhóm cho mẫu số liệu này theo mức độ thiếu máu.

**Giải**

Có 25 học sinh có chỉ số Hb từ 130 trở lên (không bị thiếu máu), 7 học sinh có chỉ số Hb từ 110 đến dưới 130 (thiếu máu mức nhẹ), 3 học sinh có chỉ số Hb từ 80 đến dưới 110 (thiếu máu mức trung bình). Ta có mẫu số liệu ghép nhóm:

|  |  |  |  |
| --- | --- | --- | --- |
|  Chỉ số Hb (g/l) |  Từ 130 trở lên  |   |    |
|  Số học   |  25  |  7  |  3  |

### **C. BÀI TẬP**

1. Từ 1/7/2019, dựa trên thu nhập bình quân đầu người (kí hiệu là GNIPC, tính theo đô la Mĩ), Ngân hàng Thế giới xác định một nền kinh tế ở mức thu nhập thấp nếu GNIPC nhỏ hơn 1026, ở mức thu nhập dưới trung bình nếu GNIPC từ 1026 đến dưới 3996, ở mức thu nhập trên trung bình nếu GNIPC từ 3996 đến dưới 12376 và ở mức thu nhập cao nếu GNIPC từ 12376 trở lên (Theo Ngân hàng Thế giới). Thu nhập bình quân đầu người của một số nền kinh tế thuộc khu vực châu Á Thái Bình Dương năm 2021 được cho như sau:



(Theo statistica.com)

a) Ghép nhóm mẫu số liệu trên theo mức thu nhập của nền kinh tế.

b) GNIPC của Việt Nam năm 2021 là 11040. Nền kinh tế Việt Nam được xếp Ở mức nào?

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  Chỉ số AQI  |    |    |    |    |  Trên 200  |
|  Số ngày  |  5  |  11  |  7  |  4  |  3  |

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

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

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

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

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

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  Thời gian (phút)  |   |    |   |   |
|  Só học sinh  |  2  |  7  |  10  |  25  |

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

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

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

### **A. KIẾN THỨC CẦN NHỚ**

**1.** Ở lớp 10, chúng ta đã biết cách tính các số đặc trưng của mẫu số liệu không ghép nhóm (còn gọi là mẫu số liệu gốc) . Tuy nhiên, trong nhiều trường hợp ta không có mẫu số liệu gốc mà chỉ có mẫu số liệu ghép nhóm dạng

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  Nhóm  |    |    |    |    |   |
|  Tần só  |    |    |    |    |    |

Khi đó, các số đặc trưng của mẫu số liệu ghép nhóm xấp xỉ cho các số đặc trưng của mẫu số liệu gốc. Các số đặc trưng đo xu thế trung tâm của mẫu số liệu ghép nhóm gồm số trung bình, trung vị, tứ phân vị và mốt.

**2.** Số trung bình của mẫu số liệu ghép nhóm là 

trong đó,  là tổng số quan sát (còn gọi là cỡ mẫu) và  gọi là giá trị đại diện của nhóm . Số trung bình của mẫu số liệu ghép nhóm xấp xỉ cho số trung bình của mẫu số liệu gốc, nó cho biết vị trí trung tâm của mẫu số liệu và có thể dùng để đại diện cho mẫu số liệu.

**3.** 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 .

Bước 2. Trung vị là 

trong đó  là cỡ mẫu. Với  ta quy ước . Trung vị chính là tứ phân vị thứ hai . Trung vị của mẫu số liệu ghép nhóm xấp xỉ cho trung vị của mẫu số liệu gốc, nó chia mẫu số liệu thành hai phần, mỗi phần chứa  giá trị.

**4.** Để tính tứ phân vị thứ nhất  của mẫu số liệu ghép nhóm trước hết ta xác định nhóm chứa , giả sử đó là nhóm thứ . Khi đó, 

trong đó  là cỡ mẫu, với  ta quy ước .

Để tính tứ phân vị thứ ba  của mẫu số liệu ghép nhóm trước hết ta xác định nhóm chứa . Giả sư đó là nhóm thứ . Khi đó, 

trong đó  là cỡ mẫu, với  ta quy ước .

Để xác định nhóm chứa tứ phân vị thứ  ta có thể dựa vào tính chất có khoảng  số giá trị nhỏ hơn tứ phân vị đó.

Các tứ phân vị  của mẫu số liệu ghép nhóm xấp xỉ cho các tứ phân vị của mẫu số liệu gốc, chúng chia mẫu số liệu thành 4 phần, mỗi phần chứa  giá trị.

**5.** Để tìm mốt 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 mốt), giả sử là nhóm .

Bước 2. Mốt được xác định là 

trong đó  là độ rộng của nhóm và ta quy ước .

Mốt của mẫu số liệu ghép nhóm xấp xỉ cho mốt của mẫu số liệu gốc, nó được dùng để đo xu thế trung tâm của mẫu số liệu.

Lưu ý. Người ta chỉ định nghĩa mốt cho mẫu ghép nhóm có độ dài các nhóm bằng nhau. Một mẫu có thể không có mốt hoặc có nhiều hơn 1 mốt. Khi tần số của các nhóm bằng nhau thì mẫu số liệu ghép nhóm không có mốt.

**6.** Đối với dữ liệu rời rạc, người ta thường cho các nhóm dưới dạng  trong đó . Nhóm  được hiểu là nhóm gồm các giá trị . Khi đó, ta cần hiệu chỉnh mẫu dữ liệu ghép nhóm trước khi thực hiện tính toán các số đặc trưng bằng cách hiệu chỉnh nhóm  thành nhóm .

### **B. VÍ DỤ**

**Ví dụ 1**. Độ bão hoà oxygen trong máu (còn được gọi là chỉ số ) biểu thị cho tỉ lệ hemoglobin có oxygen trên tổng lượng hemoglobin trong máu. Chỉ số  (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  ở một số bệnh nhân Covid-19 người ta thu được kết quả sau:

|  |  |  |  |
| --- | --- | --- | --- |
|    |    |    |    |
|  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.

**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:

|  |  |  |  |
| --- | --- | --- | --- |
|    |    |    |    |
|  Số bệnh nhân  |  12  |  31  |  7  |

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

Cỡ mẫu . Do đó, số trung bình là 

Do có  giá trị nhỏ hơn trung vị nên trung vị thuộc nhóm . Ta có, . Do đó, trung vị là 

Như vậy, chỉ số  trung bình của 50 bệnh nhân là 94,58; có 25 bệnh nhân có chỉ số  nhỏ hơn 94,76 và 25 bệnh nhân có chỉ số  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 , chọn giá trị đại diện là .

**Ví dụ 2.** Mức thưởng tết (triệu đồng) mà các công nhân một nhà máy nhận được như sau:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  Mức thưởng  |   |   |    |   |
|  Só công nhân  |  13  |  35  |  47  |  25  |

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

**Giải**

Cỡ mẫu: . Số công nhân có mức thưởng tết từ 15 đến dưới 20 triệu đồng là nhiều nhất nên nhóm chứa mốt là nhóm 

Ta có, . Do đó, mốt của mẫu số liệu là 

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

### **C. BÀI TẬP**

1. 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  |    |   |    |    |    |
|  Số cầu thủ  |  2  |  5  |  6  |  9  |  3  |

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

1. 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  |    |   |    |    |    |
|  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.

1. 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  |    |   |    |    |    |
|  Số cầu thủ  |  2  |  5  |  6  |  9  |  3  |

Tìm  sao cho có  số cầu thủ tham gia trận đấu chạy ít nhất .

1. 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  |    |    |    |    |    |
|  Só học sinh  |  23  |  8  |  5  |  3  |  1  |

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

1. 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  |    |    |    |    |    |
|  Só học sinh  |  23  |  8  |  5  |  3  |  1  |

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

1. 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  |    |    |    |    |    |
|  Só học sinh  |  23  |  8  |  5  |  3  |  1  |

Tính các tứ phân vị của mẫu số liệu ghép nhóm và cho biết ý nghĩa của các kết quả thu được.

## **BÀI TẬP CUỐI CHƯƠNG III**

### **A. TRẮC NGHIỆM**

1. Nhóm số liệu rời rạc  với  là nhóm gồm các giá trị

**A.**  và . **B.** . **C.** . **D.** .

1. Giá trị đại diện của nhóm  là

**A.** . **B.** . **C.** . **D.** .

1. Số  thoả mãn có  giá trị trong mẫu số liệu nhỏ hơn  và  giá trị trong mẫu số liệu lớn hơn  là

**A.** số trung bình. **B.** trung vị. **C.** tứ phân vị thứ nhất. **D.** tứ phân vị thứ ba.

1. Số  thoả mãn có  giá trị trong mẫu số liệu nhỏ hơn  và  giá trị trong mẫu số liệu lớn hơn  là

**A.** số trung bình. **B.** trung vị.

**C.** tứ phân vị thứ nhất. **D.** tứ phân vị thứ ba.

1. Mẫu số liệu ghép nhóm với tần số các nhóm bằng nhau có số mốt là

**A.** 0. **B.** 1. **C.** 2. **D.** 3.

1. Cho mẫu số liệu ghép nhóm về tuổi thọ (đơn vị tính là năm) của một loại bóng đèn mới như sau.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  Tuổi thọ  |    |    |   |   |
|  Số bóng đèn  |  8  |  22  |  35  |  15  |

Số trung bình của mẫu số liệu là

**A.** 5,0. **B.** 5,32. **C.** 5,75. **D.** 6,5.

1. Cho mẫu số liệu ghép nhóm về tuổi thọ (đơn vị tính là năm) của một loại bóng đèn mới như sau.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  Tuổi thọ  |    |    |   |   |
|  Số bóng đèn  |  8  |  22  |  35  |  15  |

Nhóm chứa trung vị của mẫu số liệu là

**A.** . **B.** . **C.** . **D.** .

1. Cho mẫu số liệu ghép nhóm về tuổi thọ (đơn vị tính là năm) của một loại bóng đèn mới như sau.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  Tuổi thọ  |    |    |   |   |
|  Số bóng đèn  |  8  |  22  |  35  |  15  |

Nhóm chứa tứ phân vị thứ nhất của mẫu số liệu là

**A.** . **B.** . **C.** . **D.** .

1. Cho mẫu số liệu ghép nhóm về tuổi thọ (đơn vị tính là năm) của một loại bóng đèn mới như sau.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  Tuổi thọ  |    |    |   |   |
|  Số bóng đèn  |  8  |  22  |  35  |  15  |

Nhóm chứa tứ phân vị thứ ba của mẫu số liệu là

**A.** . **B.** . **C.** . **D.** .

1. Cho mẫu số liệu ghép nhóm về tuổi thọ (đơn vị tính là năm) của một loại bóng đèn mới như sau.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  Tuổi thọ  |    |    |   |   |
|  Số bóng đèn  |  8  |  22  |  35  |  15  |

Nhóm chứa mốt của mẫu số liệu là

**A.** . **B.** . **C.** . **D.** .

1. Cho mẫu số liệu ghép nhóm về tuổi thọ (đơn vị tính là năm) của một loại bóng đèn mới như sau.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  Tuổi thọ  |    |    |   |   |
|  Số bóng đèn  |  8  |  22  |  35  |  15  |

Số mốt của mẫu số liệu ghép nhóm này là

**A.** 0. **B.** 1. **C.** 2. **D.** 3.

### **B. TỰ LUẬN**

1. 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:



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?

1. Bạn Chi vào website của một cửa hàng bán điện thoại tìm hiểu và đã thống kê số lượng một loại điện thoại theo giá bán cho kết quả như sau:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  Giá tiền (triệu đồng)  |    |   |   |  |   |
|  Số lượng  |  20  |  5  |  11  |  18  |  21  |

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

b)  loại điện thoại trên có giá dưới bao nhiêu?

1. Số nguyện vọng đăng kí vào đại học của các bạn trong lớp được thống kê trong bảng sau:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  Số nguyện vọng  |    |    |    |    |
|  Số học sinh  |  5  |  18  |  13  |  7  |

a) Trung bình một bạn trong lớp đăng kí bao nhiêu nguyện vọng.

b) Tìm các tứ phân vị của mẫu số liệu.

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