BÀI GIẢNG: MẪU SỐ LIỆU GHÉP NHÓM - TIẾT 1
CHUYÊN ĐỀ: CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CỦA MẪU SỐ LIỆU GHÉP NHÓM
MÔN: TOÁN 11
GIÁO VIÊN: NGUYỄN CÔNG CHÍNH

MỤC TIÊU

- ✓ *Đọc và giải thích mẫu số liệu ghép nhóm*
- ✓ *Ghép nhóm mẫu số liệu*

I. GIỚI THIỆU VỀ MẪU SỐ LIỆU GHÉP NHÓM

Bài toán mở đầu:

Trong kì thi tốt nghiệp Trung học phổ thông năm 2021 đợt 1 có 344752 thí sinh dự thi cả ba môn Toán, Vật lí, Hoá học (theo: vietnamnet.vn, ngày 26/07/2021). Giả sử điểm thi của các thí sinh này được cho trong bảng số liệu sau:

STT	Điểm Toán	Điểm Vật lí	Điểm Hoá học	Tổng điểm
1	8,00	7,25	7,75	23,00
2	6,00	7,75	6,50	20,25
...	...	...	...	...
344 752	4,50	5,75	6,25	16,50

Các trường đại học, cao đẳng tuyển sinh theo tổ hợp A00 quan tâm đến tổng điểm ba môn của các thí sinh này. Biểu diễn dãy số liệu về tổng điểm ba môn của các thí sinh này thế nào để các trường thấy được bức tranh tổng thể về kết quả thi?

HĐ1. Xét dữ liệu cho trong tình huống mở đầu.

a) Mẫu số liệu về tổng điểm, kí hiệu là (T), có bao nhiêu giá trị?

b) Nếu lập bảng tần số cho mẫu số liệu (T) thì có dễ hình dung được bức tranh tổng thể về kết quả thi không? Vì sao?

c) Mẫu số liệu (T) được mô tả dưới dạng bảng thống kê sau:

Tổng điểm	<6	[6,7)	[7,8)	…	[28,29)	[29,30]
Số thí sinh	23	69	192		216	12

Hãy đọc và giải thích số liệu được biểu diễn trong bảng thống kê

Cách giải:

a) (T) có tất cả 344752 số liệu

b) Rất khó để có một bức tranh tổng thể về kết quả thi vì có nhiều điểm, số điểm lẻ, đa dạng.

c) Có 23 học sinh có tổng điểm nhỏ hơn 6, có 69 học sinh có tổng điểm từ 6 đến 7,… có 12 học sinh có tổng điểm từ 29 đến 30.

Kết luận:

Mẫu số liệu ghép nhóm là mẫu số liệu cho dưới dạng bảng tần số của các nhóm số liệu. 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 [a ; b), trong đó a là đầu mút trái, b là đầu mút phải.

Nhận xét

- Mẫu số liệu ghép nhóm được dùng khi ta không thể thu thập được số liệu chính xác hoặc do yêu cầu của bài toán mà ta phải biểu diễn mẫu số liệu dưới dạng ghép nhóm để thuận lợi cho việc tổ chức, đọc và phân tích số liệu.

- Trong một số trường hợp, nhóm số liệu cuối cùng có thể lấy đầu mút bên phải.

Ví dụ 1. Mẫu số liệu sau cho biết phân bố theo độ tuổi của dân số Việt Nam năm 2019 .

Độ tuổi	Dưới 15 tuổi	Từ 15 đến dưới 65 tuổi	Từ 65 tuổi trở lên
Số người	23 371 882	65 420 451	7 416 651

(Theo: Báo cáo số liệu tổng điều tra dân số năm 2019)

a) Mẫu số liệu đã cho có là mẫu số liệu ghép nhóm hay không?

b) Nêu các nhóm và tần số tương ứng. Dân số Việt Nam năm 2019 là bao nhiêu?

Cách giải:

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

b) Có ba nhóm là: Dưới 15 tuổi, Từ 15 đến dưới 65 tuổi, Từ 65 tuổi trở lên. Có 23371882 người dưới 15 tuổi; 65420451 người từ 15 đến dưới 65 tuổi và 7416651 người từ 65 tuổi trở lên.

Dân số Việt Nam năm 2019 là 23371882 + 65420451 + 7416651 = 96208984 người.

Luyện tập 1. Cho mẫu số liệu ghép nhóm về thời gian (phút) đi từ nhà đến nơi làm việc của các nhân viên một công ty như sau:

Thời gian	[15,20)	[20,25)	[25,30)	[30,35)	[35,40)	[40,45)	[45,50)
Số nhân viên	6	14	25	37	21	13	9

Đọc và giải thích mẫu số liệu này

Cách giải:

Từ 15 phút đến 20 phút có 6 nhân viên

Từ 20 phút đến 25 phút có 14 nhân viên

….

Từ 45 phút đến 50 phút có 9 nhân viên.

II. GHÉP NHÓM MẪU SỐ LIỆU

Chỉ số BMI (đo bằng w/h^2 trong đó w là cân nặng đơn vị là kilôgam, h là chiều cao đơn vị là mét) của các học sinh trong một tổ được cho như sau:

Một người có chỉ số BMI nhỏ hơn 18,5 được xem là thiếu cân; từ 18,5 đến dưới 23 là có cân nặng lí tưởng so với chiều cao; từ 23 trở lên là thừa cân. Hãy lập mẫu số liệu ghép nhóm cho mẫu số liệu trên để biểu diễn tình trạng cân nặng so với chiều cao của các học sinh trong tổ.

Cách giải:

Chỉ số BMI	<18,5	[18,5;23)	≥23
Số học sinh	1	5	2

Kết luận:

Để chuyển mẫu số liệu không ghép nhóm sang mẫu số liệu ghép nhóm, ta làm như 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.

- Độ dài của nhóm [a ; b) là b - a.

- Không nên chia thành quá nhiều nhóm hoặc quá ít nhóm. Các nhóm không giao nhau, các nhóm nên có độ dài như nhau và tổng độ dài các nhóm lớn hơn khoảng biến thiên.

Ví dụ 2. Bảng thống kê sau cho biết thời gian chạy (phút) của 30 vận động viên (VĐV) trong một giải chạy Marathon.

Thời gian	129	130	133	134	135	136	138	141	142	143	144	145
Số VĐV	1	2	1	1	1	2	3	3	4	5	2	5

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

Cách giải:

Giá trị nhỏ nhất là 129, giá trị lớn nhất là 145 nên khoảng biến thiên là 145 - 129 = 16 . 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 [127,5 ; 130,5), [130,5 ; 133,5), ..., [142,5; 145,5). Đếm số giá trị thuộc mỗi nhóm, ta có mẫu số liệu ghép nhóm nhu sau:

Thời gian	[127,5; 130,5)	[130,5; 133,5)	[133,5; 136,5)	[136,5; 139,5)	[139,5; 142,5)	[142,5; 145,5)
Số VĐV	3	1	4	3	7	12

Luyện tập 2. Cân nặng (kg) của 35 người trưởng thành tại một khu dân cư được cho như sau:

43 51 47 62 48 40 50 62 53 56 40 48 56 53 50 42 55

52 48 46 45 54 52 50 47 44 54 55 60 63 58 55 60 58 53

Chuyển mẫu số liệu trên thành dạng ghép nhóm, các nhóm có độ dài bằng nhau, trong đó có nhóm [40 ; 45).

Cách giải:

Độ dài mỗi nhóm là 5.

GTNN là 40, GTLN là 63

Cân nặng	[40,45)	[45,50)	[50,55)	[55,60)	[60,65]
Số người	5	7	11	7	5

Vận dụng. Một công ty may quần áo đồng phục học sinh cho biết cỡ áo theo chiều cao của học sinh được tinh như sau:

Chiều cao (cm)	[150;160)	[160; 167)	[167; 170)	[170; 175)	[175; 180)
Cỡ áo	S	M	L	XL	XXL

Công ty muốn ước lượng tỉ lệ các cỡ áo khi may cho học sinh lớp 11 đã đo chiều cao của 36 học sinh nam khối 11 của một trường và thu được mẫu số liệu sau (đơn vị là centimét):

160 161 161 162 162 162 163 163 163 164 164 164 164

165 165 165 165 165 166 166 166 166 167 167 168 168

168 168 169 169 170 171 171 172 172 174

a) Lập bảng tần số ghép nhóm của mẫu số liệu với các nhóm đã cho ở bảng trên.

b) Công ty may 500 áo đồng phục cho học sinh lớp 11 thì nên may số lượng áo theo mỗi cỡ là bao nhiêu chiếc?

Cách giải:

a) Lập bảng tần số ghép nhóm

Chiều cao (cm)	[150,160)	[160,167)	[167,170)	[170,175)	[175,180)
Số học sinh	0	22	8	6	0

b) Lập bảng tần số ghép nhóm cho 500 học sinh

Tần suất $f_i = \frac{n_i}{n}.100\%$

Chiều cao (cm)	[150,160)	[160,167)	[167,170)	[170,175)	[175,180)
Số học sinh	0	306	111	83	0