**CHƯƠNG V: THỐNG KÊ**

*Thống kê là khoa học nghiên cứu các phương pháp thu thập, phân tích và xử lí các số liệu nhằm phát hiện các quy luật thống kê trong tự nhiên và xã hội. Chương trình này giúp học sinh nắm vững một số phương pháp trình bày số liệu (bằng bảng, biểu đồ) và thu gọn số liệu nhờ các số đặc trưng.*

**§ 1: BẢNG PHÂN BỐ TẦN SỐ VÀ TẦN SUẤT**

I/ Ôn tập:

1. Số liệu thống kê:

Khi thực hiện điều tra thống kê (theo mục đích đã định trước), cần xác định tập hợp các đơn vị điều tra, dấu hiệu điều tra và thu thập các số liệu.

Ví dụ 1:

Khi điều tra “Năng suất lúa hè thu năm 1998” của 31 tỉnh, người ta thu thập được các số liệu ghi trong bảng dưới đây

*Năng suất lúa hè thu (tạ/ha) năm 1998 của 31 tỉnh*

30 30 25 25 35 45 40 40 35 45

25 45 30 30 30 40 30 25 45 45

35 35 30 40 40 40 35 35 35 35 35

Tập hợp các đơn vị điều tra là tập hợp 31 tỉnh, mỗi một tỉnh là một đơn vị điều tra. Dấu hiệu điều tra là *năng suất lúa hè thu năm 1998 ở mỗi tỉnh.* Các số liệu trong bảng 1 gọi là các số liệu thống kê, còn gọi là các giá trị của dấu hiệu.

1. Tần số:

Trong 31 số liệu thống kê ở trên, ta thấy có 5 giá trị khác nhau là:

x1 = 25; x2 = 30; x3 = 35; x4 = 40; x5 = 45.

Giá trị x1 = 25 xuất hiện 4 lần, ta gọi n1 = 4 là tần số của giá trị x1/

Tương tự, n2 = 7; n3 = 9; n4 = 6; n5 = 5 lần lượt là tần số của các giá trị x2; x3; x4; x5.

II/ Tần suất:

Trong 31 số liệu thống kê ở trên, giá trị x1 có tần số là 4, do đó chiếm tỉ lệ là 4/31 hay 12,9% được gọi là tần suất của giá trị x1.

Tương tự, các giá trị x2; x3; x4; x5 lần lượt có tần suất là

7/31 = 22,6%; 9/31 = 29%; 6/31 = 19,4%; 5/31 = 16,1%.

Dựa vào các kết quả thu được, ta lập bảng sau

Năng suất lúa hè thu năm 1998 của 31 tỉnh

|  |  |  |
| --- | --- | --- |
| Năng suất lúa  (tạ/ha) | Tần số | Tần suất(%) |
| 35  30  35  40  45 | 4  7  9  6  5 | 12,9  22,6  29  19,4  16,1 |
| Cộng | 31 | 100 (%) |

***Bảng 2***

Bảng 2 phản ánh tình hình năng suất lúa của 31 tỉnh, được gọi là bảng phân bố tần số và tần suất.

Nếu trong bảng 2, bỏ cột tần số ta được bảng phân bố tần suaatsl bỏ cột tần suất ta được bảng phân bố tần số.

III/ Bảng phân bố tần số và tần suất ghép lớp:

Ví dụ 2: Đề chuẩn bị may đồng phục cho học sinh người ta đo chiều cao của 36 học sinh trong một lớp học và thu được các số liệu thống kê ghi trong bảng sau

*Chiều cao của 36 học sinh (đơn vị: cm)*

158 152 156 158 168 160 170 166 161 160 172 173

150 167 165 163 158 162 169 159 163 164 161 160

164 159 163 155 163 165 154 161 164 151 164 152

***Bảng 3***

Để xác định hợp lí số lượng quần áo cần may cho mỗi “kích cỡ” ta phân lớp các số liệu trên như sau:

Lớp 1 gồm những số đo chiều ca từ 150 cm đến dưới 156 cm, ký hiệu là [150; 156);

Lớp 2 gồm những số đo chiều ca từ 156 cm đến dưới 162 cm, ký hiệu là [156; 162);

Lớp 3 gồm những số đo chiều ca từ 162 cm đến dưới 168 cm, ký hiệu là [162; 168);

Lớp 4 gồm những số đo chiều ca từ 168 cm đến 174 cm, ký hiệu là [168; 174];

Ta thấy có 6 số liệu thuộc vào lớp 1, ta gọi n1 = 6 là tần số của lớp 1. Cũng vậy, ta gọi n2 = 12 là tần số của lớp 2, n3 = 13 là tần số của lớp 3, n4 = 5 là tần số của lớp 4.

Các tỉ số

f1 = 6/36 = 16,7%; f2 = 12/36 = 33,3%

f3 = 13/36 = 36,1%; f4 = 5/36 = 13,9%

được gọi là tần suất của các lớp tương ứng.

Các kết quả trên được trình bày gọn trong bảng đưới đây:

*Chiều cao của 36 học sinh*

|  |  |  |
| --- | --- | --- |
| Lớp số đo chiều cao  (cm) | Tần số | Tần suất (%) |
| [150 ; 156)  [156 ; 162)  [162 ; 168)  [168 ; 174] | 6  12  13  5 | 16,7  33,3  36,1  13,9 |
| Cộng | 36 | 100 (%) |

***Bảng 4***

*Bảng 4 được gọi là bảng phân bố tần số và tần suất ghép lớp. Nếu trong bảng 4 bỏ cột tần số thì sẽ có bảng phân bố tần suất ghép lớp, bỏ cột tần suất thì sẽ có bảng phân bố tần số ghép lớp.*

Bảng 4 ở trên cho ta cơ sở để xác định số lượng quần áo cần may của mỗi cỡ (tương ứng mỗi lớp). Chẳng hạn, vì số học sinh có chiều cao thuộc lớp thứ nhất chiếm 16,7% tổng số học sinhm nên số quần áo cần may thuộc cỡ tương ứng với lớp đó chiếm 16,7% số lượng quần áo cần may. Ta cũng có kết luận tương tự đốiv ới các lớp khác.

Nếu lớp học kể trên đại diện được cho toàn trường thì có thể áo dụng kết quả đó để may quần áo cho học sinh cả trường.

🕮Cho các số liệu thống kê ghi trong bảng sau

Tiền lãi (nghìn đồng) của mỗi ngày trong 30 ngày được khảo sát ở một quầy bán báo

37 74 65 31 63 58 82 67 77 63 46 30 53 73 51 44 52 92 93 53 85 77 47 42 57 57 85 55 64

***Bảng 5***

Hãy lập bảng phân bố tần suất ghép lớp với các lớp như sau

[29,5 ; 40,5), [40,5 ; 51,5), [51,5 ; 62,5), [62,5 ; 73,5), [73,5 ; 84,5), [84,5 ; 95,5]

**BÀI TẬP**

1. Cho các số liệu thống kê ghi trong bảng sau

*Tuổi thọ của 30 bóng đèn điện được thắp thử (đơn vị: giờ)*

1180 1150 1190 1170 1180 1170

1160 1170 1160 1150 1190 1180

1170 1170 1170 1190 1170 1170

1170 1180 1170 1160 1160 1160

1170 1160 1180 1180 1150 1170

1. Lập bảng phân bố tần số và bảng phân bố tần suất.
2. Dựa vào kết quả của câu a), hãy đưa ra nhận xét về tuổi thọ của các bóng đèn nói trên.
3. Cho bảng phân bố tần số ghép lớp sau

*Độ dài của 60 lá dương xỉ trưởng thành*

|  |  |
| --- | --- |
| Lớp của độ dài (cm) | Tần số |
| [10 ; 20)  [20 ; 30)  [30 ; 40)  [40 ; 50] | 8  18  24  10 |
| Cộng | 60 |

1. Lập bảng phân bố tần suất ghép lớp.
2. Dựa vào kết quả của câu a), hãy nêu rõ trong 60 lá dương xỉ được khảo sát:

Số lá có độ dài dưới 30 cm chiếm bao nhiêu phần trăm?

Số lá có độ dài từ 30 cm đến 50 cm chiếm bao nhiêu phần trăm?

1. Cho các số liệu thống kê ghi trong bảng sau

*Khối lượng của 30 củ khoai tây thu hoạch được ở nông trường T (đơn vị : g)*

90 73 88 99 100 102 111 96 79 93

81 94 96 93 95 82 90 106 103 116

109 108 112 87 74 91 84 97 85 92

Lập bảng phân bố tần số và tần suất ghép lớp, với các lớp sau

[70 ; 80), [80 ; 90), [90 ; 100), [100 ; 110), [110 ; 120]

1. Cho các số liệu thống kê ghi trong bảng sau

*Chiều cao của 35 cây bạch đàn (đơn vị : m)*

6,6 7,5 8,2 8,2 7,8 7,9 9,0 8,9 8,2

7,2 7,5 8,3 7,4 8,7 7,7 7,0 9,4 8,7

8,0 7,7 7,8 8,3 8,6 8,1 8,1 9,5 6,9

8,0 7,6 7,9 7,3 8,5 8,4 8,0 8,8

1. Lập bảng phân bố tần suất ghép lớp, với các lớp sau

[6,5 ; 7,0) ; [7,0 ; 7,5) ; [7,5 ; 8,0) ; [8,0 ; 8,5) ; [8,5 ; 9,0) ; [9,0 ; 9,5].

1. Dựa vào kết quả của câu a), hãy nêu nhận xét về chiều cao của 35 cây bạch đàn nói trên.

**§ 2: BIỂU ĐỒ**

I/ Biểu đồ tần suất hình cột và đường gấp khúc tần suất

Ta có thể mô tả một cách trực quan các bảng phân bố tần suất (hoặc tần số), bảng phân bố tần suất (hoặc tần số) ghép lớp bằng biểu đồ hoặc đường gấp khúc.

1. Biểu đồ tần suất hình cột

Ví dụ 1: Để mô tả bảng phân bố tần suất ghép lớp (bảng 4) trong § 1, có thể vẽ biểu đồ tần suất hình cột sau (h .34)

*Hình 34. Biểu đồ tần suất hình cột về chiều cao (cm) của 36 học sinh*

1. Đường gấp khúc tần suất:

Bảng phân bố tần suất ghép lớp kể trên (bảng 4) cũng có thể được mô tả bằng một đường gấp khúc, vẽ như sau.

Trên mặt phẳng tọa độ, xác định các điểm (ci; fi), I = 1,2,3,4 trong đó ci là trung bình cộng hai mút của lớp i (ta gọi ci là giá trị đại diện của lớp i).

Vẽ các đoạn thẳng nối điểm (ci;fi) với điểm (ci+1; fi+1), I =1,2,3 ta thu được một đường gấp khúc, gọi là đường gấp khúc tần suất (h.35).

*Hình 35. Đường gấp khúc tần suất về chiều cao (cm) của 36 học sinh.*

🕮**1**

Cho bảng phân bố tần suất ghép lớp sau

Nhiệt độ trung bình của tháng 12 tại thành phố Vinh từ 1961 đến 1990 (30 năm).

|  |  |
| --- | --- |
| Lớp nhiệt độ (độ C) | Tần suất (%) |
| [15 ; 17)  [17 ; 19)  [19 ; 21)  [21 ; 23) | 16,7  43,3  36,7  3,3 |
| Cộng | 100 (%) |

***Bảng 6***

Hãy mô tả bảng 6 bằng cách vẽ biểu dồ tần suất hình cột và đường gấp khúc tần suất.

1. CHÚ Ý:

*Ta cũng có thể mô tả bảng phân bố tần số ghép lớp bằng biểu đồ tần số hình cột hoặc đường gấp khúc tần số. Cách vẽ cũng như cách vẽ biểu đồ tần suất hình cột hoặc đường gấp khúc tần suất, trong đó thay trục tần suất bởi trục tần số.*

II/ Biểu đồ hình quạt:

Người ta còn dùng biểu đồ hình quạt để mô tả bảng cơ cấu trong ví dụ dưới đây

Ví dụ 2: Cho bảng 7

*Cơ cấu giá trị sản xuất công nghiệp trong nước năm 1997, phân theo thành phần kinh tế.*

|  |  |
| --- | --- |
| Các thành phần kinh tế | Số phần trăm |
| 1. Khu vực doanh nghiệp nhà nước 2. Khu vực ngoài quốc doanh 3. Khu vực đầu tư nước ngoài | 23,7  47,3  29,0 |
| Cộng | 100 (%) |

***Bảng 7***

Hình 36a dưới đây là biểu đồ hình quạt mô tả bảng 7.

a)

b)

CHÚ Ý:

*Các bảng phân bố tần suất ghép lớp cũng có thể mô tả bằng biểu đồ hình quạt, chẳng hạn hình 36b mô tả bảng 6.*

🕮**2**

Dựa vào biểu đồ hình quạt cho ở hình 37 đưới đây, hãy lập bảng cơ cấu như trong ví dụ 2.

*Hình 37. Biểu đồ hình quạt về cơ cấu giá trị sản xuất công nghiệp trong nước năm 1999, phân theo thành phần kinh tế (%)*

**BÀI TẬP**

1. Hãy mô tả bảng phân bố tần suất ghép lớp đã được lập ở bài tập số 2 của § 1 bằng cách vẽ biểu đồ tần suất hình cột và đường gấp khúc tần suất.
2. Xét bảng phân bố tần số và tần suất ghép lớp đã được lập ở bài tập số 3 của § 1.
3. Hãy vẽ biểu đồ tần suất hình cột, đường gấp khúc tần suất.
4. Hãy vẽ biểu đồ tần số hình cột, đường gấp khúc tần số.
5. Dựa vào biểu đồ tần suất hình cột đã vẽ ở câu a), hãy nêu nhận xét về khối lượng của 30 củ khoai tây được khảo sát.
6. Dựa vào biểu đồ hình quạt dưới đây (h.38), hãy lập bảng cơ cấu như trong ví dụ 2.

*Hình 38. Biểu đồ hình quạt về cơ cấu giá trị sản xuất công nghiệp trong nước năm 2000, phân theo thành phần kinh tế (%).*

**§ 3: SỐ TRUNG BÌNH CỘNG. SỐ TRUNG VỊ. MỐT**

*Để thu được các thông tin quan trọng từ các số liệu thống kê, người ta sử dụng những số đặc trưng như số trung bình cộng, số trung vị, mốt, phương sai, độ lệch chuẩn. Các số đặc trưng này phản ánh những khía cạnh khác nhau của dấu hiệu điều tra.*

I/ Số trung bình cộng (hay số trung bình)

**Ví dụ 1:**

1. Áp dụng công thức tính số trung bình cộng đã học ở lớp 7, ta tính được chiều cao trung bình của 36 học sinh trong kết quả điều tra được trình bày ở bảng 3 của § 1 là 161 cm.
2. Sử dụng bảng phân bố tần số và tần suất ghép lớp, ta tính gần đúng chiều cao trung bình của 36 học sinh trong kết quả điều tra được trình bày ở bảng 4 của § 1 theo 2 cách sau

*Cách 1:* ***Sử dụng bảng phân bố tần số ghép lớp***

*Nhân giá trị đại diện của mỗi lớp với tần số của lớp đó, cộng các kết quả lại rồi chia cho 36, ta được*

Kết quả này có nghĩa là chiều cao trung bình của 36 học sinh kể trên là 161 cm.

Ta cũng nói 162 cm là *số trung bình cộng của bảng 4.*

*Cách 2:* ***Sử dụng bảng phân bố tần suất ghép lớp***

*Nhân giá trị đại diện của mỗi lớp với tần suất của lớp đó rồi cộng các kết quả lại ta cũng được.*

*Vậy ta có thể tính số trung bình cộng cảu các số liệu thống kê theo các công thức sau đây.*

***Trường hợp bảng phân bố tần số, tần suất***

=

*Trong đó ni, fi lần lượt là tần số, tần suất của giá trị xi, n là số các số liệu thống kê (n1 + n2 + …+ nk = n).*

***Trường hợp bảng phân bố tần số, tần suất ghép lớp***

=

*Trong đó ci, ni, fi lần lượt là giá trị đại diện, tần số, tần suất của lớp thứ i, n là số các số liệu thống kê (n1+n2+…+nk = n).*

🕮**1**

Cho bảng phân bố tần số và tần suất ghép lớp sau

*Nhiệt độ trung bình của tháng 2 tại thành phố Vinh từ 1961 đến hết 1960 (30 năm).*

|  |  |  |
| --- | --- | --- |
| Lớp nhiệt độ () | Tần số | Tần suất (%) |
| [12 ; 14)  [14 ; 16)  [16 ; 18)  [18 ; 20)  [20 ; 22] | 1  3  12  9  5 | 3,33  10  40  30  16,67 |
| Cộng | 30 | 100 (%) |

***Bảng 8***

1. Hãy tính trung bình cộng của bảng 6 và bảng 8.
2. Từ kết quả đã tính được ở câu a), có nhận xét gì về nhiệt độ ở thành phố Vinh trong tháng 2 và tháng 12 (của 30 năm được khảo sát).

II/ Số trung vị:

**Ví dụ 2**: Điểm thi toán cuối năm của một nhóm 9 học sinh lớp 6 là:

1; 1; 3; 6; 7; 8; 8; 9 ; 10.

Điểm trung bình của nhóm là 5,9.

Ta thấy hầu hết học sinh (6 em) trong nhóm có số điểm vượt điểm trung bình và có những điểm vượt rất xa. Như vậy, điểm trung bình không đại diện được cho trình độ học lực của các em trong nhóm.

Khi các số liệu thống kê có sự chênh lệch lớn thì số trung bình cộng không đại diện được cho các số liệu đó. Khi đó ta chọn số đặc trưng khác đại diện thích hợp hơn, đó là số trung vị.

*Sắp thứ tự các số liệu thống kê thành dãy không giảm (hoặc không tăng).* ***Số trung vị*** *(của các số liệu thống kê đã cho) kí hiệu Me là số đứng giữa dãy nếu số phần tử là lẻ và là trung bình cộng của hai số đứng giữa dãy nếu số phần tử là chẵn.*

Trong ví dụ 2 ta có Me = 7.

**Ví dụ 3 :** Điểm thi Toán của bốn học sinh lớp 6 được xếp thành dãy không giảm là 1 ; 2,5 ; 8 ; 9,5.

Trong dãy này có 2 số đứng giữa là 2,5 và 8.

Khi đó, ta chọn số trung vị là trung bình cộng của hai số này

Me = .

🕮**2**

Trong bảng phân bố tần số, các số liệu thống kê đã được sắp thứ tự thành dãy không giảm theo các giá trị của chúng.

Hãy tìm số trung vị của các số liệu thống kê cho ở bảng 9.

*Số áo bán được trong một quý ở một cửa hang bán áo sơ mi nam*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cỡ áo | 36 | 37 | 38 | 39 | 40 | 41 | 42 | Cộng |
| Tần số  (Số áo bán được) | 13 | 45 | 126 | 110 | 126 | 40 | 5 | 465 |

***Bảng 9***

III/ Mốt:

Ở lớp 7 ta đã biết

*Mốt của một bảng phân bố tần số là giá trị tần số lớn nhất và được ký hiệu là Mo.*

Nếu trong bảng phân bố tần số có hai giá trị tần số bằng nhau và lớn hơn tần số của các giá trị khác thì chọn mốt là giá trị nào? Ta xét bảng 9 ở trên.

Trong bảng 9, có hai giá trị là 38 và 40 cùng có tần số lớn nhất là 126, trong trường hợp này ta coi rằng có hai mốt là

= 38, = 40.

Kết quả vừa thu được cho thấy rằng trong kinh doanh, cửa hang nên ưu tiên nhập hai cỡ áo số 38 và 40 nhiều hơn.

**BÀI TẬP**

1. Tính trung bình cộng của các bảng phân bố đã được lập ở bài tập 1 và bài tập 2 của § 1.
2. Trong một trường THPT, đề tìm hiểu tình hình học môn Toán của hai lớp 10A và 10B, người ta cho hai lớp thi Toán theo cùng một đề tih và lập được hai bảng phân bố tần số ghép lớp sau đây

*Điểm thi toán của lớp 10A*

|  |  |
| --- | --- |
| Lớp điểm thi | Tần số |
| [0 ; 2)  [2 ; 4)  [4 ; 6)  [6 ; 8)  [8 ; 10] | 2  4  12  28  4 |
| Cộng | 50 |

*Điểm thi toán của lớp 10B*

|  |  |
| --- | --- |
| Lớp điểm thi | Tần số |
| [0 ; 2)  [2 ; 4)  [4 ; 6)  [6 ; 8)  [8 ; 10] | 4  10  18  14  5 |
| Cộng | 51 |

Tính các số trung bình cộng của hai bảng phân bố ở trên và nêu nhận xét về kết quả làm bài thi của hai lớp.

1. Điều tra tiền lương hang tháng của 30 công nhân của một xưởng may, ta có bảng phân bố tần số sau

*Tiền lương của 30 công nhân xưởng may*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Tiền lương (nghìn đồng) | 300 | 500 | 700 | 800 | 900 | 1000 | Cộng |
| Tần số | 3 | 5 | 6 | 5 | 6 | 5 | 30 |

Tìm mốt của bảng phân bố trên. Nêu ý nghĩa của kết quả tìm được.

1. Tiền lương hang tháng của 7 nhân viên trong một công ti du lịch là: 650, 840, 690, 720, 2500, 670, 3000 (đơn vị : nghìn đồng).

Tìm số trung vị của các số liệu thống kê đã cho. Nêu ý nghĩa của kết quả đã tìm được.

1. Cho biết tình hình thu hoạch lúa vụ mùa năm 1980 của ba hợp tác xã ở địa phương V như sau

|  |  |  |
| --- | --- | --- |
| Hợp tác xã | Năng suất lúa (tạ/ha) | Diện tích trồng lúa (ha) |
| A  B  C | 40  38  36 | 150  130  120 |

Hãy tính năng suất lúa trung bình của vụ mùa năm 1980 trong toàn bộ ba hợp tác xã kể trên.

**§ 4 PHƯƠNG SAI VÀ ĐỘ LỆCH CHUẨN**

I/Phương sai:

**Ví dụ 1:** Cho biết giá trị thành phầm quy ra tiền (nghìn đồng) trong một tuần lao động của 7 công nhân ở tổ 1 là

180, 190, 190, 200, 210, 210, 220 (1)

Còn của 7 công nhân ở tổ 2 là

150, 170, 170, 200, 230, 230, 250. (2)

Ta thấy số trung bình cộng của dãy (1) và số trung bình cộng của dãy (2) bằng nhau

Tuy nhiên, khi so sánh dãy (1) và dãy (2) ta thấy các số liệu ở dãy (1) gần với số trung bình cộng hơn nên chúng đồng đều hơn. Khi đó ta nói các số liệu thống kê ở dãy (1) ít phân tán hơn dãy (2).

*Để tìm số* ***đo độ phân tán*** *(so với số trung bình cộng) của dãy (1) ta tính các độ lệch của mỗi số liệu thống kê đới với số trung bình cộng*

(180 - 200) ; (190 – 200) ; (190 – 200) ; (200 - 200) ; (210 – 200) ; (210 – 200) ; (220 – 200).

*Bình phuơng các độ lệch và tính trung bình cộng cảu chúng, ta được*

Số được gọi là phương sai của dãy (1).

Tương tự phương sai của dãy (2) là

Ta thấy phương sai của dãy (1) nhỏ hơn phương sai của dãy (2). Điều đó biểu thị độ phân tán cảu các số liệu thống kê ở dãy (1) ít hơn ở dãy (2).

**Ví dụ 2:** Tính phương sai của các số liệu thống kê cho ở bảng 4, § 1 (cũng gọi là phương sai của bảng 4).

Số trung bình cộng của bàng 4 là

Mỗi số liệu thống kê thuộc một lớp được thay thế bở giá trị đại diện của lớp đó.

1. Phương sai của bảng 4 (bảng phân bố tần số và tần suất ghép lớp) được tính như sau

Hệ thức (3) biểu thị cách tính gần đúng phương sai của bảng 4 theo tần số.

1. Từ (3) ta có

Hay

*Hệ thức (4) biểu thị cách tính gần đúng phương sai của bảng 4 theo tần suất.*

CHÚ Ý:

1. Khi hay dãy số liệu thống kê có cùng đơn vị đo và có số trung bình cộng bằng nhau hoặc xấp xỉ nhau, nếu phương sai càng nhỏ thì mức độ phân tán (so với số trung bình cộng) của các số liệu thống kê càng bé.
2. Có thể tính phương sai theo các công thức sau đây

**Trường hợp bảng phân bố tần số, tần suất**

Trong đó ni, fi lần lượt là tần số, tần suất của giá trị xi ; n là số các số liệu thống kê (n = n1+ n2 +…+nk) ; là số trung bình cộng của các số liệu đã cho.

**Trường hợp bảng phân bố tần số, tần suất ghép lớp**

Trong đó ci, ni, fi lần lượt là giá trị đại diện, tần số, tần suất của lớp thứ i ; n là số các số liệu thống kê (n = n1+n2+…+nk) ; là số trung bình cộng của các số liệu thống kê đã cho.

Ngoài ra, người ta còn chứng minh được công thức sau

Trong đó là trung bình cộng của các bình phương số liệu thống kê, tức là

(Đối với bảng phân bố tần số, tần suất),

(Đối với bảng phân bố tần số, tần suất ghép lớp)

🕮**1**

Hãy tính phương sai của bảng 6 (ở § 2).

II/ Độ lệch chuẩn:

Trong ví dụ 2 ở trên, ta đã tính được phương sai của bảng 4 (ở § 1) bằng Nếu để ý đến đơn vị đo thì ta thấy đơn vị đo của là (bình phương đơn vị đo của dấu hiệu được nghiên cứu). Muốn tránh điều này, có thể dung căn bậc 2 của phương sai gọi là độ lệch chuẩn (của bảng 4) và ký hiệu là s. Vậy

s =

*Phương sai và độ lệch chuẩn s đều được dung để đánh giá mức độ phân tán của các số liệu thống kê (so với số trung bình cộng). Nhưng khi cần chú ý đến đơn vị đo thì ta dung s, vì s có cùng đơn vị đo với dấu hiệu được nghiên cứu.*

🕮**2**

Hãy tính độ lệch chuẩn của bảng 6 (ở § 2).

**BÀI TẬP**

1. Tính phương sai và độ lệch chuẩn của bảng phân bố tần số đã được lập ở bài tập 1 và của bảng phân bố tần số ghép lớp cho ở bài tập 2 của § 1.
2. Hai lớp 10C, 10D của một trường Trung học phổ thong đồng thời làm bài thi môn Ngữ văn theo cùng một đề thi. Kết quả thi được trình bày ở hai bảng phân bố tần số sau đây

*Điểm thi Ngữ văn của lớp 10C*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Điểm thi | 5 | 6 | 7 | 8 | 9 | 10 | Cộng |
| Tần số | 3 | 7 | 12 | 14 | 3 | 1 | 40 |

*Điểm thi Ngữ văn của lớp 10D*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Điểm thi | 6 | 7 | 8 | 9 | Cộng |
| Tần số | 8 | 18 | 10 | 4 | 40 |

1. Tính các số trung bình cộng, phương sai, độ lệch chuẩn của các bảng phân bố tần số đã cho.
2. Xét xem kết quả làm bài thi của môn Ngữ văn ở lớp nào là đồng đều hơn?
3. Cho hai bảng phân bố tần số ghép lớp

*Khối lượng của nhóm cá mè thứ 1*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Lớp khối lượng (kg) | [0,6;0,8) | [0,8;1,0) | [1,0;1,2) | [1,2;1,4] | Cộng |
| Tần số | 4 | 6 | 6 | 4 | 20 |

*Khối lượng của nhóm cá mè thứ 2*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Lớp khối lượng (kg) | [0,5;0,7) | [0,7;0,9) | [0,9;1,1) | [1,1;1,3) | [1,3;1,5] | Cộng |
| Tần số | 3 | 4 | 6 | 4 | 3 | 20 |

1. Tính các số trung bình cộng của các bảng phân bố tần số ghép lớp đã cho.
2. Tính phương sai của các bảng phân bố tần số ghép lớp đã cho.
3. Xét xem nhóm cá nào có khối lượng đồng đều hơn?

**ÔN TẬP CHƯƠNG V**

1. Chỉ rõ các bước để
2. Lập bảng phân bố tần suất ghép lớp;
3. Lập bảng phân bố tần số ghép lớp;
4. Nêu rõ cách tính số trung bình cộng, số trung vị, mốt, phương sai và độ lệch chuẩn.
5. Kết quả điều tra 59 hộ gia đình ở một vùng dân cư về số con của mỗi hộ gia đình được ghi trong bảng sau

3 2 1 1 1 1 0 2 4 0 3 0

1 3 0 2 2 2 1 3 2 2 3 3

2 2 4 3 2 2 4 3 2 4 1 3

2 1 2 0 4 2 3 1 1 2 0

1. Lập bảng phân bố tần số và tần suất;
2. Nêu nhận xét về số con của 59 gia đình đã được điều tra;
3. Tính số trung bình cộng, số trung vị, mốt của các số liệu thống kê đã cho.
4. Cho các số liệu thống kê được ghi trong hai bảng sau đây

*Khối lượng (tính theo gam) của nhóm cá thứ 1*

645 650 645 644 650 635 650 654

650 650 650 643 650 630 647 650

645 650 645 642 652 635 647 652

*Khối lượng (tính theo gam) của nhóm cá thứ 2*

640 650 645 650 643 645 650 650 642

640 650 645 650 641 650 650 649 645

640 645 650 650 644 650 650 645 640

1. Lập bảng phân bố tần số và tần suất ghép lớp theo nhóm cá thứ 1 với các lớp là

[630 ; 635) ; [635 ; 640) ; [640 ; 645) ; [645 ; 650) ; [650 ; 655]

1. Lập bảng phân bố tần số và tần suất ghép lớp theo nhóm cá thứ 2 với các lớp là

[638 ; 642) ; [642 ; 646) ; [646 ; 650) ; [650 ; 654);

1. Mô tả bảng phân bố tần suất ghép lớp đã được lập ở câu a) bằng cách vẽ biểu đồ tần suất hình cột và đường gấp khúc tần suất;
2. Mô tả bảng phân bố tần suất ghép lớp đã được lập ở câu b) bằng cách vẽ biểu đồ tần suất hình cột và đường gấp khúc tần suất;
3. Tính số trung bình cộng, phương sai và độ lệch chuẩn của các bảng phân bố tần số và tần suất ghép lớp đã lập được.

Từ đó, xét xem nhóm cá nào có khối lượng đồng đều hơn.

1. Cho các số liệu thống kê được ghi trong bảng sau

Mức lương hàng năm của các cán bộ và nhân viên trong một công ti (đơn vi : nghìn đồng)

|  |  |  |  |
| --- | --- | --- | --- |
| 20910 | 76000 | 20350 | 20060 |
| 21410 | 20110 | 21410 | 21360 |
| 20350 | 21130 | 20960 | 125000 |

Tìm mức lương bình quân của các cán bộ và nhân viên trong công ti, số trung vị của các số liệu thống kê đã cho.

Nêu ý nghĩa của số trung vị.

1. Người ta đã tiến hành thăm dò ý kiến của khách hàng về các mẫu 1,2,3,4,5 của một loại sản phẩm mới được sản xuất ở một nhà máy. Dưới đây là bảng phân bố tần số theo số phiếu tín nhiệm dành cho các mẫu kể trên.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Mẫu | 1 | 2 | 3 | 4 | 5 | Cộng |
| Tần số | 2100 | 1860 | 1950 | 2000 | 2090 | 10000 |

1. Tìm mốt của bảng phân bố tần số đã cho.
2. Trong sản xuất, nhà máy nên ưu tiên cho mẫu nào?

**BÀI TẬP TRẮC NGHIỆM**

Chọn phương án đúng trong các bài tập sau

1. Cho bảng phân bố tần số

*Tiền thưởng (triệu đồng) cho cán bộ và nhân viên trong một công ti.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tiền thưởng | 2 | 3 | 4 | 5 | 6 | Cộng |
| Tần số | 5 | 15 | 10 | 6 | 7 | 43 |

Mốt của bảng phân bố tần số đã cho là

(A)2 triệu đồng; (B) 6 triệu đồng;

(C)3 triệu đồng; (D)5 triệu đồng.

8. Cho bảng phân bố tần số

*Tuổi của 169 đoàn viên thanh niên*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tuổi | 18 | 19 | 20 | 21 | 22 | Cộng |
| Tần số | 10 | 50 | 70 | 29 | 10 | 169 |

Số trung vị của bang phân bố tần số đã cho là

(A)18 tuổi; (B)20 tuổi;

(C)19 tuổi; (D)21 tuổi.

9. Cho dãy số liệu thống kê: 21,23,25,25,22,20.

Số trung bình cộng của các số liệu thống kê đã cho là

(A)23,5; (B)22;

(C)22,5; (D)14.

10. Cho dãy số liệu thống kê: 1,2,3,4,5,6,7.

Phương sai các số liệu thống kê đã cho là

(A)1; (B)2;

(C)3; (D)4.

11. Ba nhóm học sinh gồm 10 người, 15 người, 25 người. Khối lượng trung bình của mỗi nhóm lần lượt là: 50kg , 38kg, 40kg. Khối lượng trung bình của cả ba nhóm học sinh là

(A)41,4kg; (B)42,4kg;

(C)26kg; (D)37kg.

**ĐÁP ÁN**

§ 1.

2.b) 43,3% ; 56,7%

§ 3.

1. 1170 giờ ; 31 cm.
2. 6,1 điểm ; 5,2 điểm

Điểm trung bình cộng của lớp 10A cao hơn, nên có thể nói học sinh lớp 10A có kết quả làm bài thi cao hơn.

1. Có hai mốt là
2. Số trung vị Me = 720 nghìn đồng.
3. 38,15 tạ/ha

§ 4.

1. a) Dãy số liệu về điểm thi của lớp 10C có

Dãy số liệu về điểm thi của lớp 10D có

b)Điểm số của các bài thi ở lớp 10D đồng đều hơn.

3. a)Nhóm cá thứ 1 có

Nhóm cá thứ 2 có

b)Nhóm cá thứ 1 có =0,042;

Nhóm cá thứ 2 có =0,064;

c)Nhóm cá thứ 1 có khối lượng đồng đều hơn.

**Ôn tập chương V**

1. c)
2. e)Nhóm cá thứ 1 có

Nhóm cá thứ 2 có

Nhóm cá thứ 2 có khối lượng đồng đều hơn.

1. a)Mốt là mẫu 1.
2. (C); 8.(B); 9.(C); 10.(D); 11.(A).