



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

a) Áp dụng công thức tính số trung bình cộng đã học ở lớp 7, ta tính được chiều cao trung bình \bar{x} của 36 học sinh trong kết quả điều tra được trình bày ở bảng 3 của §1 là $\bar{x} \approx 161$ cm.

b) Sử dụng bảng phân bố tần số và tần suất ghép lớp, ta tính gần đúng chiều cao trung bình \bar{x} của 36 học sinh trong kết quả điều tra được trình bày ở bảng 4 của §1 theo hai 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

$$\frac{6 \times 153 + 12 \times 159 + 13 \times 165 + 5 \times 171}{36} \approx 162 \text{ (cm).}$$

Kết quả này có nghĩa là chiều cao trung bình của 36 học sinh kể trên là $\bar{x} \approx 162$ 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

$$\bar{x} \approx \frac{16,7}{100} \times 153 + \frac{33,3}{100} \times 159 + \frac{36,1}{100} \times 165 + \frac{13,9}{100} \times 171 \approx 162 \text{ (cm).}$$

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

Trường hợp bảng phân bố tần số, tần suất

$$\bar{x} = \frac{1}{n}(n_1x_1 + n_2x_2 + \dots + n_kx_k) = f_1x_1 + f_2x_2 + \dots + f_kx_k$$

trong đó n_i, f_i lần lượt là tần số, tần suất của giá trị x_i , n là số các số liệu thống kê ($n_1 + n_2 + \dots + n_k = n$).

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

$$\bar{x} = \frac{1}{n}(n_1c_1 + n_2c_2 + \dots + n_kc_k) = f_1c_1 + f_2c_2 + \dots + f_kc_k$$

trong đó c_i, n_i, f_i lần lượt là giá trị đại diện, tần số, tần suất của lớp thứ i , n là số các số liệu thống kê ($n_1 + n_2 + \dots + n_k = n$).



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 1990 (30 năm).

Lớp nhiệt độ (°C)	Tần số	Tần suất (%)
[12 ; 14)	1	3,33
[14 ; 16)	3	10,00
[16 ; 18)	12	40,00
[18 ; 20)	9	30,00
[20 ; 22]	5	16,67
Cộng	30	100%

Bảng 8

- a) Hãy tính số trung bình cộng của bảng 6 và bảng 8.
 b) 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 cả nhóm là $\bar{x} \approx 5,9$.

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 \bar{x} 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 M_e là số đứng giữa dãy nếu số phân tử là lẻ và là trung bình cộng của hai số đứng giữa dãy nếu số phân tử là chẵn.

Trong ví dụ 2 ta có $M_e = 7$.

Ví dụ 3. Điểm thi Toán của bốn học sinh lớp 6 được xếp thành dãy không giảm là 1 ; $\boxed{2,5 ; 8}$; 9,5.

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

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

$$M_e = \frac{2,5 + 8}{2} = 5,25.$$



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 hàng 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 có tần số lớn nhất và được kí hiệu là M_0 .

Nếu trong bảng phân bố tần số có hai giá trị có 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

$$M_O^{(1)} = 38, \quad M_O^{(2)} = 40.$$

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

Bài tập

1. Tính số trung bình cộng của các bảng phân bố đã được lập ở bài tập số 1 và bài tập số 2 của §1.
2. Trong một trường THPT, để tìm hiểu tình hình học môn Toán của hai lớp 10A và 10B, người ta cho hai lớp thi Toán theo cùng một đề thi 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
[2 ; 4)	4
[4 ; 6)	12
[6 ; 8)	28
[8 ; 10]	4
Cộng	50

Điểm thi Toán của lớp 10B

Lớp điểm thi	Tần số
[0 ; 2)	4
[2 ; 4)	10
[4 ; 6)	18
[6 ; 8)	14
[8 ; 10]	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.

3. Điều tra tiền lương hàng 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 đã tìm được.

4. Tiền lương hàng 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.

5. 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	40	150
B	38	130
C	36	120

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