

Bài 2. CÁC SỐ ĐẶC TRƯNG ĐO XU THẾ TRUNG TÂM CHO MẪU SỐ LIỆU KHÔNG GHEP NHOM

|FanPage: Nguyễn Bảo Vương

PHẦN A. LÝ THUYẾT

I. Số trung bình cộng (số trung bình)

1. Định nghĩa

Số trung bình cộng của một mẫu n số liệu thống kê bằng tổng của các số liệu chia cho số các số liệu đó. Số trung bình cộng của mẫu số liệu $x_1, x_2, \dots, x_n$ bằng

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}.$$

Ví dụ 1. Kết quả 4 lần kiểm tra môn Toán của bạn Hoa là $7; 9; 8; 9$. Tính số trung bình cộng $\bar{x}$ của mẫu số liệu trên.

Giải

$$\bar{x} = \frac{7+9+8+9}{4} = \frac{33}{4} = 8,25$$

Số trung bình cộng của mẫu số liệu trên là

Nhận xét: Công thức tính số trung bình cộng $\bar{x}$ khi có các số liệu thống kê bằng nhau có thể viết lại ở dạng:

$$\bar{x} = \frac{7+8+2 \cdot 9}{1+1+2} = \frac{33}{4} = 8,25.$$

Ta có thể tính số trung bình cộng theo các công thức sau:

- Số trung bình cộng của mẫu số liệu thống kê trong bảng phân bố tần số là:

$$\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_k x_k}{n_1 + n_2 + \dots + n_k}.$$

Giá trị	x_1	x_2	$\dots$	x_k
Tần số	n_1	n_2	$\dots$	n_k

- Số trung bình cộng của mẫu số liệu thống kê trong bảng phân bố tần số tương đối là:

$$\bar{x} = f_1 x_1 + f_2 x_2 + \dots + f_k x_k,$$

Giá trị	x_1	x_2	$\dots$	x_k
Tần số tương đối	f_1	f_2	$\dots$	f_k

trong đó $f_1 = \frac{n_1}{n}, f_2 = \frac{n_2}{n}, \dots, f_k = \frac{n_k}{n}$, với $n = n_1 + n_2 + \dots + n_k$.

2. Ý nghĩa

Trong thực tiễn, để tìm hiểu một đối tượng thống kê ta đưa ra tiêu chí thống kê và tiến hành thu thập nhiều lần số liệu thống kê theo tiêu chí đó, tạo thành mẫu số liệu. Căn cứ vào mẫu số liệu đó, ta rút ra những kết luận có ích về đối tượng thống kê. Để kết luận rút ra phản ánh đúng đắn bản chất của đối tượng, ta cần nhận biết được hình thái và xu thế thay đổi của mẫu số liệu. Với cách nhìn nhận như thế, số trung bình cộng của mẫu số liệu có ý nghĩa sau:

Khi các số liệu trong mẫu ít sai lệch với số trung bình cộng, ta có thể giải quyết được vấn đề trên bằng cách lấy số trung bình cộng làm đại diện cho mẫu số liệu.

Chẳng hạn, để dự báo lượng mưa trong tháng 8 tại Hà Nội người ta tiến hành đo lượng mưa của từng ngày trong tháng 8 tại Hà Nội, ta được mẫu số liệu gồm 31 số liệu. Số trung bình cộng của mẫu số liệu đó được xem như lượng mưa trung bình tháng 8 của Hà Nội. Thống kê lượng mưa trung bình tháng 8 của Hà Nội trong nhiều năm liên tiếp sẽ cho ta những dự báo (ngày càng chính xác hơn) lượng mưa trung bình tháng 8 của Hà Nội trong những năm sắp tới.

II. Trung vị

1. Định nghĩa

Sắp thứ tự mẫu số liệu gồm n số liệu thành một dãy không giảm (hoặc không tăng).

- Nếu n là lẻ thì số liệu đứng ở vị trí thứ $\frac{n+1}{2}$ (số đứng chính giữa) gọi là *trung vị*.

- Nếu n là chẵn thì số trung bình cộng của hai số liệu đứng ở vị trí thứ $\frac{n}{2}$ và $\frac{n}{2}+1$ gọi là *trung vị*.

Trung vị kí hiệu là M_e .

Ví dụ 2. Thời gian (tính theo phút) mà 10 người đợi ở bến xe buýt là:

2,8 1,2 3,4 14,6 1,3 2,5 4,2 1,9 3,5 0,8

Tìm trung vị của mẫu số liệu trên

Giải

Bước 1. Sắp xếp các số liệu của mẫu theo thứ tự không giảm

0,8 1,2 1,3 1,9 2,5 2,8 3,4 3,5 4,2 14,6

Bước 2. Xác định xem số các số liệu là số chẵn hay số lẻ để tìm trung vị:

Mẫu số liệu trên có 10 số. Số thứ năm và số thứ sáu lần lượt là 2,5 và 2,8.

$$M_e = \frac{2,5 + 2,8}{2} = 2,65$$

Vì vậy (phút).

Nhận xét

- Trung vị không nhất thiết là một số trong mẫu số liệu và dễ tính toán.

- Khi các số liệu trong mẫu không có sự chênh lệch lớn thì số trung bình cộng và trung vị xấp xỉ nhau.

2. Ý nghĩa

Nếu những số liệu trong mẫu có sự chênh lệch lớn thì ta nên chọn thêm trung vị làm đại diện cho mẫu số liệu đó nhằm điều chỉnh một số hạn chế khi sử dụng số trung bình cộng. Những kết luận về đối tượng thống kê rút ra khi đó sẽ tin cậy hơn.

Chẳng hạn, số trung bình cộng của mẫu số liệu thống kê trong **Ví dụ 2** là:

$$\bar{x} = \frac{2,8 + 1,2 + 3,4 + 14,6 + 1,3 + 2,5 + 4,2 + 1,9 + 3,5 + 0,8}{10} = 3,62 \quad (\text{phút})$$

Vì thế, nếu chọn thêm trung vị $M_e = 2,65$ (phút) làm đại diện cho mẫu số liệu đó thì kết luận về thời gian đợi ở bến xe buýt sẽ tin cậy hơn.

III. Tứ phân vị

1. Định nghĩa

Sắp thứ tự mẫu số liệu gồm N số liệu thành một dãy không giảm.

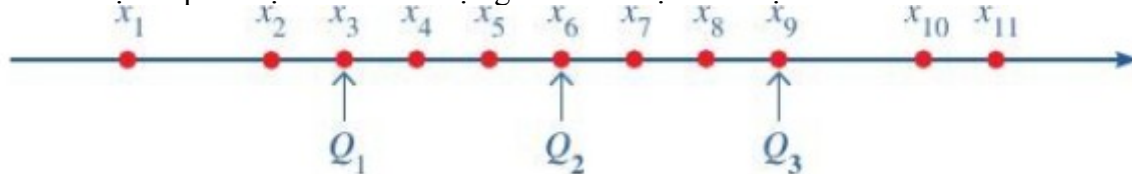
Tứ phân vị của mẫu số liệu trên là bộ ba giá trị: tứ phân vị thứ nhất, tứ phân vị thứ hai và tứ phân vị thứ ba; ba giá trị này chia mẫu số liệu thành bốn phần có số lượng phần tử bằng nhau.

- Tứ phân vị thứ hai Q_2 bằng trung vị.

- Nếu N là số chẵn thì tứ phân vị thứ nhất Q_1 bằng trung vị của nửa dãy phía dưới và tứ phân vị thứ ba Q_3 bằng trung vị của nửa dãy phía trên.

- Nếu N là số lẻ thì tứ phân vị thứ nhất Q_1 bằng trung vị của nửa dãy phía dưới (không bao gồm Q_2) và tứ phân vị thứ ba Q_3 bằng trung vị của nửa dãy phía trên (không bao gồm Q_2).

Ta minh họa tứ phân vị của mẫu số liệu gồm 11 số liệu trên trục số như sau:



Ví dụ 3. Tìm tứ phân vị của mẫu số liệu: 21 35 17 43 8 59 72 119

Biểu diễn tứ phân vị đó trên trục số.

Giải

Mẫu số liệu trên được sắp xếp theo thứ tự tăng dần như sau: 8 17 21 35 43 59 72 119

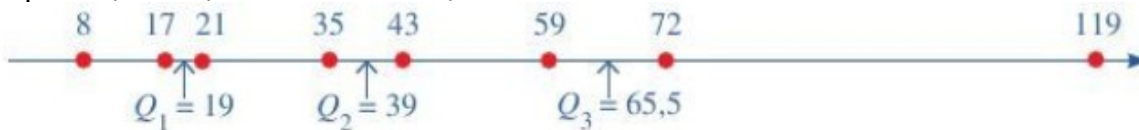
$$\text{Trung vị của mẫu số liệu trên là } \frac{35 + 43}{2} = 39$$

Trung vị của dãy 8 17 21 35 là $\frac{17+21}{2}=19$.

Trung vị của dãy 43 59 72 119 là: $\frac{59+72}{2}=65,5$.

Vậy $Q_1=19, Q_2=39, Q_3=65,5$.

Tứ phân vị đó được biểu diễn trên trục số như sau:



2. Ý nghĩa

- Trong thực tiễn, có những mẫu số liệu mà nhiều số liệu trong mẫu đó vẫn còn sự chênh lệch lớn so với trung vị. Ta nên chọn thêm những số khác cùng làm đại diện cho mẫu đó. Bằng cách lấy thêm trung vị của từng dãy số liệu tách ra bởi trung vị của mẫu nói trên, ta nhận được tứ phân vị đại diện cho mẫu số liệu đó.

- Bộ ba giá trị Q_1, Q_2, Q_3 trong tứ phân vị phản ánh độ phân tán của mẫu số liệu. Nhưng mỗi giá trị Q_1, Q_2, Q_3 lại đo xu thế trung tâm của phần số liệu tương ứng của mẫu đó.

IV. Một

1. Định nghĩa

Mốt của mẫu số liệu là giá trị có tần số lớn nhất trong bảng phân bố tần số và kí hiệu là M_o .

Chú ý: Một mẫu số liệu có thể có một hoặc nhiều mốt.

Ví dụ 4. Một trong bảng tần số thống kê số áo bán ra trong tháng đầu tiên của cửa hàng Bắc Tâm là bao nhiêu?

Cỡ áo	37	38	39	40	41	42	43
Tần số (Số áo bán được)	15	46	62	81	51	20	3

Giải.

Vì tần số lớn nhất là 81 và 81 tương ứng với cỡ áo 40 nên mốt của bảng trên là 40.

2. Ý nghĩa

Mốt của một mẫu số liệu đặc trưng cho số lần lặp đi lặp lại nhiều nhất tại một vị trí của mẫu số liệu đó. Dựa vào mốt, ta có thể đưa ra những kết luận (có ích) về đối tượng thống kê.

Chẳng hạn, trong Ví dụ 4, mốt trong bảng tần số thống kê số áo bán ra trong tháng đầu tiên của cửa hàng là 40. Do vậy, bác Tâm nên nhập về nhiều hơn cỡ áo 40 để bán trong tháng tiếp theo.

V. Tính hợp lý của số liệu thống kê

Ví dụ 5. Mẫu số liệu sau ghi lại cân nặng của 40 học sinh lớp 10 của một trường trung học phổ thông (đơn vị: ki-lô-gam):

30	32	45	45	45	47	48	44	44	49
49	49	52	51	50	50	53	55	54	54
54	56	57	57	58	58,5	58,5	60	60	60
60	63,5	63	62	69	58,5	88	85	72	71

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

b) Từ kết quả câu a), bước đầu xác định những số liệu bất thường trong mẫu số liệu trên.

Giải

a) Mẫu số liệu trên được sắp xếp theo thứ tự tăng dần như sau:

30	32	44	44	45	45	45	47	48	49
49	49	50	50	51	52	53	54	54	54
55	56	57	57	58	58,5	58,5	60	60	60
60	62	63	63,5	68,5	69	71	72	85	88

$$\frac{54+55}{2}=54,5$$

- Trung vị của mẫu số liệu trên là:

- Trung vị của nửa dãy phía dưới

30 32 44 44 45 45 45 47 48 49 49 49 50 50 51 52 53 54 54 54 là: $\frac{49+49}{2}=49$.

- Trung vị của nửa dãy phía trên 55 56 57 57 58 58,5 58,5 60 60 60

60 62 63 63,5 68,5 69 71 72 85 88 là: $\frac{60+60}{2}=60$

Vậy $Q_1=49; Q_2=54,5; Q_3=60$

b) Dựa vào trung vị, tứ phân vị của mẫu số liệu đã cho, bước đầu ta có thể thấy những số liệu bất thường trong mẫu số liệu đó là: 30 32 85 88.

Chú ý: Trong thực tiễn, những số liệu bất thường của mẫu số liệu được xác định bằng những công cụ toán học sâu sắc hơn.

PHẦN B. BÀI TẬP TỰ LUẬN

Câu 1. điểm thi HKI môn toán của tổ học sinh lớp 10C (quy ước làm tròn đến 0,5 điểm) liệt kê như sau: 2; 5; 7,5; 8; 5; 7; 6,5; 9; 4,5; 10.

Tính điểm trung bình của 10 học sinh đó (quy tròn đến chữ thập phân thứ nhất)

Câu 2. Cho các số liệu thống kê về sản lượng chè thu được trong 1 năm (kg/sào) của 20 hộ gia đình

111	112	112	113	114	114	115	114	115	116
112	113	113	114	115	114	116	117	113	115

Tính số trung vị

Câu 3. điểm điều tra về chất lượng sản phẩm mới (thang điểm 100) như sau:

80 65 51 48 45 61 30 35 84 83 60 58 75
72 68 39 41 54 61 72 75 72 61 50 65

Hãy tìm các tứ phân vị.

Câu 4. Cho các số liệu thống kê về sản lượng chè thu được trong 1 năm (kg/sào) của 20 hộ gia đình

111	112	112	113	114	114	115	114	115	116
112	113	113	114	115	114	116	117	113	115

Câu 5. điểm điều tra về chất lượng sản phẩm mới (thang điểm 100) như sau:

80 65 51 48 45 61 30 35 84 83 60 58 75
72 68 39 41 54 61 72 75 72 61 50 65

Tìm một của bảng số liệu trên.

PHẦN C. BÀI TẬP TRẮC NGHIỆM

Câu 1. Số áo bán được trong một quý ở cửa hàng bán áo sơ mi nam được thống kê như sau:

Cỡ áo	36	37	38	39	40	41	42
Tần số (Số áo bán được)	13	45	126	125	110	40	12

Giá trị một của bảng phân bố tần số trên bằng

A. 38.

B. 126.

C. 42.

D. 12.

Câu 2. Tiền lương hàng tháng của 7 nhân viên trong một công ty du lịch lần lượt là: 6,5; 8,4; 6,9; 7,2; 2,5; 6,7; 3,0 (đơn vị: triệu đồng). Số trung vị của dãy số liệu thống kê trên bằng

- A. 6,7 triệu đồng. B. 7,2 triệu đồng. C. 6,8 triệu đồng. D. 6,9 triệu đồng.

Câu 3. Điểm kiểm tra môn Toán cuối năm của một nhóm gồm 9 học sinh lớp 6 lần lượt là 1; 1; 3; 6; 7; 8; 8; 9; 10. Điểm trung bình của cả nhóm gần nhất với số nào dưới đây?

- A. 7,5. B. 7. C. 6,5. D. 5,9.

Câu 4. Các giá trị xuất hiện nhiều nhất trong mẫu số liệu được gọi là

- A. Một. B. Số trung bình. C. Số trung vị. D. Độ lệch chuẩn.

Câu 5. Thời gian chạy 50m của 20 học sinh được ghi lại trong bảng dưới đây:

Thời gian (giây)	8,3	8,4	8,5	8,7	8,8
Tần số	2	3	9	5	1

Hỏi trung bình mỗi học sinh chạy 50m hết bao lâu ?

- A. 8,54. B. 4. C. 8,50. D. 8,53.

Câu 6. Một tổ học sinh gồm 10 học sinh có điểm kiểm tra giữa học kì 2 môn toán như sau: 5;6;7;5;8;8;10;9;7;8. Tính điểm trung bình của tổ học sinh đó.

- A. 7. B. 8. C. 7,3. D. 7,5.

Câu 7. Một tổ học sinh gồm 10 học sinh có điểm kiểm tra cuối học kì 1 môn toán như sau: 7;5;6;6;6;8;7;5;6;9. Tìm một của dãy trên.

- A. $M_0 = 6$. B. $M_0 = 7$. C. $M_0 = 5$. D. $M_0 = 8$.

Câu 8. Một tổ học sinh gồm 10 học sinh có điểm kiểm tra giữa học kì 2 môn toán như sau: 5;6;7;5;8;8;10;9;7;8. Tính điểm trung bình của tổ học sinh đó.

- A. 7. B. 8. C. 7,3. D. 7,5.

Câu 9. Cân nặng của 40 học sinh lớp 10 trường THPT A được cho bởi bảng sau

Lớp cân nặng (kg)	[35; 37)	[37; 39)	[39; 41)	[41; 43]	
Tần số	6	9	11	14	N = 40

Tính số trung bình cộng của mẫu số liệu trên.

- A. $\bar{x} = 38,26$. B. $\bar{x} = 40,25$. C. $\bar{x} = 39,65$. D. $\bar{x} = 40,83$.

Câu 10. Kết quả điểm kiểm tra 15' môn Toán của 100 em học sinh được trình bày ở bảng sau:

Điểm	3	4	5	6	7	8	9	10	Cộng
Tần số	3	5	11	17	30	19	10	5	100

Số trung bình cộng của bảng phân bố tần số nói trên là

- A. 6,88. B. 7,12. C. 6,5. D. 7,22.

Câu 11. Một học sinh có điểm các bài kiểm tra Toán như sau: 8;4;9;8;6;6;9;9;9. Điểm trung bình môn Toán của học sinh đó (làm tròn đến 1 chữ số thập phân) là

- A. 7,3. B. 6,8. C. 8,5. D. 7,6.

Câu 12. Thống kê điểm kiểm tra môn Lịch Sử của 45 học sinh lớp 10A như sau:

Điểm	5	6	7	8	9	10
Số học sinh	2	11	9	16	4	3

Số trung vị trong điểm các bài kiểm tra đó là

- A. $8,1$ điểm. B. $7,4$ điểm. C. $7,5$ điểm. D. 8 điểm.

Câu 13. Cho mẫu số liệu thống kê $\{2; 4; 6; 8; 10\}$. Số trung bình của mẫu số liệu trên là:

- A. 7. B. 12. C. 6,5. D. 6.

Câu 14. Điểm kiểm tra của 24 học sinh được ghi lại trong bảng sau:

7	2	3	5	8	2
8	5	8	4	9	6
6	1	9	3	6	7
3	6	6	7	2	9

Tìm một của điểm điều tra.

- A. 2. B. 7. C. 6. D. 9.

Câu 15. Kết quả điểm kiểm tra 45 phút môn Hóa Học của 100 em học sinh được trình bày ở bảng sau:

Điểm	3	4	5	6	7	8	9	10	Cộng
Tần số	3	5	14	14	30	22	7	5	100

Số trung bình cộng của bảng phân bố tần số nói trên là

- A. $6,82$. B. 4. C. $6,5$. D. $7,22$.

Câu 16. Điều tra tiền lương một tháng của 100 người lao động trên địa bàn một xã ta có bảng phân bố tần số sau:

Tiền lương (VND)	5.000.000	6.000.000	7.000.000	8.000.000	9.000.000	9.500.000
Tần số	26	34	20	10	5	5

Tìm **một** của bảng phân bố tần số trên.

- A. 5.000.000. B. 6.000.000. C. 7.500.000. D. 9.500.000.

Câu 17. Cho bảng phân bố tần số sau: *khối lượng*²⁰ *học sinh lớp*^{10A}

Khối lượng (kg)	Tần số
50	4
51	5
52	6
55	3
56	2

Số trung bình cộng $\bar{x}$ của bảng số liệu đã cho là

- A. $\bar{x} = 53$. B. $\bar{x} = 52,8$. C. $\bar{x} = 52,2$. D. $\bar{x} = 52$.

Câu 18. Kết quả thi môn Toán giữa kì 1¹ của lớp ^{10A₃} trường THPT Ba Vì được thống kê như sau:

Điểm thi	5	6	7	8	9	10	Cộng
Tần số	5	7	8	12	8	5	45

Giá trị một M_0 của bảng phân bố tần số trên bằng

A. 5. B. 7. C. 8. D. 12.

Câu 19. Điểm thi toán cuối năm của một nhóm gồm 7 học sinh lớp 11 là 1; 3; 4; 5; 7; 8; 9. Số trung vị của dãy số liệu đã cho là

A. 6. B. 4. C. 7. D. 5.

Câu 20. Điểm thi toán cuối năm của một nhóm gồm 7 học sinh lớp 11 là 1; 3; 4; 5; 7; 8; 9. Số trung vị trên của dãy số liệu đã cho là

A. 8. B. 3. C. 7. D. 5.

Câu 21. Cho dãy số liệu thống kê $5, 7, 8, 11, 14, 15, 17, 20$. Số trung bình cộng của dãy số liệu trên là

A. 11. B. 12. C. 12.5. D. 12.125

Câu 22. Thời gian chạy $50m$ của 20 học sinh được ghi lại trong bảng dưới đây:

Thời gian (giây)	8,3	8,4	8,5	8,7	8,8
Tần số	2	3	9	5	1

Số trung bình cộng thời gian chạy của học sinh là

A. 8,54. B. 4. C. 8,50. D. 8,53.

Câu 23. Cho mẫu số liệu 10, 8, 6, 2, 4. Số trung bình cộng của mẫu là

A. 2,8. B. 2,4. C. 6. D. 8.

Câu 24. Một của một bảng phân bố tần số là

- A. tần số lớn nhất trong bảng phân bố tần số.
- B. giá trị có tần số lớn nhất trong bảng phân bố tần số.
- C. giá trị có tần số nhỏ nhất trong bảng phân bố tần số.
- D. tần số nhỏ nhất trong bảng phân bố tần số.

Câu 25. Cho bảng số liệu thống kê chiều cao của một nhóm học sinh như sau:

150	153	153	154	154	155	160	160	162	162	163	163	163	165	165	167
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Số trung vị của bảng số liệu nói trên là

A. 161. B. 153. C. 163. D. 156.

Câu 26. Cho bảng số liệu thống kê chiều cao của một nhóm học sinh như sau:

150	153	153	154	154	155	160	160	162	162	163	163	163	165	165	167
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Số trung vị dưới của bảng số liệu nói trên là

A. 161. B. 154. C. 163. D. 156.

Câu 27. Cho bảng phân bố tần số như sau:

Giá trị	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
Tần số	15	$9n-1$	12	n^2+7	14	10	$9n-20$	17

Tìm n để $M_o^{(1)} = x_2; M_o^{(2)} = x_4$ là hai một của bảng số liệu trên.

A. $n=1; n=8$. B. $n=8$. C. $n=1$. D. $n=9$.

Câu 28. Nhiệt độ trung bình hàng tháng trong một năm được ghi lại trong bảng sau

Tháng	1	2	3	4	5	6	7	8	9	10	11	12
Nhiệt độ	16	20	25	28	30	30	28	25	25	20	18	16

Một của dấu hiệu là

A. 20.

B. 25.

C. 28.

D. 30.

Câu 29. Cho bảng số liệu điểm kiểm tra môn Toán của 20 học sinh.

Điểm	4	5	6	7	8	9	10	Cộng
Số học sinh	1	2	3	4	5	4	1	20

Số trung vị của bảng số liệu trên là

A. 7.

B. 8.

C. 7,5.

D. 7,3.

Câu 30. Cho bảng số liệu điểm kiểm tra môn Toán của 20 học sinh.

Điểm	4	5	6	7	8	9	10	Cộng
Số học sinh	1	2	3	4	5	4	1	20

Số trung vị trên của bảng số liệu trên là

A. 7.

B. 8.

C. 8,5.

D. 7,3.