

VII

THỐNG KÊ

1. Số trung bình

- Giả sử ta có một mẫu số liệu là $x_1, x_2, \dots, x_n$.

Số trung bình (hay **số trung bình cộng**) của mẫu số liệu này, kí hiệu là $\bar{x}$, được tính bởi công thức $\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$.

- Giả sử mẫu số liệu được cho dưới dạng bảng tần số

Giá trị	x_1	x_2	...	x_k
Tần số	n_1	n_2	...	n_k

Khi đó, công thức tính số trung bình trở thành $\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_k x_k}{n}$, trong đó

$n = n_1 + n_2 + \dots + n_k$. Ta gọi n là cỡ mẫu.

Chú ý: Kí hiệu $f_k = \frac{n_k}{n}$ là tần số tương đối (hay còn gọi là tần suất) của x_k trong mẫu số liệu thì số trung bình còn có thể biểu diễn là $\bar{x} = f_1 x_1 + f_2 x_2 + \dots + f_k x_k$.

- Ý nghĩa của số trung bình: Số trung bình của mẫu số liệu được dùng làm đại diện cho các số liệu của mẫu. Nó là một số đo xu thế trung tâm của mẫu đó.

2. Trung vị và tứ phân vị

Sắp xếp lại mẫu số liệu theo thứ tự không giảm, ta được $x_1 \leq x_2 \leq \dots \leq x_n$.

- **Trung vị** của mẫu, kí hiệu là M_e , là giá trị ở chính giữa dãy $x_1, x_2, \dots, x_n$. Cụ thể:

+ Nếu $n = 2k + 1, k \in \mathbb{N}$ thì trung vị mẫu là $M_e = x_{k+1}$.

+ Nếu $n = 2k, k \in \mathbb{N}$ thì trung vị mẫu là $M_e = \frac{1}{2}(x_k + x_{k+1})$.

- Ý nghĩa của trung vị: Trung vị được dùng để đo xu thế trung tâm của mẫu số liệu. Trung vị là giá trị nằm ở chính giữa của mẫu số liệu theo nghĩa: luôn có ít nhất 50% số liệu trong mẫu lớn hơn hoặc bằng trung vị và ít nhất 50% số liệu trong mẫu nhỏ hơn hoặc bằng trung vị. Khi trong mẫu xuất hiện thêm một giá trị rất lớn hoặc rất nhỏ thì số trung bình sẽ bị thay đổi đáng kể nhưng trung vị thì ít thay đổi.

- **Tứ phân vị** của một mẫu ngẫu nhiên gồm 3 giá trị, đó là **tứ phân vị thứ nhất, thứ hai và thứ ba** (lần lượt kí hiệu là Q_1, Q_2, Q_3). Ba giá trị này chia tập hợp dữ liệu đã sắp xếp thành bốn phần đều nhau. Cụ thể:

+ Giá trị tứ phân vị thứ hai, Q_2 , chính là trung vị của mẫu.

+ Giá trị tứ phân vị thứ nhất, Q_1 , là trung vị của nửa số liệu đã sắp xếp bên trái Q_2 (không bao gồm Q_2 nếu n lẻ).

+ Giá trị tứ phân vị thứ ba, Q_3 , là trung vị của nửa số liệu đã sắp xếp bên phải Q_2 (không bao gồm Q_2 nếu n lẻ).

- **Ý nghĩa của tứ phân vị:** Các điểm tứ phân vị Q_1, Q_2, Q_3 chia mẫu số liệu đã sắp xếp theo thứ tự từ nhỏ đến lớn thành bốn phần, mỗi phần chứa khoảng 25% tổng số số liệu đã thu thập được. Tứ phân vị thứ nhất Q_1 còn được gọi là tứ phân vị dưới và đại diện cho nửa mẫu số liệu phía dưới. Tứ phân vị thứ ba Q_3 còn được gọi là tứ phân vị trên và đại diện cho nửa mẫu số liệu phía trên.

3. Một

- Cho một mẫu số liệu dưới dạng bảng tần số. Giá trị có tần số lớn nhất được gọi là **mốt** của mẫu số liệu, kí hiệu là M_o .

- **Ý nghĩa của mốt:** Mốt đặc trưng cho giá trị xuất hiện nhiều nhất trong mẫu.

Chú ý: Một mẫu số liệu có thể có nhiều mốt. Khi tất cả các giá trị trong mẫu số liệu có tần số xuất hiện bằng nhau thì mẫu số liệu đó không có mốt.

4. Khoảng biến thiên và khoảng tứ phân vị

Sắp xếp lại mẫu số liệu theo thứ tự không giảm, ta được $x_1 \leq x_2 \leq \dots \leq x_n$.

- **Khoảng biến thiên** của một mẫu số liệu, kí hiệu là R , là hiệu giữa giá trị lớn nhất và giá trị nhỏ nhất của mẫu số liệu đó, tức là: $R = x_n - x_1$.

- **Khoảng tứ phân vị**, kí hiệu là Δ_Q , là hiệu giữa Q_3 và Q_1 , tức là $\Delta_Q = Q_3 - Q_1$.

- **Ý nghĩa của khoảng biến thiên và khoảng tứ phân vị**

+ Khoảng biến thiên đặc trưng cho độ phân tán của toàn bộ mẫu số liệu.

+ Khoảng tứ phân vị đặc trưng cho độ phân tán của một nửa các số liệu, có giá trị thuộc đoạn từ Q_1 đến Q_3 trong mẫu.

+ Khoảng tứ phân vị không bị ảnh hưởng bởi các giá trị rất lớn hoặc rất bé trong mẫu.

- **Giá trị ngoại lệ**

+ Khoảng tứ phân vị được dùng để xác định các **giá trị ngoại lệ** trong mẫu, đó là các giá trị quá nhỏ hay quá lớn so với đa số các giá trị của mẫu. Cụ thể, phần tử x trong mẫu là giá trị ngoại lệ nếu $x > Q_3 + 1,5\Delta_Q$ hoặc $x < Q_1 - 1,5\Delta_Q$.

+ Khi mẫu có giá trị ngoại lệ, người ta thường sử dụng trung vị và khoảng tứ phân vị để đo mức độ tập trung và mức độ phân tán của đa số các phần tử trong mẫu số liệu.

5. Phương sai và độ lệch chuẩn

Giả sử ta có một mẫu số liệu là $x_1, x_2, \dots, x_n$.

- Phương sai của mẫu số liệu này, kí hiệu là S^2 , được tính bởi công thức

$$S^2 = \frac{1}{n} \left[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2 \right] = \frac{1}{n} (x_1^2 + x_2^2 + \dots + x_n^2) - \bar{x}^2.$$

trong đó $\bar{x}$ là số trung bình của mẫu số liệu.

- Căn bậc hai của phương sai được gọi là **độ lệch chuẩn**, kí hiệu là S .

Giả sử mẫu số liệu được cho dưới dạng bảng tần số

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

- Khi đó, công thức tính phương sai trở thành

$$S^2 = \frac{1}{n} \left[n_1 (x_1 - \bar{x})^2 + n_2 (x_2 - \bar{x})^2 + \dots + n_k (x_k - \bar{x})^2 \right]$$

trong đó $n = n_1 + n_2 + \dots + n_k$.

- Có thể biến đổi công thức tính phương sai trên thành $S^2 = \frac{1}{n} (n_1 x_1^2 + n_2 x_2^2 + \dots + n_k x_k^2) - \bar{x}^2$

- Ý nghĩa của phương sai và độ lệch chuẩn

+ Phương sai là trung bình cộng của các bình phương độ lệch từ mỗi giá trị của mẫu số liệu đến số trung bình.

+ Phương sai và độ lệch chuẩn được dùng để đo mức độ phân tán của các số liệu trong mẫu quanh số trung bình. Phương sai và độ lệch chuẩn càng lớn thì các giá trị của mẫu càng cách xa nhau (có độ phân tán lớn).

TÍNH SAI SỐ TUYỆT ĐỐI, ĐỘ CHÍNH XÁC CỦA MỘT SỐ GẦN ĐÚNG.

♦ Nếu a là số gần đúng của $\bar{a}$ thì $\Delta_a = |\bar{a} - a|$ được gọi là sai số tuyệt đối của số gần đúng a

♦ Nếu $\Delta_a \leq d$ thì $a - d \leq \bar{a} \leq a + d$ khi đó ta viết $\bar{a} = a \pm d$

d gọi là độ chính xác của số gần đúng.

Câu 1: Cho số $x = \frac{2}{7}$ và các giá trị gần đúng của x là 0,28; 0,29; 0,286; 0,3. Hãy xác định sai số tuyệt đối trong từng trường hợp và cho biết giá trị gần đúng nào là tốt nhất.

Câu 2: Hình chữ nhật có các cạnh $x = 10 \pm 0.01$ m, $y = 5 \pm 0.02$ m. Tính chu vi hình chữ nhật và sai số tuyệt đối của giá trị đó.

Câu 3: Chiều dài của một chi tiết máy được ghi $l = 20\text{cm} \pm 0,01\text{cm}$. Chiều dài đúng của cây thước là bao nhiêu?

- Câu 4:** Số lượng que diêm của một bao diêm được ghi trên bao bì là 45 ± 5 que. Số lượng que diêm của một bao diêm đúng quy định là bao nhiêu?
- Câu 5:** Cho $\bar{a} = \frac{12}{7} = 1,71428517...$ Hãy xác định số gần đúng của $\bar{a}$ với độ chính xác $d = 0,002$.
- Câu 6:** Cho $\bar{b} = \frac{1-\sqrt{5}}{2} = -0,61803398...$ Hãy xác định số gần đúng của $\bar{b}$ với độ chính xác $d = 0,0005$
- Câu 7:** An tính diện tích hình tròn có bán kính $r = 4\text{cm}$ bằng công thức $S = 3,145.4^2 = 50,32\text{cm}^2$. Biết rằng $3,14 < \pi < 3,15$, hãy ước lượng độ chính xác của S .
- Câu 8:** Độ dài đường chéo của một hình vuông có cạnh bằng 10cm được tính bởi $d = 10.1,415 = 14,15$. Cho biết $1,41 < \sqrt{2} < 1,42$. Hãy xác định độ chính xác của kết quả tìm được.
- Câu 9:** Cho tam giác ABC có độ dài ba cạnh đo được như sau $a = 12\text{cm} \pm 0,2\text{cm}$; $b = 10,2\text{cm} \pm 0,2\text{cm}$; $c = 8\text{cm} \pm 0,1\text{cm}$. Tính chu vi P của tam giác và đánh giá sai số tuyệt đối
- Câu 10:** Hình vuông có cạnh $x = 2 \pm 0,02\text{m}$. Tính diện tích hình vuông và sai số tuyệt đối của giá trị đó.

SAI SỐ TƯƠNG ĐỐI CỦA SỐ GẦN ĐÚNG

Sai số tương đối của số gần đúng a , kí hiệu là δ_a là tỉ số giữa sai số tuyệt đối và $|a|$.

$$\circ \delta_a = \frac{\Delta_a}{|a|}.$$

$$\circ \delta_a \leq \frac{d}{|a|}.$$

Do đó $\frac{d}{|a|}$ càng nhỏ thì chất lượng của phép đo đạc hay tính toán càng cao.

- Câu 11:** Trong một cuộc điều tra dân số, người ta viết dân số của một huyện là $3574625\text{người} \pm 25000$ người. Hãy ước lượng sai số tương đối của số gần đúng này.
- Câu 12:** Một thửa ruộng hình chữ nhật có chiều dài là $x = 23\text{m} \pm 0,01\text{m}$ và chiều rộng là $y = 15\text{m} \pm 0,01\text{m}$. Tính chu vi của thửa ruộng và ước lượng sai số tương đối của số gần đúng của chu vi qua phép đo đó.
- Câu 13:** Độ cao của một ngọn núi được ghi lại như sau $\bar{h} = 1372\text{m} \pm 0,2\text{m}$. Hãy ước lượng sai số tương đối của số gần đúng của phép đo đã cho.
- Câu 14:** Đo chiều dài của một cây thước, ta được kết quả $\bar{a} = 45 \pm 0,3(\text{cm})$. Hãy ước lượng sai số tương đối của số gần đúng trong phép đo trên.
- Câu 15:** Cho tam giác ABC có độ dài ba cạnh đo được như sau $a = 12\text{cm} \pm 0,2\text{cm}$; $b = 10,2\text{cm} \pm 0,2\text{cm}$; $c = 8\text{cm} \pm 0,1\text{cm}$. Tính chu vi P của tam giác và ước lượng sai số tuyệt đối, sai số tương đối của số gần đúng của chu vi qua phép đo.
- Câu 16:** Bạn Nam đo chiều dài của một cây cầu được kết quả là $150 \pm 0,2\text{m}$. Bạn Bình đo chiều cao của một dãy nhà của trường mà hai bạn học được kết quả $14 \pm 0,1\text{m}$. Hãy đánh giá độ chính xác về kết quả đo của hai bạn Nam và Bình theo sai số tương đối.

Tìm giá trị bất thường trong mẫu số liệu trên bằng cách sử dụng biểu đồ hộp.

Câu 25: Trên hai con đường A và B , trạm kiểm soát đã ghi lại tốc độ (km/h) của 30 chiếc ô tô trên mỗi con đường như sau:

Con đường A:	60	65	70	68	62	75	80	83	82	69	73	75	85	72	67
	88	90	85	72	63	75	76	85	84	70	61	60	65	73	76

Con đường B :

76	64	58	82	72	70	68	75	63	67	74	70	79	74	60
80	73	75	71	68	72	73	79	80	63	62	71	70	69	63

- a) Tính số trung bình, phương sai và độ lệch chuẩn của tốc độ ô tô trên mỗi con đường A, B.
- b) Theo em thì chạy xe trên con đường nào an toàn hơn?

Câu 26: Thống kê 100 học sinh tham gia kì thi học sinh giỏi toán. Kết quả được thống kê trong bảng sau:

Điểm	9	10	11	12	13	14	15	16	17	18	19	
Tần số	1	1	3	5	8	13	19	24	14	10	2	$N = 100$

Tính độ lệch chuẩn của bảng số liệu thống kê.

Câu 27: Sản lượng lúa các năm từ 2014 đến 2018 của hai tỉnh Thái Bình và Hậu Giang được cho ở bảng sau

Năm Tỉnh	2014	2015	2016	2017	2018
Thái Bình	1061,9	1061,9	1053,6	942,6	1030,4
Hậu Giang	1204,6	1293,1	1231,0	1261,0	1246,1

- a) Hãy tính độ lệch chuẩn của sản lượng lúa từng tỉnh.
- b) Tỉnh nào có sản lượng lúa ổn định hơn? Tại sao?

Câu 28: Số quần jeans bán được trong một quý ở một cửa hàng thời trang được thống kê ở bảng sau:

Size	26	27	28	29	30	31	32	Cộng
Tần số	128	105	119	16	8	15	9	400

Tìm phương sai và độ lệch chuẩn cho mẫu số liệu này.

Câu 29: Tiền thưởng của 50 cán bộ và nhân viên trong một công ty được cho ở bảng dưới đây:

Tiền thưởng	1	2	3	4	5	Cộng
Tần số	10	8	20	10	2	50

Tìm phương sai và đô lệch chuẩn cho mẫu số liêu này.

Câu 30: Cho mẫu số liệu thống kê: $\{2, 4, 6, 8, 10\}$. Tính phương sai và độ lệch chuẩn của mẫu số liệu trên.

Câu 31: Người ta tiến hành phỏng vấn một số người về chất lượng của một loại sản phẩm mới. Người điều tra yêu cầu cho điểm sản phẩm kết quả như sau:

80 65 51 48 45 61 30 35 84 83 60 50 75

72 68 39 41 54 61 72 75 72 61 58 65

b) Nhân xét gì về các kết quả nhân được.

Điểm	0	1	2	3	4	5	6	7	8	9	10	
Tần số	1	1	3	5	8	13	19	24	14	10	2	N=100

b) Tìm số trung bình, phương sai và độ lệch chuẩn.

Tháng	1	2	3	4	5	6	7	8	9	10	11	12
Vĩnh Phúc	16,1	17,0	20,2	23,7	27,3	28,8	29,9	28,2	27,2	24,6	21,4	18,2
TP Thủ Đức	25,8	26,7	27,9	28,9	28,3	27,5	27,1	27,1	26,8	26,7	26,4	25,7

b) Có nhận xét gì về sự biến động của nhiệt độ trung bình các tháng trong năm của tỉnh Vĩnh Phúc và thành phố Thủ Đức?

Tháng	1	2	3	4	5	6	7	8	9	10	11	12
Tuyên Quang	25	89	72	117	106	177	156	203	227	146	117	145
Cà Mau	180	223	257	245	191	111	141	134	130	122	157	173

b) Nêu nhận xét về sự thay đổi tổng số giờ nắng theo từng tháng ở mỗi tỉnh.

4	6	8	6	9	10	7	7	6	5
5	2	5	9	9	8	3	8	8	10

Tính khoảng biến thiên, khoảng tứ phân vị, phương sai và độ lệch chuẩn của mẫu số liệu.

Câu 37: Đo chiều cao (cm) của 11 học sinh tổ 2 của lớp 10A ta có mẫu số liệu: 145, 164, 164, 155, 154, 174, 158, 146, 180, 160, 157. Xác định các giá trị bất thường của mẫu số liệu.

3	5	7	9	10	6	8	3	4	6	5	7	8	10	7
9	3	6	4	7	8	9	10	6	9	7	4	5	3	8

86 | Ths. Lê Việt – THPT Thăng Long, Hà Nội – Phone: 0835221183 - Ôn thi cam kết điểm số

b. Cửa hàng đó nên nhập về nhiều hơn cỡ giày nào để bán trong tháng tiếp theo?

Câu 45: Bảng 2 cho biết nhiệt độ trung bình các tháng trong năm ở Hà Nội.

Tháng	1	2	3	4	5	6	7	8	9	10	11	12
Nhiệt độ ($^{\circ}\text{C}$)	16,4	17,0	20,2	23,7	27,3	28,8	28,9	28,2	27,2	24,6	21,4	18,2

a. Nhiệt độ trung bình trong năm ở Hà Nội là bao nhiêu?

b. Nhiệt độ trung bình của tháng có giá trị thấp nhất là bao nhiêu độ C? Cao nhất là bao nhiêu độ C?

Câu 46: Bảng 3 cho biết tổng diện tích rừng từ năm 2008 đến năm 2019 ở nước ta.

Năm	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Tổng diện tích rừng (triệu ha)	13,1	13,2	13,4	13,5	13,9	14,0	13,8	14,1	14,4	14,4	14,5	14,6

a. Diện tích rừng trung bình của nước ta từ năm 2008 đến năm 2019 là bao nhiêu?

b. Từ năm 2008 đến năm 2019, diện tích rừng của năm có giá trị thấp nhất là bao nhiêu triệu héc-ta? Cao nhất là bao nhiêu triệu héc-ta?

c. So với năm 2008, tỉ lệ tổng diện tích rừng của nước ta năm 2019 tăng lên được bao nhiêu phần trăm? Theo em, tỉ lệ tăng đó là cao hay thấp?

d. Hãy tìm hiểu số liệu về tổng diện tích rừng của tỉnh em đang sống trong một số năm gần đây.

Câu 47: Bốn bạn Bình, Cường, Hoa, Kiên cùng thi vào trường phổ thông chất lượng cao Bình Minh. Kết quả thi được cho bởi bảng thống kê sau:

Học sinh	Điểm Toán	Điểm Ngữ Văn	Điểm Tiếng Anh
Bình	10	8	9
Cường	6	7	5
Hoa	10	10	4
Kiên	9	5	10

Tính điểm trung bình kết quả thi 3 môn Toán, Ngữ Văn, Tiếng Anh của mỗi bạn và cho biết bạn nào trúng tuyển. Biết rằng, nếu muốn trúng tuyển, điểm trung bình các môn thi ở trên phải lớn hơn hoặc bằng 8 và không môn nào dưới 5 điểm.

Câu 48: Đầu năm học, nhà trường cho học sinh khám sức khỏe. Mẫu số liệu thống kê kết quả đo cân nặng (đơn vị: ki-lô-gam) của 7 bạn nam đầu tiên như sau:

64 58 62,1 55 67 61 60,5

Trung vị của mẫu số liệu trên là bao nhiêu?

Câu 49: Mẫu số liệu thống kê chiều cao (đơn vị: xăng-ti-mét) của 10 bạn tổ I lớp 10A như sau:

164 156 170 168 158 173 167 161 157 174

Trung vị của mẫu số liệu trên là bao nhiêu?

Câu 50: Mẫu số liệu thống kê số cân nặng (đơn vị: ki-lô-gam) tăng thêm của 7 trẻ sơ sinh trong ba tháng đầu tiên như sau:

0,9 1,0 1,1 1,14 1,18 1,2 1,3

Tứ phân vị của mẫu số liệu trên là bao nhiêu?

Câu 51: Mẫu số liệu thống kê thời gian (đơn vị: phút) đọc hết một cuốn sách của 9 bạn tổ I lớp 10A như sau: 102 130 118 127 115 138 121 109 132

Tứ phân vị của mẫu số liệu trên là bao nhiêu?

Câu 52: Một cửa hàng bán giày thống kê số đôi giày bán được trong Quý III năm 2020 như sau:

Cỡ giày	37	38	39	40	41	42	43	44
Số đôi giày bán được (Tần số)	41	49	50	71	53	46	27	5

a) Một trong bảng tần số thống kê số giày bán ra trong Quý III năm 2020 của cửa hàng trên là bao nhiêu?

b) Cửa hàng đó nên nhập về nhiều hơn cỡ giày nào để bán tiếp?

Câu 53: Theo báo cáo của WTTC (World Travel and Tourism Council), mức tăng đóng góp của ngành du lịch cho GDP năm 2021 so với năm 2020 tại một số khu vực (đơn vị: %) như sau: -42 -58 -41 -52 -50 -56 -37 -53 -45 -54.

a) Tính số trung bình, trung vị của dãy số liệu trên.

b) Giải thích ý nghĩa các giá trị thu được.

Câu 54: Tính các tứ phân vị cho dữ liệu về diện tích đất (đơn vị: km^2) của 266 quốc gia và vùng lãnh thổ cho số liệu như sau:

$$Q_1 = 20574,1; \quad Q_2 = 194690; \quad Q_3 = 1249825.$$

(Theo World Bank)

a) Có bao nhiêu quốc gia, vùng lãnh thổ có diện tích đất lớn hơn $194690 km^2$?

b) Diện tích đất của Việt Nam khoảng $310070 km^2$ có thuộc nhóm 25% quốc gia và vùng lãnh thổ có diện tích đất lớn nhất không?

Câu 55: Tìm số trung bình, trung vị, một và tứ phân vị của mỗi mẫu số liệu sau đây:

a) Số điểm mà năm vận động viên bóng rổ ghi được trong một trận đấu:

9 8 15 8 20

b) Giá của một số loại giày (đơn vị nghìn đồng): 350 300 650 300 450 500 300 250

c) Số kênh được chiếu của một số hãng truyền hình cáp: 36 38 33 34 32 30 34 35

Câu 56: Hãy chọn số đặc trưng đo xu thế trung tâm của mỗi mẫu số liệu sau. Giải thích và tính giá trị của số đặc trưng đó.

a) Số mặt trăng đã biết của các hành tinh:

Hành tinh	Thủy tinh	Kim tinh	Trái Đất	Hoả tinh	Mộc tinh	Thổ tinh	Thiên Vương tinh	Hải Vương tinh
-----------	-----------	----------	----------	----------	----------	----------	------------------	----------------

Số mặt trắng	0	0	1	2	63	34	27	13
-----------------	---	---	---	---	----	----	----	----

b) Số đường chuyền thành công trong một trận đấu của một số cầu thủ bóng đá:

32 24 20 14 23.

c) Chỉ số IQ của một nhóm học sinh:

80 10 83 10 10 94 11 10 10 100
2 3 8 0 6 4

d) Các sai số trong một phép đo: 10 15 18 15 14 13 42 15 12 14 42.

Câu 57: Số lượng học sinh giỏi Quốc gia năm học 2018-2019 của 10 trường Trung học phổ thông được cho như sau:

0 0 4 0 0 0 10 0 6 0

a) Tìm số trung bình, mốt, các tứ phân vị của mẫu số liệu trên.

b) Giải thích tại sao tứ phân vị thứ nhất và trung vị trùng nhau.

Câu 58: Bảng sau đây cho biết số chỗ ngồi của một số sân vận động được sử dụng trong Giải Bóng đá Vô địch Quốc gia Việt Nam năm 2018 (số liệu gần đúng).

Sân vận động	Cẩm phả	Thiên trường	Hàng Dẫy	Thanh Hóa	Mỹ Đình
Số chỗ ngồi	20120	21315	23405	20120	37546

(Theo vov, vn)

Các giá trị số trung bình, trung vị, mốt bị ảnh hưởng thế nào nếu bỏ đi số liệu chỗ ngồi của Sân vận động Quốc gia Mỹ Đình?

Câu 59: Tổng số ca mắc Covid-19 tính đến ngày 26-8-2021 tại Thành phố Hồ Chí Minh và một số tỉnh lân cận được thống kê như sau:

190174 81182 19728 19048 8155 6103 5807
4544 3760 3297 2541 2000 1934 1602 1195.

(Theo Bo Y tế?)

a) Tính số trung bình và trung vị cho dãy số liệu trên.

b) Giải thích tại sao số trung bình và trung vị lại khác nhau nhiều?

Câu 60: Lan thống kê số anh, chị, em ruột của các bạn trong lớp thu được bảng số liệu sau:

Số anh, chị, em ruột	0	1	2	3
Số bạn	4	25	5	1

Xác định mốt cho mẫu số liệu trên và giải thích ý nghĩa.

Câu 61: Thống kê GDP năm 2020 (đơn vị: tỉ đô la Mỹ) của 10 nước tại khu vực Đông Nam Á được kết quả như sau:

Brunei	Campuchia	Indonesia	Lào	Malaysia
12,02	25,95	1059,64	19,08	338,28
Myanmar	Philippines	Singapore	Thái Lan	Việt Nam
81,26	362,24	339,98	501,89	340,82

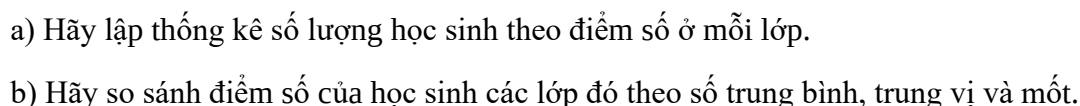
(Theo statista.com)

Số bóng đỏ	0	1	2	3
Số lần	10	30	40	20

Thời gian (đơn vị: phút)	5	6	7	8	35
Số thí sinh	1	3	5	2	1

Bác Dũng	2	7	3	6	1	4	1	4	5	1
Bác Thu	1	3	1	2	3	4	1	2	20	2

Năm	Tổng điểm	Năm	Tổng điểm	Năm	Tổng điểm	Năm	Tổng điểm
2020	150	2015	151	2010	133	2005	143
2019	177	2014	157	2009	161	2004	196
2018	148	2013	180	2008	159	2003	172
2017	155	2012	148	2007	168	2002	166
2016	151	2011	113	2006	131	2001	139



a) Hãy tính lượng mưa trung bình tại trạm quan trắc trên từ năm 2010 đến 2020.
b) Hãy tìm các tứ phân vị của mẫu số liệu đó.

93 | Ths. Lê Việt – THPT Thăng Long, Hà Nội – Phone: 0835221183 - Ôn thi cam kết điểm số

Câu 75: Minh và Thuỷ ghi lại số thư điện tử mà mỗi người nhận được mỗi ngày trong 10 ngày được lựa chọn ngẫu nhiên từ tháng 01/ 2021 ở bảng sau:

Minh	6	7	3	6	1	4	1	4	5	1
Thuỷ	2	3	1	2	3	4	1	2	20	2

- Hãy tìm số trung bình, trung vị và mốt của số thư điện tử mà mỗi bạn nhận được theo số liệu trên.
- Nếu so sánh theo số trung bình thì ai nhận được nhiều thư điện tử hơn?
- Nếu so sánh theo trung vị thì ai nhận được nhiều thư điện tử hơn?
- Nên dùng số trung bình hay số trung vị để so sánh xem ai nhận được nhiều thư điện tử hơn mỗi ngày?

Câu 76: Bạn Út ghi lại khối lượng của một số quả xoài Keo và xoài Thanh Ca ở bảng sau (đơn vị: gam).

Xoài Keo	370	320	350	290	300	350	310	330	340	370	390	
Xoài Thanh Ca	350	310	410	390	380	370	320	350	330	340	370	400

- Sử dụng số trung bình, hãy so sánh khối lượng của hai loại xoài.
- Sử dụng trung vị, hãy so sánh khối lượng của hai loại xoài.
- Hãy tính các tứ phân vị của hai mẫu số liệu trên.
- Nếu bạn Út mua 5 kg xoài Keo thì sẽ được khoảng bao nhiêu quả?



Nếu bạn Út mua 5 kg xoài Thanh Ca thì sẽ được khoảng bao nhiêu quả?

Câu 77: Số đơn vị hành chính cấp quận/huyện/thị xã của các tỉnh/thành phố khu vực Đồng bằng sông Hồng và khu vực Trung du và miền núi phía Bắc vào năm 2019 được cho như sau:

Đồng bằng sông Hồng: 30; 7; 7; 10; 10; 15; 9; 7; 5; 9; 6.

Trung du và miền núi phía Bắc: 10; 12; 7; 6; 8; 8; 7; 10; 9; 12; 9; 7; 11; 10.

(Nguồn: Tổng cục Thống kê)

- Mỗi khu vực nêu trên có bao nhiêu tỉnh/thành phố?
- Sử dụng số trung bình, hãy so sánh số đơn vị hành chính cấp quận/huyện/thị xã của các tỉnh/thành phố ở hai khu vực.
- Sử dụng trung vị, hãy so sánh số đơn vị hành chính cấp quận/huyện/thị xã của các tỉnh/thành phố ở hai khu vực.
- Hãy giải thích tại sao lại có sự khác biệt khi so sánh bằng số trung bình và trung vị.
- Hãy tìm tứ phân vị và mốt của hai khu vực.