

BÀI 15. CÁC SỐ ĐẶC TRƯNG ĐO ĐỘ PHÂN TÁN

• | Fanpage: Nguyễn Bảo Vương

A. KIẾN THỨC CẦN NHỚ

1. KHOẢNG BIẾN THIÊN VÀ KHOẢNG TỨ PHÂN VỊ

Khoảng biến thiên, kí hiệu là R , là hiệu số giữa giá trị lớn nhất và giá trị nhỏ nhất trong mẫu số liệu.

Ý nghĩa. Khoảng biến thiên dùng để đo độ phân tán của mẫu số liệu. Khoảng biến thiên càng lớn thì mẫu số liệu càng phân tán.

Ví dụ 1. Điểm kiểm tra học kì môn Toán của các bạn Tổ 1, Tổ 2 lớp 10 A được cho như sau:

Tổ 1: 7 8 8 9 8 8 8

Tổ 2: 10 6 8 9 9 7 8 7 8.

a) Điểm kiểm tra trung bình của hai tổ có như nhau không?

b) Tính các khoảng biến thiên của hai mẫu số liệu. Căn cứ trên chỉ số này, các bạn tổ nào học đồng đều hơn?

Lời giải

a) Điểm kiểm tra trung bình của hai tổ đều bằng 8.

b) Đối với Tổ 1: Điểm kiểm tra thấp nhất, cao nhất tương ứng là 7 ; 9. Do đó khoảng biến thiên là:

$$R_1 = 9 - 7 = 2.$$

Đối với Tổ 2: Điểm kiểm tra thấp nhất, cao nhất tương ứng là 6; 10. Do đó khoảng biến thiên là:

$$R_2 = 10 - 6 = 4.$$

Do $R_2 > R_1$ nên ta nói các bạn Tổ 1 học đều hơn các bạn Tổ 2.

Nhận xét. Sử dụng khoảng biến thiên có ưu điểm đơn giản, dễ tính toán song khoảng biến thiên chỉ sử dụng thông tin giá trị lớn nhất và giá trị nhỏ nhất mà bỏ qua thông tin từ tất cả các giá trị khác. Do đó, khoảng biến thiên rất dễ bị ảnh hưởng bởi các giá trị bất thường.

Khoảng tứ phân vị, kí hiệu là Δ_Q , là hiệu số giữa tứ phân vị thứ ba và tứ phân vị thứ nhất, tức là:

$$\Delta_Q = Q_3 - Q_1.$$

Ý nghĩa. Khoảng tứ phân vị cũng là một số đo độ phân tán của mẫu số liệu. Khoảng tứ phân vị càng lớn thì mẫu số liệu càng phân tán.

Chú ý. Một số tài liệu gọi khoảng biến thiên là biên độ và khoảng tứ phân vị là độ trải giữa

Ví dụ 2. Mẫu số liệu sau cho biết số ghế trống tại một rạp chiếu phim trong 9 ngày:

7 8 22 20 15 18 19 13 11.

Tìm khoảng tứ phân vị cho mẫu số liệu này.

Lời giải

Trước hết, ta sắp xếp mẫu số liệu theo thứ tự không giảm:

7 8 11 13 15 18 19 20 22.

Mẫu số liệu gồm 9 giá trị nên trung vị là số ở vị trí chính giữa $Q_2 = 15$.

Nửa số liệu bên trái là 7, 8, 11, 13 gồm 4 giá trị, hai phần tử chính giữa là 8, 11.

Do đó, $Q_1 = (8 + 11) : 2 = 9,5$.

Nửa số liệu bên phải là 18, 19, 20, 22 gồm 4 giá trị, hai phần tử chính giữa là 19, 20.

Do đó, $Q_3 = (19 + 20) : 2 = 19,5$.

Vậy khoảng tứ phân vị cho mẫu số liệu là $\Delta_Q = 19,5 - 9,5 = 10$.

2. PHƯƠNG SAI VÀ ĐỘ LỆCH CHUẨN

Phương sai là giá trị $s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$

Căn bậc hai của phương sai, $s = \sqrt{s^2}$, được gọi là độ lệch chuẩn

Chú ý. Người ta còn sử dụng đại lượng để đo độ phân tán của mẫu số liệu:

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n-1}$$

Ý nghĩa. Nếu số liệu càng phân tán thì phương sai và độ lệch chuẩn càng lớn.

Ví dụ 3. Mẫu số liệu sau đây cho biết sĩ số của 5 lớp khối 10 tại một trường Trung học:

43 45 46 41 40.

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

Lời giải

Số trung bình của mẫu số liệu là: $\bar{x} = \frac{43 + 45 + 46 + 41 + 40}{5} = 43$.

Ta có bảng sau:

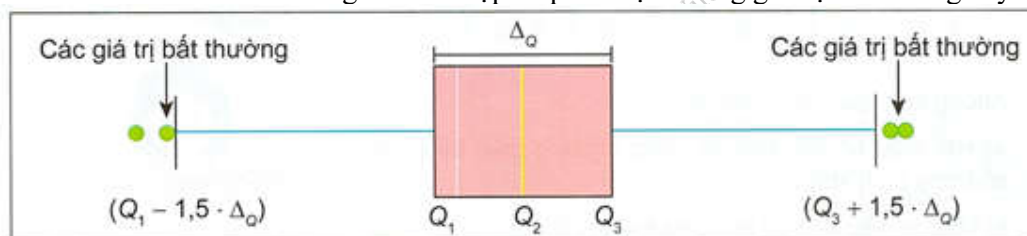
Giá trị	Độ lệch	Bình phương độ lệch
43	$43 - 43 = 0$	0
45	$45 - 43 = 2$	4
46	$46 - 43 = 3$	9
41	$41 - 43 = -2$	4
40	$40 - 43 = -3$	9
	Tổng	26

Mẫu số liệu gồm 5 giá trị nên $n = 5$. Do đó phương sai là: $s^2 = \frac{26}{5} = 5,2$.

Độ lệch chuẩn là: $s = \sqrt{5,2} \approx 2,28$.

3. PHÁT HIỆN SỐ LIỆU BẤT THƯỜNG HOẶC KHÔNG CHÍNH XÁC BẰNG BIỂU ĐỒ HỘP

Trong mẫu số liệu thống kê, có khi gặp những giá trị quá lớn hoặc quá nhỏ so với đa số các giá trị khác. Những giá trị này được gọi là giá trị bất thường. Chúng xuất hiện trong mẫu số liệu có thể do nhầm lẫn hay sai sót nào đó. Ta có thể dùng biểu đồ hộp để phát hiện những giá trị bất thường này.



Các giá trị lớn hơn $Q_3 + 1,5 \cdot \Delta_Q$ hoặc bé hơn $Q_1 - 1,5 \cdot \Delta_Q$ được xem là giá trị bất thường.

Ví dụ 4. Hàm lượng Natri (đơn vị mg) trong 100 g một số loại ngũ cốc được cho như sau:

0 340 70 140 200 180 210 150 100 130

140 180 190 160 290 50 220 180 200 210.

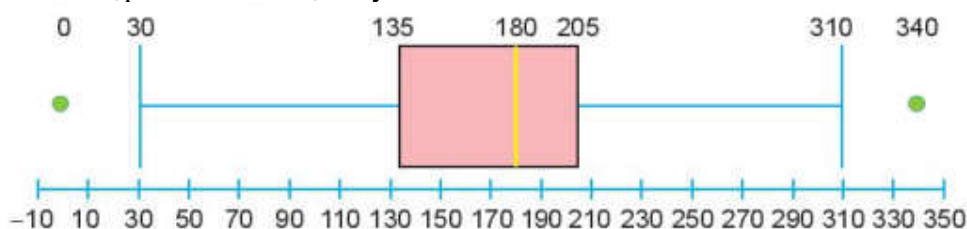
Tìm giá trị bất thường trong mẫu số liệu trên bằng cách sử dụng biểu đồ hộp.

Lời giải

Từ mẫu số liệu ta tính được $Q_1 = 135$ và $Q_3 = 205$. Do đó, khoảng tứ phân vị là:

$$\Delta_Q = 205 - 135 = 70.$$

Biểu đồ hộp cho mẫu số liệu này là:



Ta có $Q_1 - 1,5 \cdot \Delta_Q = 30$ và $Q_3 + 1,5 \cdot \Delta_Q = 310$ nên trong mẫu số liệu có hai giá trị được xem là bất thường là 340mg (lớn hơn 310mg) và 0mg (bé hơn 30mg).

B. CÁC DẠNG TOÁN THƯỜNG GẶP**BÀI TẬP SÁCH GIÁO KHOA, SÁCH BÀI TẬP**

- Câu 1.** Mẫu số liệu thống kê tiền lương (đơn vị: triệu đồng/tháng) của 8 cán bộ trong một tổ của công ty là: 8 8,5 10 9 10,5 9,5 11 12
 Tìm khoảng biến thiên của mẫu số liệu trên.

Lời giải

Trong mẫu số liệu trên, số lớn nhất là 12 và số nhỏ nhất là 8. Vậy khoảng biến thiên của mẫu số liệu trên là: $R = x_{\max} - x_{\min} = 12 - 8 = 4$ (triệu đồng/tháng).

- Câu 2.** Mẫu số liệu thống kê chiều cao (đơn vị: mét) của 12 cây thông là: 30,5 31 30,1 33,2 30,7 34,8 35 34,5 31,6 32,8 31,5 34,9
 Tìm khoảng tứ phân vị của mẫu số liệu trên.

Lời giải

Sắp xếp mẫu số liệu trên theo thứ tự không giảm, ta được:
 30,1 30,5 30,7 31 31,5 31,6 32,8 33,2 34,5 34,8 34,9 35

Trung vị của mẫu số liệu trên là: $\frac{31,6 + 32,8}{2} = 32,2$.

Trung vị của dãy 30,1; 30,5; 30,7; 31; 31,5; 31,6 là: $\frac{30,7 + 31}{2} = 30,85$.

Trung vị của dãy 32,8; 33,2; 34,5; 34,8; 34,9; 35 là: $\frac{34,5 + 34,8}{2} = 34,65$.

Vậy $Q_1 = 30,85$, $Q_2 = 32,2$, $Q_3 = 34,65$. Do đó khoảng tứ phân vị của mẫu số liệu trên là:
 $\Delta_Q = Q_3 - Q_1 = 34,65 - 30,85 = 3,8(m)$.

- Câu 3.** Kết quả 5 lần nhảy xa (đơn vị: mét) của bạn Huy và bạn Tùng cho ở bảng sau:

Huy	2,2	2,5	2,4	2,6	2,3
Tùng	2,0	2,8	2,5	2,4	2,3

- a) Kết quả trung bình của hai bạn có bằng nhau hay không?
 b) Tính phương sai và độ lệch chuẩn của mẫu số liệu thống kê kết quả 5 lần nhảy xa của mỗi bạn.
 Từ đó cho biết bạn nào có kết quả nhảy xa ổn định hơn?

Lời giải

- a) Gọi kết quả trung bình của bạn Huy và bạn Tùng lần lượt là \bar{x}_H, \bar{x}_T . Ta có:

$$\bar{x}_H = \frac{2,2 + 2,5 + 2,4 + 2,6 + 2,3}{5} = 2,4(m)$$

$$\bar{x}_T = \frac{2,0 + 2,8 + 2,5 + 2,4 + 2,3}{5} = 2,4(m)$$

Vậy kết quả trung bình của hai bạn bằng nhau.

- b) Gọi phương sai tương ứng với mẫu số liệu thống kê kết quả 5 lần nhảy của Huy và Tùng lần lượt là: s_H^2, s_T^2 . Ta có:

$$s_H^2 = \frac{(2,2 - 2,4)^2 + (2,5 - 2,4)^2 + (2,4 - 2,4)^2 + (2,6 - 2,4)^2 + (2,3 - 2,4)^2}{5} = 0,02;$$

$$s_T^2 = \frac{(2,0 - 2,4)^2 + (2,8 - 2,4)^2 + (2,5 - 2,4)^2 + (2,4 - 2,4)^2 + (2,3 - 2,4)^2}{5} = 0,068.$$

Ta cũng có độ lệch chuẩn của mẫu số liệu thống kê kết quả 5 lần nhảy của Huy và Tùng lần lượt là:

$$s_H = \sqrt{s_H^2} = \sqrt{0,02}(m); \quad s_T = \sqrt{s_T^2} = \sqrt{0,068}(m).$$

Do $s_H^2 = 0,02 < s_T^2 = 0,068$ nên bạn Huy có kết quả nhảy xa ổn định hơn bạn Tùng.

- Câu 4.** Nêu các giá trị bất thường của mẫu số liệu thống kê sau:
 0 1 13 16 17 18 19 20 21 22 23 24 28 37 38

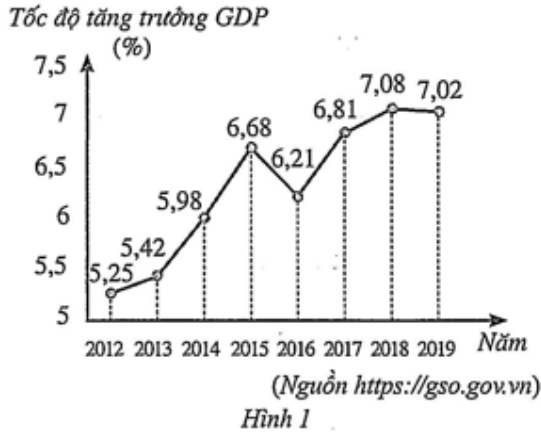
Lời giải

Mẫu số liệu trên có tứ phân vị là $Q_1 = 16; Q_2 = 20; Q_3 = 24$. Suy ra $\Delta_Q = Q_3 - Q_1 = 24 - 16 = 8$.

Các giá trị $0,1$ (nhỏ hơn $Q_1 - \frac{3}{2}\Delta_Q = 16 - \frac{3}{2} \cdot 8 = 4$) và các giá trị $37,38$ (lớn hơn

$Q_3 + \frac{3}{2}\Delta_Q = 24 + \frac{3}{2} \cdot 8 = 36$) là các giá trị bất thường của mẫu số liệu đã cho.

Câu 5. Biểu đồ đoạn thẳng ở Hình 1 Tốc độ tăng trưởng GDP liệu đó. (Nguồn <https://gso.gov.vn>)



d) Tính phương sai và độ lệch chuẩn Hình 1 của mẫu số liệu đó.

Lời giải

a) Mẫu số liệu thống kê tốc độ tăng trưởng GDP nhận được từ biểu đồ trên là: 5,25 5,42 5,98 6,68 6,21 6,81 7,08 7,02

b) Trong mẫu số liệu trên, số lớn nhất là 7,08 và số nhỏ nhất là 5,25. Vậy khoảng biến thiên của mẫu số liệu đó là: $R = x_{\max} - x_{\min} = 7,08 - 5,25 = 1,83(\%)$.

c) Sắp xếp mẫu số liệu trên theo thứ tự tăng dần, ta được: 5,25 5,42 5,98 6,21 6,68 6,81 7,02 7,08

Vậy ta có tứ phân vị của mẫu số liệu đó là:

$$Q_1 = \frac{5,42 + 5,98}{2} = 5,7(\%), \quad Q_2 = \frac{6,21 + 6,68}{2} = 6,445(\%),$$

$$Q_3 = \frac{6,81 + 7,02}{2} = 6,915(\%).$$

Do đó khoảng tứ phân vị của mẫu số liệu đó là: $\Delta_Q = Q_3 - Q_1 = 6,915 - 5,7 = 1,215(\%)$.

d) Số trung bình cộng của mẫu số liệu trên là:

$$\bar{x} = \frac{5,25 + 5,42 + 5,98 + 6,68 + 6,21 + 6,81 + 7,08 + 7,02}{8} = 6,30625(\%).$$

Ta có:

$$(5,25 - 6,30625)^2 + (5,42 - 6,30625)^2 + (5,98 - 6,30625)^2 + (6,68 - 6,30625)^2 + (6,21 - 6,30625)^2 + (6,81 - 6,30625)^2 + (7,08 - 6,30625)^2 + (7,02 - 6,30625)^2 = 3,5183875.$$

$$\text{Phương sai của mẫu số liệu trên là: } s^2 = \frac{3,5183875}{8} \approx 0,44.$$

$$\text{Độ lệch chuẩn của mẫu số liệu trên là: } s \approx \sqrt{0,44} \approx 0,66 (\%).$$

Câu 6. Mỗi khẳng định sau **đúng** hay **sai**?

- (1) Nếu các giá trị của mẫu số liệu càng tập trung quanh giá trị trung bình thì độ lệch chuẩn càng lớn.
- (2) Khoảng biến thiên chỉ sử dụng thông tin của giá trị lớn nhất và bé nhất, bỏ qua thông tin của các giá trị còn lại.
- (3) Khoảng tứ phân vị có sử dụng thông tin của giá trị lớn nhất, giá trị bé nhất.
- (4) Khoảng tứ phân vị chính là khoảng biến thiên của nửa dưới mẫu số liệu đã sắp xếp.
- (5) Các số đo độ phân tán đều không âm.

Lời giải

Khẳng định (1): Nếu các giá trị của mẫu số liệu càng tập trung quanh giá trị trung bình thì độ lệch của mỗi giá trị so với giá trị trung bình càng nhỏ (tức là $x_i - \bar{x}$ càng nhỏ, với $i = 1; 2; \dots; n$), dẫn đến độ lệch chuẩn càng nhỏ.

\Rightarrow Sai

Khẳng định (2): Khoảng biến thiên R bằng hiệu số giữa giá trị lớn nhất và giá trị nhỏ nhất nên chỉ sử dụng thông tin của giá trị lớn nhất và bé nhất \Rightarrow (2) Đúng.

Khẳng định (3): Khoảng tứ phân vị $\Delta_Q = Q_3 - Q_1$, các giá trị Q_1, Q_3 không bị ảnh hưởng bởi giá trị của giá trị lớn nhất và giá trị nhỏ nhất (với $n > 4$) \Rightarrow Sai

Khẳng định (4): Khoảng tứ phân vị chính là khoảng biến thiên của 50% số liệu chính giữa của mẫu số liệu đã sắp xếp

\Rightarrow Sai.

Khẳng định (5): Các số đo độ phân tán là

Khoảng biến thiên $R = \text{Số lớn nhất} - \text{Số nhỏ nhất} > 0$

Trước khi tính khoảng tứ phân vị thì mẫu số liệu được sắp xếp theo thứ tự không giảm

$\Rightarrow Q_3 > Q_1 \Rightarrow \Delta_Q = Q_3 - Q_1 > 0$

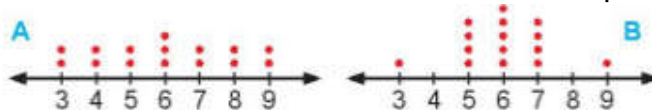
$$\text{Phương sai } s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} > 0$$

Độ lệch chuẩn: $s = \sqrt{s^2} > 0$

\Rightarrow Các số đo độ phân tán đều không âm

\Rightarrow (5) Đúng.

Câu 7. Cho hai biểu đồ chấm điểm biểu diễn hai mẫu số liệu A, B như sau:



Số chấm trên mỗi giá trị biểu diễn cho tần số của giá trị đó.

Không tính toán, hãy cho biết:

a) Hai mẫu số liệu này có cùng khoảng biến thiên và số trung bình không?

b) Mẫu số liệu nào có phương sai lớn hơn?

Lời giải

a) Cả 2 mẫu đều có $n = 15$.

Ta có cả 2 mẫu đều có giá trị nhỏ nhất là 3, giá trị lớn nhất là 9

Do đó cả 2 mẫu cùng khoảng biến thiên.

Cả 2 biểu đồ này có dạng đối xứng nên giá trị trung bình của hai mẫu A và B bằng nhau.

b) Từ biểu đồ ta thấy, mẫu A có các số liệu đồng đều và ổn định hơn mẫu B nên phương sai của mẫu A nhỏ hơn mẫu B.

Câu 8. Cho mẫu số liệu gồm 10 số dương không hoàn toàn giống nhau. Các số đo độ phân tán (khoảng biến thiên, khoảng tứ phân vị, độ lệch chuẩn) sẽ thay đổi như thế nào nếu:

a) Nhân mỗi giá trị của mẫu số liệu với 2.

b) Cộng mỗi giá trị của mẫu số liệu với 2.

Lời giải

$n=10$

Giả sử sau khi sắp xếp 10 số dương theo thứ tự không giảm thì được:



\Rightarrow Trung vị là giá trị trung bình của số thứ 5 và thứ 6.

$\Rightarrow Q_1$ là số thứ 3 và Q_3 là số thứ 8.

a) Khi nhân mỗi giá trị của mẫu số liệu với 2 thì:

+ Số lớn nhất tăng 2 lần và số nhỏ nhất tăng 2 lần

$\Rightarrow R$ tăng 2 lần

$+Q_1$ và Q_3 tăng 2 lần

=> Khoảng tứ phân vị $\Delta_Q = Q_3 - Q_1$ tăng 2 lần.

+ Giá trị trung bình tăng 2 lần

=> Độ lệch của mỗi giá trị so với giá trị trung bình $|x_i - \bar{x}|$ cũng tăng 2 lần

=> $(x_i - \bar{x})^2$ tăng 4 lần

=> Phương sai tăng 4 lần

=> Độ lệch chuẩn tăng 2 lần.

Vậy R tăng 2 lần, khoảng tứ phân vị tăng 2 lần và độ lệch chuẩn tăng 2 lần.

b) Cộng mỗi giá trị của mẫu số liệu với 2 thì

+ Số lớn nhất tăng 2 đơn vị và số nhỏ nhất tăng 2 đơn vị

=> R không đổi vì phần tăng thêm bị triệt tiêu cho nhau.

$+Q_1$ và Q_3 tăng 2 đơn vị

=> Khoảng tứ phân vị $\Delta_Q = Q_3 - Q_1$ không đổi vì phần tăng thêm bị triệt tiêu cho nhau.

+ Giá trị trung bình tăng 2 đơn vị

=> Độ lệch của mỗi giá trị so với giá trị trung bình $|x_i - \bar{x}|$ không đổi vì phần tăng thêm bị triệt tiêu cho nhau.

=> $(x_i - \bar{x})^2$ không đổi

=> Phương sai không đổi.

=> Độ lệch chuẩn không đổi.

Vậy khoảng biến thiên, khoảng tứ phân vị và độ lệch chuẩn đều không đổi.

Câu 9.

Từ mẫu số liệu về thuế thuốc lá của 51 thành phố tại một quốc gia, người ta tính được:

Giá trị nhỏ nhất bằng 2,5; $Q_1 = 36, Q_2 = 60, Q_3 = 100$; giá trị lớn nhất bằng 205.

a) Tỷ lệ thành phố có thuế thuốc lá lớn hơn 36 là bao nhiêu?

b) Chỉ ra hai giá trị sao cho có 50% giá trị của mẫu số liệu nằm giữa hai giá trị này.

c) Tìm khoảng tứ phân vị của mẫu số liệu.

Lời giải

a) Tỷ lệ thành phố có thuế thuốc lá lớn hơn 36 là tỷ lệ thành phố có thuế thuốc lá lớn hơn Q_1
Suy ra Có 75%

b) Ta thấy từ giá trị nhỏ nhất đến Q_2 có 50% giá trị của mẫu số liệu nằm giữa hai giá trị này

=> Ta chọn giá trị thứ nhất là 2,5 và 36.

c) Khoảng tứ phân vị $\Delta_Q = Q_3 - Q_1 = 100 - 36 = 64$

Câu 10.

Mẫu số liệu sau đây cho biết cân nặng của 10 trẻ sơ sinh (đơn vị kg):

2,977 3,155 3,920 3,412 4,236

2,593 3,270 3,813 4,042 3,387

Hãy tính khoảng biến thiên, khoảng tứ phân vị và độ lệch chuẩn cho mẫu số liệu này.

Lời giải

Sắp xếp theo thứ tự không giảm.

2,593 2,977 3,155 3,270 3,387 3,412 3,813 3,920 4,042 4,236

Khoảng biến thiên $R = 4,236 - 2,593 = 1,643$

Vì $n=10$ nên ta có:

$Q_1 = 3,155; Q_3 = 3,920$

Khoảng tứ phân vị $\Delta_Q = Q_3 - Q_1 = 3,920 - 3,155 = 0,765$

$\bar{x} \approx 3,481$

Ta có:

Giá trị	Độ lệch	Bình phương độ lệch
2,593	0,888	0,789
2,977	0,504	0,254

3,155	0,326	0,106
3,270	0,211	0,045
3,387	0,094	0,009
3,412	0,069	0,005
3,813	0,332	0,110
3,920	0,439	0,193
4,042	0,561	0,315
4,236	0,755	0,570

Độ lệch chuẩn: $s = \sqrt{0,2396} \approx 0,489$ Phương sai là: $s_2 = \frac{2,396}{10} = 0,2396$

Câu 11. Tỷ lệ thất nghiệp ở một số quốc gia vào năm 2007 (đơn vị %) được cho như sau:

7,8 3,2 7,7 8,7 8,6 8,4 7,2 3,6

5,0 4,4 6,7 7,0 4,5 6,0 5,4

Hãy tìm các giá trị bất thường (nếu có) của mẫu số liệu trên.

Lời giải

Sắp xếp theo thứ tự không giảm.

3,2 3,6 4,4 4,5 5,0 5,4 6,0 6,7 7,0 7,2 7,7 7,8 8,4 8,6 8,7

Vì $n=15$ nên $Q_2 = 6,7$

$Q_1 = 4,5; Q_3 = 7,8$

$\Delta_Q = Q_3 - Q_1 = 7,8 - 4,5 = 3,3$

$Q_3 + 1,5\Delta_Q = 12,75$

$Q_1 - 1,5\Delta_Q = -0,45$

Ta thấy không có giá trị nào dưới -0,45 và trên 12,75 nên không có giá trị bất thường.

Câu 12. Nhiệt độ trung bình (đơn vị: $^{\circ}\text{C}$) các tháng trong năm tại Hà Nội và Thành phố Hồ Chí Minh được cho trong bảng sau:

Tháng	1	2	3	4	5	6	7	8	9	10	11	12
Hà Nội	16,4	17,0	20,2	23,7	27,3	28,8	28,9	28,2	27,2	24,6	21,4	18,2
Thành phố Hồ Chí Minh	25,8	26,7	27,9	28,9	28,3	27,5	27,1	27,1	26,8	26,7	26,4	25,7

(Theo weatherspark.com)

a) Tính khoảng biến thiên, khoảng tứ phân vị và độ lệch chuẩn cho mỗi dãy số liệu trên.

b) Có nhận xét gì về sự biến động của nhiệt độ trung bình các tháng trong năm tại hai thành phố này?

Lời giải

a) Với dãy số liệu về nhiệt độ trung bình các tháng tại Hà Nội:

Giá trị nhỏ nhất là 16,4.

Giá trị lớn nhất là 28,9.

Khoảng biến thiên là: $R = 28,9 - 16,4 = 12,5$.

Dãy số liệu sắp xếp theo thứ tự không giảm:

16,4 17,0 18,2 20,2 21,4 23,7 24,6 27,2 27,3 28,2 28,8 28,9.

Trung vị là $Q_2 = (23,7 + 24,6) : 2 = 24,15$.

Nửa dữ liệu bên trái Q_2 là: 16,4 17,0 18,2 20,2 21,4 23,7

Do đó, $Q_1 = (18,2 + 20,2) : 2 = 19,2$.

Nửa dữ liệu bên phải Q_2 là: 24,6 27,2 27,3 28,2 28,8 28,9

Do đó, $Q_3 = (27,3 + 28,2) : 2 = 27,75$.

Khoảng tứ phân vị cho mẫu số liệu là: $\Delta_Q = Q_3 - Q_1 = 27,75 - 19,2 = 8,55$.

Số trung bình của mẫu số liệu là: $\bar{x} = \frac{16,4 + 17,0 + \dots + 18,2}{12} \approx 23,49$.

Độ lệch chuẩn: $s_1 = \sqrt{\frac{(16,4 - 23,49)^2 + \dots + (18,2 - 23,49)^2}{12}} \approx 4,52$.

Làm tương tự với dãy số liệu về nhiệt độ trung bình cho các tháng tại Thành phố Hồ Chí Minh ta có:

Khoảng biến thiên: $R = 3,2$.

Khoảng tứ phân vị là: $\Delta_Q = 27,7 - 26,55 = 1,15$.

Độ lệch chuẩn $s_2 = 0,91$.

b) Khoảng biến thiên, khoảng tứ phân vị, độ lệch chuẩn của dãy số liệu về nhiệt độ trung bình các tháng tại Thành phố Hồ Chí Minh đều nhỏ hơn các số đặc trưng này tại Hà Nội nên ta khẳng định rằng nhiệt độ trung bình các tháng ở Thành phố Hồ Chí Minh ít biến động hơn.

Câu 13. Điểm thi môn Toán của các bạn trong lớp được cho trong bảng sau:

Điểm	0	5	6	7	10
Tần số	1	10	20	10	1

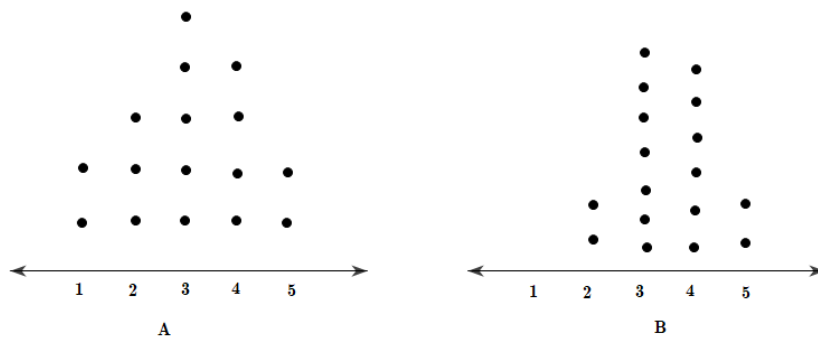
Tìm khoảng biến thiên của mẫu số liệu trên. Có nên dùng đại lượng này để đo độ phân tán của mẫu số liệu trên không?

Lời giải

Điểm thi thấp nhất là 0, cao nhất là 10. Do đó, khoảng biến thiên là $10 - 0 = 10$.

Hầu hết các bạn trong lớp có điểm 5, 6, 7 vì vậy dùng khoảng biến thiên để đo độ phân tán của dãy số liệu này sẽ không hợp lý.

Câu 14. Cho hai biểu đồ chấm điểm biểu diễn hai mẫu số liệu A, B như sau:



trong đó, mỗi chấm biểu diễn một giá trị trong mẫu số liệu.

Không tính, hãy cho biết:

- Độ lệch chuẩn của mẫu số liệu nào lớn hơn.
- Khoảng biến thiên của hai mẫu số liệu có như nhau không.

Lời giải

- Độ lệch chuẩn của dãy số liệu A lớn hơn.
- Không như nhau.

Câu 15. Cho hai dãy số liệu sau:

$A: 4 \quad 5 \quad 7 \quad 9 \quad 10;$

$B: 9 \quad 10 \quad 12 \quad 14 \quad 15.$

Không tính, hãy cho biết:

- Khoảng biến thiên của hai dãy có như nhau không.
- Độ lệch chuẩn của hai dãy có như nhau không.

Lời giải

HD. Dãy B có được là do cộng mỗi giá trị của dãy A với 5.

Câu 16. Điểm số của hai vận động viên bắn cung trong 10 lần bắn thử để chuẩn bị cho Olympic Tokyo 2020 được ghi lại như sau:

$VĐV A: 10 \quad 9 \quad 8 \quad 10 \quad 9 \quad 9 \quad 9 \quad 10 \quad 9 \quad 8;$

$VĐV B: 5 \quad 10 \quad 10 \quad 10 \quad 10 \quad 7 \quad 9 \quad 10 \quad 10 \quad 10.$

- a) Tìm khoảng biến thiên và độ lệch chuẩn của mỗi dãy số liệu trên.
b) Vận động viên nào có thành tích bắn thử ổn định hơn?

Lời giải

- a) Vận động viên A: Khoảng biến thiên = 2, Độ lệch chuẩn = 0,7.
Vận động viên B: Khoảng biến thiên = 5, Độ lệch chuẩn $\approx 1,64$.
b) Vì khoảng biến thiên, độ lệch chuẩn về thành tích của vận động viên A đều nhỏ hơn của vận động viên B nên dựa trên các tiêu chí này ta có thể kết luận vận động viên A có thành tích ổn định hơn.

Câu 17. Trong các dãy số liệu sau, dãy nào có độ lệch chuẩn lớn nhất?

- (a) 98 99 100 101 102
(b) 2 4 6 8 10
(c) 2 10.

Lời giải

Dãy số liệu (c).

Câu 18. Mẫu số liệu sau là chiều cao (đơn vị: cm) của các bạn trong tổ của Lan:

165 168 157 162 165 165 179 148 170 167.

- a) Tính khoảng tứ phân vị của mẫu số liệu trên.
b) Khoảng tứ phân vị có bị ảnh hưởng bởi chiều cao của bạn cao nhất, bạn thấp nhất không?

Lời giải

a) Sắp xếp dãy số liệu theo thứ tự không giảm:

148 157 162 165 165 165 167 168 170 179.

Vì $n = 10$ nên trung vị là trung bình cộng của 2 số ở chính giữa (vị trí 5 và 6):

$$Q_2 = \frac{165 + 165}{2} = 165$$

Ta tìm Q_1 là trung vị của nửa dữ liệu bên trái Q_2 là:

148 157 162 165 165 và tìm được $Q_1 = 162$.

nửa dữ liệu bên phải Q_2 là: 165 167 168 170 179 và tìm được $Q_3 = 168$.

Do đó, khoảng tứ phân vị là $\Delta_Q = 168 - 162 = 6$.

b) Khoảng tứ phân vị đo độ phân tán của 50% dữ liệu ở giữa nên không bị ảnh hưởng bởi giá trị nhỏ nhất, giá trị lớn nhất.

Câu 19. Bình dùng đồng hồ đo thời gian để một vật rơi tự do (đơn vị: giây) từ vị trí A đến vị trí B trong 10 lần cho kết quả như sau: Bình nghĩ là giá trị 0,290 ở lần đo thứ 9 không chính xác. Hãy kiểm tra nghi ngờ của Bình.

Lời giải

HD. Tìm khoảng tứ phân vị Δ_Q tương tự như Bài 5.17, tìm số trung bình \bar{x} sau đó kiểm tra xem giá trị 0,290 có nhỏ hơn $Q_1 - 1,5 \cdot \Delta_Q$ không.

Câu 20. Hãy chọn ngẫu nhiên trong lớp ra 5 bạn nam và 5 bạn nữ rồi đo chiều cao các bạn đó. So sánh xem chiều cao của các bạn nam hay các bạn nữ đồng đều hơn.

Lời giải

Chiều cao 5 HS nam	170	164	172	168	176
Chiều cao 5 HS nữ	155	152	157	162	160

+) Khoảng biến thiên chiều cao của các học sinh nam là:

$$176 - 164 = 12$$

+) Tứ phân vị: Q_1, Q_2, Q_3

Bước 1: Sắp xếp mẫu số liệu theo thứ tự không giảm: 164, 168, 170, 172, 176

Bước 2: $n = 5$, là số lẻ nên $Q_2 = M_e = 170$

Q_1 là trung vị của nửa số liệu 164, 168. Do đó $Q_1 = \frac{1}{2}(164 + 168) = 166$

Q_3 là trung vị của nửa số liệu 172, 176. Do đó $Q_3 = \frac{1}{2}(172 + 176) = 174$

Khoảng tứ phân vị $\Delta_Q = 174 - 166 = 8$

+) Khoảng biến thiên chiều cao của các học sinh nữ là: $162 - 152 = 10$

+) Tứ phân vị: Q_1, Q_2, Q_3

Bước 1: Sắp xếp mẫu số liệu theo thứ tự không giảm: 152, 155, 157, 160, 162

Bước 2: $n = 5$, là số lẻ nên $Q_2 = M_e = 157$

Q_1 là trung vị của nửa số liệu 152, 155. Do đó $Q_1 = \frac{1}{2}(152 + 155) = 153,5$

Q_3 là trung vị của nửa số liệu 160, 162. Do đó $Q_3 = \frac{1}{2}(160 + 162) = 161$

Khoảng tứ phân vị $\Delta_Q = 161 - 153,5 = 7,5$

Kết luận: So sánh khoảng biến thiên hay tứ phân vị thì theo mẫu số liệu trên, chiều cao của 5 bạn nữ là đồng đều hơn.

Câu 21. Hãy tìm độ lệch chuẩn, khoảng biến thiên, khoảng tứ phân vị và các giá trị ngoại lệ của các mẫu số liệu sau:

a) 6; 8; 3; 4; 5; 6; 7; 2; 4.

b) 13; 37; 64; 12; 26; 43; 29; 23.

Lời giải

a)

+) Số trung bình $\bar{x} = \frac{6 + 8 + 3 + 4 + 5 + 6 + 7 + 2 + 4}{9} = 5$

+) phương sai $S^2 = \frac{1}{9}(6^2 + 8^2 + \dots + 4^2) - 5^2 = \frac{10}{3} \Rightarrow$ Độ lệch chuẩn $S = \sqrt{\frac{10}{3}} \approx 1,8$

Sắp xếp mẫu số liệu theo thứ tự không giảm: 2; 3; 4; 4; 5; 6; 6; 7; 8.

+) Khoảng biến thiên: $R = 8 - 2 = 6$

Tứ phân vị: Q_1, Q_2, Q_3

$Q_2 = M_e = 5$

Q_1 là trung vị của nửa số liệu 2; 3; 4; 4. Do đó $Q_1 = 3,5$

Q_3 là trung vị của nửa số liệu: 6; 6; 7; 8. Do đó $Q_3 = 6,5$

+) Khoảng tứ phân vị: $\Delta_Q = 6,5 - 3,5 = 3$

+) x là giá trị ngoại lệ trong mẫu nếu $x > 6,5 + 1,5.3 = 11$ hoặc $x < 3,5 - 1,5.3 = -1$

Vậy không có giá trị ngoại lệ trong mẫu số liệu trên.

b)

+) Số trung bình $\bar{x} = \frac{13 + 37 + 64 + 12 + 26 + 43 + 29 + 23}{8} = 30,875$

+) phương sai hoặc $S^2 = \frac{1}{8}(13^2 + 37^2 + \dots + 23^2) - 30,875^2 \approx 255,8$

\Rightarrow Độ lệch chuẩn $S \approx 16$

Sắp xếp mẫu số liệu theo thứ tự không giảm: 12; 13; 23; 26; 29; 37; 43; 64.

+) Khoảng biến thiên: $R = 64 - 12 = 52$

Tứ phân vị: Q_1, Q_2, Q_3

$Q_2 = M_e = 27,5$

Q_1 là trung vị của nửa số liệu 12;13;23;26. Do đó $Q_1 = 18$

Q_3 là trung vị của nửa số liệu: 29;37;43;64. Do đó $Q_3 = 40$

+) Khoảng tứ phân vị: $\Delta_Q = 40 - 18 = 22$

+) x là giá trị ngoại lệ trong mẫu nếu $x > 40 + 1,5.22 = 73$ hoặc $x < 18 - 1,5.22 = -15$

Vậy không có giá trị ngoại lệ trong mẫu số liệu trên.

Câu 22. Hãy tìm độ lệch chuẩn, khoảng biến thiên, khoảng tứ phân vị của các mẫu số liệu sau:

a)

Giá trị	-2	-1	0	1	2
Tần số	10	20	30	20	10

b)

Giá trị	0	1	2	3	4
Tần số	0,1	0,2	0,4	0,2	0,1

Lời giải

a) +) Số trung bình $\bar{x} = \frac{-2 \cdot 10 + (-1) \cdot 20 + 0 \cdot 30 + 1 \cdot 20 + 2 \cdot 10}{10 + 20 + 30 + 20 + 10} = 0$

+) phương sai $S^2 = \frac{1}{9} (10 \cdot (-2)^2 + 20 \cdot (-1)^2 + \dots + 10 \cdot 2^2) - 0^2 \approx 13,33$

\Rightarrow Độ lệch chuẩn $S \approx 3,65$

+) Khoảng biến thiên: $R = 2 - (-2) = 4$

Tứ phân vị: $Q_2 = 0; Q_1 = -1; Q_3 = 1$

+) Khoảng tứ phân vị: $\Delta_Q = 1 - (-1) = 2$

b) Giả sử cỡ mẫu $n = 10$. Khi đó mẫu số liệu trở thành:

Giá trị	0	1	2	3	4
Tần số	1	2	4	2	1

+) Số trung bình $\bar{x} = \frac{0 \cdot 0,1 + 1 \cdot 0,2 + 2 \cdot 0,4 + 3 \cdot 0,2 + 4 \cdot 0,1}{0,1 + 0,2 + 0,4 + 0,2 + 0,1} = 2$

+) phương sai $S^2 = \frac{1}{1} (0 \cdot 1 \cdot 0^2 + 0,2 \cdot 1^2 + \dots + 0,1 \cdot 4^2) - 2^2 = 1,2$

\Rightarrow Độ lệch chuẩn $S \approx 1,1$

+) Khoảng biến thiên: $R = 4 - 0 = 4$

Tứ phân vị: $Q_2 = 2; Q_1 = 1; Q_3 = 3$

+) Khoảng tứ phân vị: $\Delta_Q = 3 - 1 = 2$

Câu 23. Hãy so sánh số trung bình, phương sai và độ lệch chuẩn của ba mẫu số liệu sau:

Mẫu 1: 0,1; 0,3; 0,5; 0,5; 0,3; 0,7.

Mẫu 2: 1,1; 1,3; 1,5; 1,5; 1,3; 1,7.

Mẫu 3: 1; 3; 5; 5; 3; 7.

Lời giải

Mẫu 1:

+) Số trung bình: $\bar{x} = \frac{0,1 + 0,3 + 0,5 + 0,5 + 0,3 + 0,7}{6} = 0,4$

+) Phương sai $S^2 = \frac{1}{6} (0,1^2 + 0,3^2 + 0,5^2 + 0,5^2 + 0,3^2 + 0,7^2) - 0,4^2 \approx 0,0367$

+) Độ lệch chuẩn $S = \sqrt{S^2} \approx 0,19$

Mẫu 2:

+) Số trung bình: $\bar{x} = \frac{1,1 + 1,3 + 1,5 + 1,5 + 1,3 + 1,7}{6} = 1,4$

+) Phương sai $S^2 = \frac{1}{6} (1,1^2 + 1,3^2 + 1,5^2 + 1,5^2 + 1,3^2 + 1,7^2) - 1,4^2 \approx 0,0367$

+) Độ lệch chuẩn $S = \sqrt{S^2} \approx 0,19$

Mẫu 3:

+) Số trung bình: $\bar{x} = \frac{1+3+5+5+3+7}{6} = 4$

+) Phương sai $S^2 = \frac{1}{6}(1^2 + 3^2 + 5^2 + 5^2 + 3^2 + 7^2) - 4^2 \approx 3,67$

+) Độ lệch chuẩn $S = \sqrt{S^2} \approx 1,9$

Kết luận:

Số liệu ở mẫu 2 hơn số liệu ở mẫu 1 là 1 đơn vị, số trung bình của mẫu 2 hơn số trung bình mẫu 1 là 1 đơn vị, còn phương sai và độ lệch chuẩn là như nhau.

Số liệu ở mẫu 3 gấp 10 lần số liệu mẫu 1, số trung bình, phương sai và độ lệch chuẩn của mẫu 3 lần lượt gấp 10 lần, 100 lần và 10 lần mẫu 1.

Câu 24. Sản lượng lúa các năm từ 2014 đến 2018 của hai tỉnh Thái Bình và Hậu Giang được cho ở bảng sau (đơn vị nghìn tấn):

Năm	2014	2015	2016	2017	2018
Tỉnh					
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 và khoảng biến thiên của sản lượng lúa từng tỉnh.

b) Tỉnh nào có sản lượng lúa ổn định hơn? Tại sao?

Lời giải

a)

Tỉnh Thái Bình:

Số trung bình $\bar{x} = \frac{1061,9 + 1061,9 + 1053,6 + 942,6 + 1030,4}{5} = 1030,08$

Phương sai $S^2 = \frac{1}{5}(1061,9^2 + 1061,9^2 + 1053,6^2 + 942,6^2 + 1030,4^2) - 1030,08^2 = 2046,2$

= Độ lệch chuẩn $S = \sqrt{S^2} \approx 45,2$

+) Khoảng biến thiên $R = 1061,9 - 942,6 = 119,3$

Tỉnh Hậu Giang:

Số trung bình $\bar{x} = \frac{1204,6 + 1293,1 + 1231,0 + 1261,0 + 1246,1}{5} = 1247,16$

Phương sai $S^2 = \frac{1}{5}(1204,6^2 + 1293,1^2 + 1231,0^2 + 1261,0^2 + 1246,1^2) - 1247,16^2 = 875,13$

=> Độ lệch chuẩn $S = \sqrt{S^2} \approx 29,6$

+) Khoảng biến thiên $R = 1293,1 - 1204,6 = 88,5$

b)

So sánh khoảng biến thiên và độ lệch chuẩn ta đều thấy tỉnh Hậu Giang có sản lượng lúa ổn định hơn.

Câu 25. Kết quả điều tra mức lương hằng tháng của một số công nhân của hai nhà máy A và B được cho ở bảng sau (đơn vị: triệu đồng):

Công nhân nhà máy A	4	5	5	47	5	6	4	4	
Công nhân nhà máy B	2	9	9	8	10	9	9	11	9

nhân nhà máy B									
----------------------	--	--	--	--	--	--	--	--	--

- a) Hãy tìm số trung bình, mốt, tứ phân vị và độ lệch chuẩn của hai mẫu số liệu lấy từ nhà máy A và nhà máy B.
- b) Hãy tìm các giá trị ngoại lệ trong mỗi mẫu số liệu trên. Công nhân nhà máy nào có mức lương cao hơn? Tại sao?

Lời giải

a) Nhà máy A:

+) Số trung bình: $\bar{x} = \frac{4+5+5+47+5+6+4+4}{8} = 10$

+) Mốt: $M_o = 4, M_o = 5$

+) Tứ phân vị: Q_1, Q_2, Q_3

Sắp xếp mẫu số liệu theo thứ tự không giảm: 4; 4; 4; 5; 5; 5; 6; 47. $Q_2 = M_e = 5$

Q_1 là trung vị của nửa số liệu: 4; 4; 4; 5. Do đó $Q_1 = 4$

Q_3 là trung vị của nửa số liệu: 5; 5; 6; 47. Do đó $Q_3 = 5,5$

+) Phương sai $S^2 = \frac{1}{8}(4^2 + 5^2 + \dots + 4^2) - 10^2 = 196 \Rightarrow$ Độ lệch chuẩn $S = \sqrt{S^2} = 14$

Nhà máy B:

+) Số trung bình: $\bar{x} = \frac{2+9+9+8+10+9+9+11+9}{9} = 8,4$

+) Mốt: $M_o = 9$

+) Tứ phân vị: Q_1, Q_2, Q_3

Sắp xếp mẫu số liệu theo thứ tự không giảm: 2; 8; 9; 9; 9; 9; 10; 11

$Q_2 = M_e = 9$

Q_1 là trung vị của nửa số liệu: 2; 8; 9; 9. Do đó $Q_1 = 8,5$

Q_3 là trung vị của nửa số liệu: 9; 9; 10; 11. Do đó $Q_3 = 9,5$

+) Phương sai $S^2 = \frac{1}{9}(2^2 + 9^2 + \dots + 9^2) - 8,4^2 = 6,55 \Rightarrow$ Độ lệch chuẩn $S = \sqrt{S^2} = 2,56$

b)

Nhà máy A có: $\Delta_Q = 1,5$

Vậy giá trị ngoại lệ $x > 5,5 + 1,5.1,5 = 7,75$ hoặc $x < 4 - 1,5.1,5 = 1,75$ là 47.

Nhà máy B có: $\Delta_Q = 1$

Vậy giá trị ngoại lệ $x > 9,5 + 1,5.1 = 11$ hoặc $x < 8,5 - 1,5.1 = 7$ là 2.

Ta so sánh trung vị: $9 > 5$, do đó công nhân nhà máy B có mức lương cao hơn.

Chú ý

Ta không so sánh số trung bình vì có giá trị 47 quá lớn so với các giá trị còn lại.

Câu 26. Kiểm tra khối lượng của một số quả măng cụt của hai lô hàng A và B được kết quả như sau (đơn vị: gam)

Lô A	85	82	84	83	80	82	84	85	80	81	80	82	85	85
Lô B	81	80	82	84	82	82	85	80	80	83	84	86	78	87

- a) Hãy tìm khoảng biến thiên và khoảng tứ phân vị của khối lượng măng cụt ở mỗi lô.
- b) Hãy tìm phương sai và độ lệch chuẩn của khối lượng măng cụt ở mỗi lô.
- c) Khối lượng của măng cụt ở lô hàng nào đều hơn?

Lời giải

Sắp xếp khối lượng các quả măng cụt ở lô A và lô B theo thứ tự không giảm, ta được:

Lô	80	80	80	81	82	82	82	83	84	84	85	85	85	85
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

A														
Lô B	78	80	80	80	81	82	82	82	83	84	84	85	86	87

a) Đối với lô A, khối lượng cao nhất và thấp nhất tương ứng là 85 và 80. Do đó khoảng biến thiên của lô A là $R(A) = 85 - 80 = 5$.

Đối với lô B, khối lượng cao nhất và thấp nhất tương ứng là 87 và 78. Do đó khoảng biến thiên của lô B là $R(B) = 87 - 78 = 9$.

Đối với lô A, $Q_1^A = 81, Q_3^A = 85$ nên $\Delta_Q^A = 85 - 81 = 4$.

Đối với lô B, $Q_1^B = 80, Q_3^B = 84$ nên $\Delta_Q^B = 84 - 80 = 4$.

b) Khối lượng trung bình của cân nặng măng cụt lô A là $\bar{x}(A) = \frac{1}{14}(3.80 + 81 + 3.82 + 83 + 2.84 + 4.85) = \frac{579}{7}$.

Phương sai của cân nặng măng cụt lô A là $S(A)^2 = \frac{1}{14}(3.80^2 + 81^2 + 3.82^2 + 83^2 + 2.84^2 + 4.85^2) - \frac{579^2}{7^2} \approx 3,63$.

Độ lệch chuẩn của măng cụt lô A là $S(A) = \sqrt{S(A)^2} \approx \sqrt{3,63} \approx 1,91$.

Khối lượng trung bình của cân nặng măng cụt lô B là $\bar{x}(A) = \frac{1}{14}(78 + 3.80 + 81 + 3.82 + 83 + 2.84 + 85 + 86 + 87) = \frac{577}{7}$.

Phương sai của cân nặng măng cụt lô B là $S(A)^2 = \frac{1}{14}(78^2 + 3.80^2 + 81^2 + 3.82^2 + 83^2 + 2.84^2 + 85^2 + 86^2 + 87^2) - \frac{577^2}{7^2} \approx 6,10$.

Độ lệch chuẩn của măng cụt lô B là $S(B) = \sqrt{S(B)^2} \approx \sqrt{6,10} \approx 2,47$.

c) Sử dụng khoảng biến thiên và độ lệch chuẩn ta đều thấy khối lượng măng cụt ở lô A đều hơn lô B.

Câu 27. Một bệnh viện thống kê số ca nhập viện do tai nạn giao thông mỗi ngày trong tháng 9/2020 ở bảng sau:

Số ca	0	1	2	3	4	5	6	7	8	9	12	15
Số ngày	2	3	4	6	3	2	2	3	2	1	1	1

a) Hãy tìm khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu.

b) Hãy tìm phương sai và độ lệch chuẩn của mẫu số liệu.

c) Xác định các giá trị ngoại lệ (nếu có) của mẫu số liệu.

Lời giải

a) Khoảng biến thiên của mẫu số liệu là $15 - 0 = 15$.

Do cỡ mẫu $n = 30$ nên tứ phân vị thứ nhất bằng số hạng lớn thứ 8 trong dãy số liệu, vậy $Q_1 = 2$.

Tứ phân vị thứ ba bằng số hạng thứ 22 trong dãy số liệu, vậy $Q_3 = 7$.

Khoảng tứ phân vị của mẫu số liệu là $\Delta_Q = 7 - 2 = 5$.

b) Trung bình của mẫu số liệu là $\bar{x} = \frac{136}{30}$.

Phương sai của mẫu số liệu là $S^2 = 11,72$.

Độ lệch chuẩn của mẫu số liệu là $S = 3,42$.

c) Ta có $Q_1 - 1,5\Delta_Q = 2 - 1,5.5 = -5,5$ và $Q_3 + 1,5\Delta_Q = 7 + 1,5.5 = 14,5$ nên mẫu có 1 giá trị ngoại lệ là 15.

Câu 28. Kết quả bài thi môn Toán của các bạn học sinh tổ 1 và tổ 2 cho ở bảng sau:

Tổ 1	7	8	9	6	7	8	7	9	10	7	8	6	8	9	8
------	---	---	---	---	---	---	---	---	----	---	---	---	---	---	---

Tổ 2	6	7	8	7	9	5	8	8	9	10	7	8	0	9	7
------	---	---	---	---	---	---	---	---	---	----	---	---	---	---	---

- a) Sử dụng số trung bình, hãy so sánh điểm thi của các bạn tổ 1 và tổ 2.
b) Sau khi bỏ đi các giá trị ngoại lệ (nếu có) ở các điểm thi mỗi tổ, hãy so sánh lại điểm thi của các bạn tổ 1 và tổ 2.
c) Nên dùng số trung bình hay trung vị để so sánh điểm thi của các bạn tổ 1 và tổ 2.

Lời giải

- a) Điểm trung bình của học sinh tổ 1 và tổ 2 lần lượt là 7,8 và 7,2. Nếu so sánh theo số trung bình thì điểm thi các bạn tổ 1 cao hơn điểm thi các bạn tổ 2.
b) Tổ 1 có $Q_1 = 7; Q_2 = 8; Q_3 = 9; \Delta_Q = 9 - 7 = 2$. Điểm số các bạn tổ 1 không có giá trị ngoại lệ nào.

Tổ 2 có $Q_1 = 7; Q_2 = 8; Q_3 = 9; \Delta_Q = 9 - 7 = 2$. Điểm số các bạn tổ 2 có 1 giá trị ngoại lệ là 0.

Sau khi bỏ đi điểm 0 này thì điểm trung bình của các bạn tổ 2 là 7,71.

Vậy điểm các bạn tổ 2 gần bằng điểm các bạn tổ 1.

- c) Nên dùng số trung vị để so sánh điểm thi của các bạn tổ 1 và tổ 2 vì trong điểm thi của các bạn tổ 2 có xuất hiện giá trị ngoại lệ.

Câu 29. Bảng sau ghi giá bán ra lúc 11 giờ của 2 mã cổ phiếu A và B trong 10 ngày liên tiếp (đơn vị: nghìn đồng).

A	94	93,2	95	96,6	96	94	97	95,8	98	99,4
B	80	80,3	80,5	80,5	80,1	80,1	79,7	79,5	79,6	80

Người ta lập bảng sau để theo dõi độ dao động giá của từng mã cổ phiếu sau mỗi ngày giao dịch.

A	-0,8	1,8							
B	0,3	0,2							

- a) Hãy điền các số liệu còn lại vào bảng trên.
b) Hãy tính độ lệch chuẩn, khoảng biến thiên, khoảng tứ phân vị của độ dao động giá mỗi ngày của hai mã cổ phiếu trên.
c) Một cổ phiếu được gọi là có rủi ro cao nếu nó có biên độ dao động giá lớn. Hãy cho biết trong hai mã cổ phiếu trên, mã nào có độ rủi ro cao hơn.

Lời giải

a)

A	-0,8	1,8	1,6	-0,6	-2	3	-1,2	2,2	1,4
B	0,3	0,2	0	-0,4	0	-0,4	-0,2	0,1	0,4

- b) Độ lệch chuẩn của độ dao động giá mã cổ phiếu A và B lần lượt là 1,66 và 0,27.
Khoảng biến thiên của độ dao động giá mã cổ phiếu A và B lần lượt là 5,0 và 0,8.
Khoảng tứ phân vị của độ dao động giá mã cổ phiếu A và B lần lượt là 3 và 0,55.
c) Mã cổ phiếu A có độ rủi ro cao hơn mã cổ phiếu B.

Câu 30. Hãy tìm phương sai, khoảng biến thiên, khoảng tứ phân vị và giá trị ngoại lệ (nếu có) của mỗi mẫu số liệu sau:

- a) 90;56;50;45;46;48;52;43.
b) 19;11;1;16;19;12;14;10;11.
c) 6,7;6,2;9,7;6,3;6,8;6,1;6,2.
d) 0,79;0,68;0,35;0,38;0,05;0,35.

Lời giải

- a) $S^2 = 202,69; R = 47; \Delta_Q = 8,5$. Giá trị ngoại lệ là 90.
b) $S^2 = 26,91; R = 18; \Delta_Q = 7$. Không có giá trị ngoại lệ.
c) $S^2 = 1,41; R = 3,6; \Delta_Q = 0,6$. Giá trị ngoại lệ là 9,7.
d) $S^2 = 0,059; R = 0,74; \Delta_Q = 0,33$. Không có giá trị ngoại lệ.

Câu 31. Hãy tìm phương sai, khoảng biến thiên, khoảng tứ phân vị và giá trị ngoại lệ (nếu có) của mỗi mẫu số liệu cho bởi bảng tần số sau:

a)

Giá trị	0	4	6	9	10	17
Tần số	1	3	5	4	2	1

b)

Giá trị	2	23	24	25	26	27
Tần số	1	6	8	9	4	2

Lời giải

a) $S^2 = 13,40$; $R = 17$; $\Delta_Q = 4$. Giá trị ngoại lệ là 17.

b) $S^2 = 17,74$; $R = 25$; $\Delta_Q = 1$. Giá trị ngoại lệ là 2 và 27.

Câu 32. Một kĩ thuật viên thống kê lại số lần máy bị lỗi từng ngày trong tháng 5/2021 ở bảng sau:

Số lỗi	0	1	2	3	4	5	6	7	12	15
Số ngày	2	3	4	6	6	3	2	3	1	1

a) Hãy tìm khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu.

b) Xác định các giá trị ngoại lệ (nếu có) của mẫu số liệu.

c) Hãy tìm phương sai và độ lệch chuẩn của mẫu số liệu.

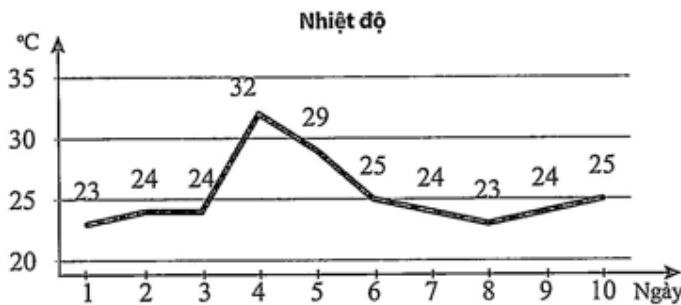
Lời giải

a) $R = 15$; $Q_1 = 2$; $Q_3 = 5$, $\Delta_Q = 3$.

b) Do $Q_1 - 1,5\Delta_Q = -2,5$ và $Q_3 + 1,5\Delta_Q = 9,5$ nên mẫu có 2 giá trị ngoại lệ là 12 và 15.

c) $S^2 = 9,79$; $S = 3,13$.

Câu 33. Biểu đồ sau ghi lại nhiệt độ lúc 12 giờ trưa tại một trạm quan trắc trong 10 ngày liên tiếp (đơn vị: $^{\circ}\text{C}$).



a) Hãy viết mẫu số liệu thống kê nhiệt độ từ biểu đồ trên.

b) Hãy tìm khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu đó.

c) Hãy tìm phương sai và độ lệch chuẩn của mẫu số liệu đó.

Lời giải

a)

Nhiệt độ	23	24	24	32	29	25	24	23	24	25
----------	----	----	----	----	----	----	----	----	----	----

b) $R = 32 - 23 = 9$; $Q_1 = 24$; $Q_3 = 25$; $\Delta_Q = 1$.

c) $S^2 = 7,61$; $S \approx 2,76$.

Câu 34. Khuê và Trọng ghi lại số tin nhắn điện thoại mà mỗi người nhận được từ ngày 1/9 đến ngày 15/9 năm 2020 ở bảng sau:

Khuê	2	4	3	4	6	2	3	2	4	5	3	4	6	7	3
Trọng	3	4	1	2	2	3	4	1	2	30	2	2	2	3	6

a) Hãy tìm phương sai của từng dãy số liệu.

b) Sau khi bỏ đi các giá trị ngoại lệ (nếu có), hãy so sánh số lượng tin nhắn mỗi bạn nhận được theo số trung bình và theo trung vị.

Lời giải

a) Khuê: $S^2 = 2,25$; Trọng: $S^2 = 48,12$.

b) Khuê: $Q_1 = 3$; $Q_3 = 5$; $\Delta_Q = 2$; $Q_1 - 1,5\Delta_Q = 0$; $Q_3 + 1,5\Delta_Q = 8$. Mẫu số liệu của Khuê không có giá trị ngoại lệ.

Trọng: $Q_1 = 2$; $Q_3 = 4$; $\Delta_Q = 2$; $Q_1 - 1,5\Delta_Q = -1$; $Q_3 + 1,5\Delta_Q = 7$. Mẫu số liệu của Trọng có một giá trị ngoại lệ là 30.

Sau khi bỏ đi giá trị ngoại lệ thì số trung bình của mẫu của Khuê và của Trọng lần lượt là 3,87 và 2,64 ; trung vị của mẫu của Khuê và của Trọng lần lượt là 4 và 2. Do đó so sánh theo cả trung bình và trung vị thì Khuê có nhiều tin nhắn mỗi ngày hơn Trọng.

Câu 35. Bảng sau ghi giá bán ra lúc 11 giờ trưa của 2 mã cổ phiếu A và B trong 10 ngày liên tiếp (đơn vị: nghìn đồng).

Ngày	1	2	3	4	5	6	7	8	9	10
A	45	45,1	45,3	35,5	45,6	45,5	45,4	45,5	45,4	45,2
B	47	47,4	47,8	68,4	49	48,8	48,8	48,8	48,6	49,2

- a) Biết có 1 trong 10 ngày trên có sự bất thường trong giá cổ phiếu. Hãy tìm ngày đó và giải thích.
b) Sau khi bỏ đi ngày có giá bất thường, hãy cho biết giá cổ phiếu nào ổn định hơn. Tại sao?

Lời giải

- a) Kiểm tra được giá trị ngoại lệ rơi vào ngày thứ 4.
b) Bỏ đi giá cổ phiếu ngày thứ 4 rồi so sánh phương sai mẫu, ta thấy giá của mã cổ phiếu A ổn định hơn giá của mã cổ phiếu B .

Câu 36. Trong 5 lần nhảy xa, hai bạn Hùng và Trung có kết quả (đơn vị: mét) lần lượt là

Hùng	2,4	2,6	2,4	2,5	2,6
Trung	2,4	2,5	2,5	2,5	2,6

- a. Kết quả trung bình của hai bạn có bằng nhau hay không?
b. Tính phương sai của mẫu số liệu thống kê kết quả 5 lần nhảy xa của mỗi bạn. Từ đó cho biết bạn nào có kết quả nhảy xa ổn định hơn.

Lời giải

- a.
- Kết quả trung bình của Hùng là: $\bar{x} = 2,5$
- Kết quả trung bình của Trung là: $\bar{x} = 2,5$
Vậy kết quả trung bình của hai bạn có bằng nhau.
b. Ta có:

$$s_H^2 = \frac{(2,4 - 2,5)^2 + (2,6 - 2,5)^2 + (2,4 - 2,5)^2 + (2,5 - 2,5)^2 + (2,6 - 2,5)^2}{5} = 0,008$$

$$s_T^2 = \frac{(2,4 - 2,5)^2 + (2,5 - 2,5)^2 + (2,5 - 2,5)^2 + (2,5 - 2,5)^2 + (2,6 - 2,5)^2}{5} = 0,004$$

Bạn Trung có kết quả nhảy xa ổn định hơn.

Câu 37. Biểu đồ đoạn thẳng ở Hình biểu diễn tốc độ tăng trưởng GDP của Việt Nam giai đoạn 2012 - 2019.



- a. Viết mẫu số liệu thống kê tốc độ tăng trưởng GDP nhận được từ biểu đồ ở Hình 3.
b. Tìm khoảng biến thiên của mẫu số liệu đó.
c. Tìm khoảng tứ phân vị của mẫu số liệu đó.
d. Tính phương sai và độ lệch chuẩn của mẫu số liệu đó.

Lời giải

- a. 5,25 5,42 5,98 6,68 6,21 6,81 7,08 7,02
b. Mẫu số liệu được sắp xếp theo thứ tự không giảm là: 5,25 5,42 5,98 6,21 6,68 6,81 7,02 7,08
 $R = 7,08 - 5,25 = 1,83$

c. Trung vị của mẫu số liệu là: $Q_2 = \frac{6,21+6,68}{2} = 6,445$

Trung vị của dãy 5,25 5,42 5,98 6,21 là $Q_1 = \frac{5,42+5,98}{2} = 5,7$

Trung vị của dãy 6,68 6,81 7,02 7,08 là $Q_3 = \frac{6,81+7,02}{2} = 6,915$

$\Rightarrow \Delta_Q = Q_3 - Q_1 = 6,915 - 5,7 = 1,215$

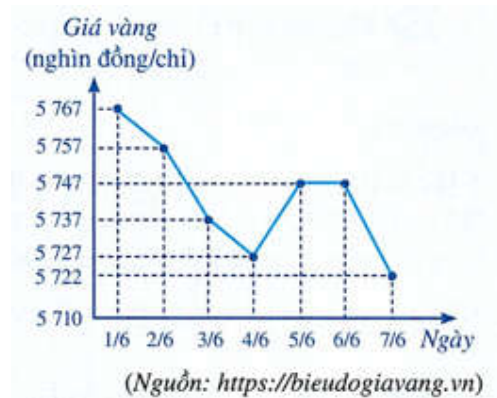
d. Ta có: $\bar{x} = 6,31$

$$s^2 = \frac{(5,25-6,31)^2 + (5,42-6,31)^2 + (5,98-6,31)^2 + (6,68-6,31)^2 + (6,21-6,31)^2 + (6,81-6,31)^2 + (7,08-6,31)^2 + (7,02-6,31)^2}{8}$$

$$= 0,4398125$$

$$\Rightarrow s = \sqrt{s^2} \approx 0,66$$

Câu 38. Biểu đồ đoạn thẳng ở Hình biểu diễn giá vàng bán ra trong bảy ngày đầu tiên của tháng 6 năm 2021.



- Viết mẫu số liệu thống kê giá vàng bán ra nhận được từ biểu đồ ở Hình.
- Tìm khoảng biến thiên của mẫu số liệu đó.
- Tìm khoảng tứ phân vị của mẫu số liệu đó.
- Tính phương sai và độ lệch chuẩn của mẫu số liệu đó.

Lời giải

a. 5767 5757 5737 5727 5747 5747 5722

b. Mẫu số liệu được sắp xếp theo thứ tự không giảm là: 5722 5727 5737 5747 5747 5757 5767

$$R = 5767 - 5722 = 45$$

c.

- Trung vị của mẫu số liệu là: $Q_2 = 5747$

- Trung vị của dãy 572257275737 là $Q_1 = 5727$

- Trung vị của dãy 574757575767 là $Q_3 = 5757$

$$\Rightarrow \Delta_Q = Q_3 - Q_1 = 5757 - 5727 = 30$$

d. Ta có: $\bar{x} \approx 5743,43$

$$s^2 = \frac{(5767-5743,43)^2 + (5757-5743,43)^2 + \dots + (5747-5743,43)^2 + (5722-5743,43)^2}{7}$$

$$\approx 219,39 \Rightarrow s = \sqrt{s^2} \approx 14,81$$

Câu 39. Để biết cây đậu phát triển như thế nào sau khi gieo hạt, bạn Châu gieo 5 hạt đậu vào 5 chậu riêng biệt và cung cấp cho chúng lượng nước, ánh sáng như nhau. Sau hai tuần, 5 hạt đậu đã nảy mầm

và phát triển thành 5 cây con. Bạn Châu đo chiều cao từ rễ đến ngọn của mỗi cây (đơn vị: mi-li-mét) và ghi kết quả là mẫu số liệu sau:

112 102 106 94 101

- Tính phương sai và độ lệch chuẩn của mẫu số liệu trên.
- Theo em, các cây có phát triển đồng đều hay không?

Lời giải

a. Ta có: $\bar{x} = 103$

$$\Rightarrow s^2 = \frac{(112-103)^2 + (102-103)^2 + (106-103)^2 + (94-103)^2 + (101-103)^2}{5} = 35,2 \Rightarrow s \approx 5,94$$

b. Không đồng đều

Câu 40. Cho mẫu số liệu 1 11 13 15 17 21

- Tìm khoảng biến thiên của mẫu số liệu trên.
- Tìm khoảng tứ phân vị của mẫu số liệu trên.
- Tính phương sai và độ lệch chuẩn của mẫu số liệu trên.
- Tìm giá trị bất thường của mẫu số liệu trên.

Lời giải

a) 20. b) 6. c) Phương sai là $\frac{116}{3}$, độ lệch chuẩn là $\frac{2\sqrt{87}}{3}$. d) 1.

Câu 41. Kết quả dự báo nhiệt độ cao nhất trong 10 ngày liên tiếp ở Nghệ An cuối tháng 01 năm 2022 được cho ở bảng sau:

Ngày	22	23	24	25	26	27	28	29	30	31
Nhiệt độ ($^{\circ}\text{C}$)	23	25	26	27	27	27	27	21	19	18

(Nguồn: <https://nchmf.gov.vn>)

- Viết mẫu số liệu thống kê nhiệt độ nhận được từ bảng trên.
- Tính số trung bình cộng, phương sai và độ lệch chuẩn của mẫu số liệu đó.

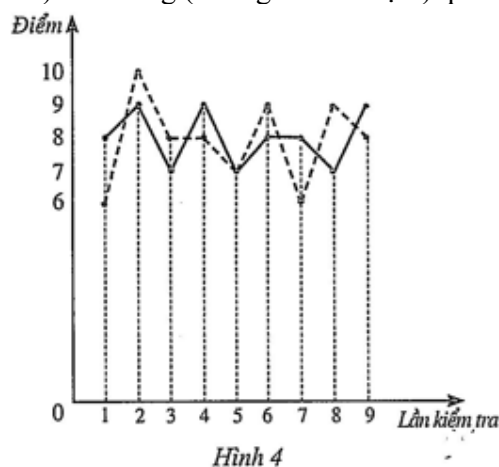
Lời giải

a) 23 25 26 27 27 27 27 21 19 18

b) Số trung bình cộng là $24(^{\circ}\text{C})$.

Phương sai là 11,2. Độ lệch chuẩn là $\frac{2\sqrt{70}}{5}(^{\circ}\text{C})$.

Câu 42. Biểu đồ đoạn thẳng ở Hình 4 cho biết kết quả thi Ngoại ngữ ở câu lạc bộ của Dũng (đường nét liền) và Hoàng (đường nét đứt đậm) qua 9 lần kiểm tra.



- Viết mẫu số liệu thống kê kết quả thi ngoại ngữ của Dũng và Hoàng nhận được từ biểu đồ ở Hình 4.
- Tìm khoảng biến thiên và khoảng tứ phân vị của mỗi mẫu số liệu đó.
- Tính phương sai và độ lệch chuẩn của hai mẫu số liệu đó. Cho biết kết quả thi của bạn nào ổn định hơn?

Lời giải

a) Mẫu số liệu kết quả thi của bạn Dũng là: 8 9 7 9 7 8 8 7 9 (1)

Mẫu số liệu kết quả thi của bạn Hoàng là: 6 10 8 8 7 9 6 9 8 (2)

b) Khoảng biến thiên của mẫu số liệu (1) và (2) lần lượt là 2 và 4.

Khoảng tứ phân vị của mẫu số liệu (1) và (2) lần lượt là 2 và 2,5.

c) Phương sai của mẫu số liệu (1) và (2) lần lượt là $\frac{2}{3}$ và $\frac{134}{81}$.

Độ lệch chuẩn của mẫu số liệu (1) và (2) lần lượt là $\frac{\sqrt{6}}{3}$ và $\frac{\sqrt{134}}{9}$.

Ta có: $\frac{2}{3} < \frac{134}{81}$ nên kết quả thi của bạn Dũng ổn định hơn.

BÀI TẬP BỔ SUNG

Câu 43. Mẫu số liệu sau đây cho biết sản lượng lúa (đv tạ) của 5 thửa ruộng thí nghiệm có cùng diện tích 2021222324

a) Tìm sản lượng trung bình

b) Tìm phương sai và độ lệch chuẩn.

c) Tìm khoảng biến thiên, khoảng tứ phân vị

Lời giải

a) Số trung bình của mẫu số liệu là: $\bar{x} = \frac{20 + 21 + 22 + 23 + 24}{5} = 22$.

b) Ta có bảng sau:

Giá trị	Độ lệch	Bình phương độ lệch
20	$20 - 22 = -2$	4
21	$21 - 22 = -1$	1
22	$22 - 22 = 0$	0
23	$23 - 22 = 1$	1
24	$24 - 22 = 2$	4
	Tổng	10

Mẫu số liệu gồm 5 giá trị nên $n = 5$. Do đó phương sai là: $s^2 = \frac{10}{5} = 2$.

Độ lệch chuẩn là: $s = \sqrt{2} \approx 1,41$.

c) Khoảng biến thiên bằng $24 - 20 = 4$

Khoảng tứ phân vị $23,5 - 20,5 = 3$

C. BÀI TẬP TRẮC NGHIỆM

BÀI TẬP SÁCH GIÁO KHOA, SÁCH BÀI TẬP

Câu 1. Cho mẫu số liệu: 21 22 23 24 25

a) Khoảng biến thiên của mẫu số liệu trên là:

A. 1. B. 2. C. 3. D. 4.

b) Khoảng tứ phân vị của mẫu số liệu trên là:

A. 1. B. 2. C. 3. D. 4.

c) Phương sai của mẫu số liệu trên là:

A. 1. B. 2. C. 3. D. 4.

d) Độ lệch chuẩn của mẫu số liệu trên là:

A. 1. B. $\sqrt{2}$. C. $\sqrt{3}$. D. 4.

Lời giải

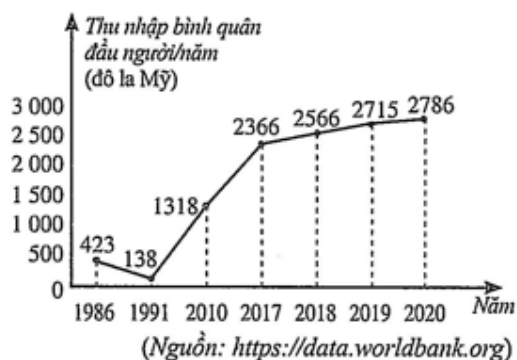
a) D.

b) C.

c) B.

d) B.

Câu 2. Biểu đồ đoạn thẳng ở Hình 2 biểu diễn thu nhập bình quân đầu người/năm của Việt Nam ở một số năm trong giai đoạn từ 1986 đến 2020. Mẫu số liệu nhận được từ biểu đồ ở Hình 2 có khoảng biến thiên là bao nhiêu?



Hình 2

A. 71.

B. 85.

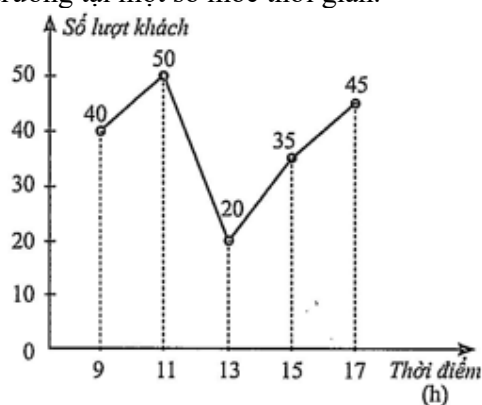
C. 1180.

D. 2648.

Lời giải

Chọn D

Câu 3. Biểu đồ đoạn thẳng ở hình 3 biểu diễn số lượt khách vào một cửa hàng trong ngày đầu khai trương tại một số mốc thời gian.



Hình 3

Mẫu số liệu nhận được từ biểu đồ ở Hình 3 có khoảng tứ phân vị là bao nhiêu?

A. 10.

B. 15.

C. 20.

D. 5.

Lời giải

Chọn C

BÀI TẬP BỔ SUNG

Câu 4. Năng suất lúa hè thu (tạ/ha) năm 1998 của 31 tỉnh ở Việt Nam được thống kê trong bảng sau:

Năng suất lúa (tạ/ha)	25	30	35	40	45
Tần số	4	7	9	6	5

Giá trị $x_3 = 35$ có tần số bằng

A. 6.

B. 4.

C. 7.

D. 9.

Lời giải

Chọn D

Câu 5. Chọn khẳng định **sai** trong các khẳng định sau:

A. Phương sai luôn là một số không âm.

B. Phương sai là bình phương của độ lệch chuẩn.

C. Phương sai càng lớn thì độ phân tán quanh số trung bình càng lớn.

D. Phương sai luôn lớn hơn độ lệch chuẩn.

Lời giải

Chọn D

♦ Phương sai S_x^2 còn độ lệch chuẩn $S_x = \sqrt{S_x^2}$ nhưng không thể khẳng định phương sai luôn lớn hơn độ lệch chuẩn.

Câu 6. Để đánh giá mức độ phân tán của các số liệu thống kê so với số trung bình, ta dùng đại lượng nào sau đây?

- A.** Số trung bình. **B.** Số trung vị **C.** Mốt. **D.** Phương sai.

Lời giải

Chọn D

♦ Dựa vào ý nghĩa của phương sai và độ lệch chuẩn để đo mức độ phân tán của các số liệu trong mẫu quanh số trung bình.

Câu 7. Chọn câu đúng trong các câu trả lời sau đây: Phương sai bằng:

- A.** Một nửa của độ lệch chuẩn **B.** Căn bậc hai của độ lệch chuẩn.
C. Hai lần của độ lệch chuẩn. **D.** Bình phương của độ lệch chuẩn.

Lời giải

Chọn D

Ta có phương sai là: s_x^2

Độ lệch chuẩn: $s_x = \sqrt{s_x^2}$

Suy ra phương sai bằng bình phương của độ lệch chuẩn

Câu 8. Cho phương sai của các số liệu bằng 4. Tìm độ lệch chuẩn.

- A.** 4. **B.** 2. **C.** 16. **D.** 8.

Lời giải

Ta có độ lệch chuẩn là căn bậc hai của phương sai

Nên $s_x = \sqrt{s_x^2} = \sqrt{4} = 2$.

Câu 9. Độ lệch chuẩn là

- A.** Căn bậc hai của phương sai. **B.** Bình phương của phương sai.
C. Một nửa của phương sai. **D.** Không phải các công thức trên.

Lời giải

Chọn A

Câu 10. Nếu đơn vị đo của số liệu là kg thì đơn vị của độ lệch chuẩn là

- A.** kg. **B.** kg^2 . **C.** Không có đơn vị. **D.** $\frac{\text{kg}}{2}$.

Lời giải

Chọn A

Câu 11. Tìm phát biểu đúng về phương sai của một mẫu số liệu.

- A.** Phương sai được sử dụng làm đại diện cho các số liệu của mẫu.
B. Phương sai được sử dụng để đánh giá mức độ phân tán của các số liệu thống kê (so với số trung bình).
C. Phương sai được tính bằng tổng số phần tử của một mẫu số liệu.
D. Phương sai là số liệu xuất hiện nhiều nhất (số liệu có tần số lớn nhất) trong bảng các số liệu thống kê.

Lời giải

Ý nghĩa của phương sai: Phương sai được sử dụng để đánh giá mức độ phân tán của các số liệu thống kê (so với số trung bình). (SGK)

- Câu 12.** Theo kết quả thống kê điểm thi giữa kỳ 2 môn toán khối 11 của một trường THPT, người ta tính được phương sai của bảng thống kê đó là $s_x^2 = 0,573$. Độ lệch chuẩn của bảng thống kê đó bằng
- A.** 0,812. **B.** 0,757. **C.** 0,936. **D.** 0,657.

Lời giải

Ta có công thức tính độ lệch chuẩn là $s_x = \sqrt{s_x^2} = \sqrt{0,573} \approx 0,757$.

- Câu 13.** Cho mẫu số liệu x_1, x_2, \dots, x_N có số trung bình là \bar{x} . Phương sai được tính theo công thức nào trong các công thức sau

A. $\frac{1}{N} \sum_{i=1}^N x_i$. **B.** $\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})}$. **C.** $\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$. **D.** $\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$.

Lời giải

Phương sai được tính theo công thức $s^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$ hoặc $s^2 = \frac{1}{N} \sum_{i=1}^N x_i^2 - \frac{1}{N^2} \left(\sum_{i=1}^N x_i \right)^2$.

- Câu 14.** Phương sai của dãy số 2;3;4;5;6 là

A. $S_x^2 = 4$. **B.** $S_x^2 = \sqrt{2}$. **C.** $S_x^2 = 2$. **D.** $S_x^2 = -2$.

Lời giải

Chọn C

♦ Ta có: $\bar{x} = \frac{2+3+4+5+6}{5} = 4$.

♦ Suy ra: $S_x^2 = \frac{1}{5} \left[(2-4)^2 + (3-4)^2 + (5-4)^2 + (6-4)^2 \right] = 2$.

- Câu 15.** Khoảng tứ phân vị của dãy số 2;3;4;5;6 là

A. $\Delta_Q = 3$. **B.** $\Delta_Q = \sqrt{2}$. **C.** $\Delta_Q = 2$. **D.** $\Delta_Q = -2$.

Lời giải

Chọn A

$\Delta_Q = Q_3 - Q_1 = \frac{11}{2} - \frac{5}{2} = 3$

- Câu 16.** Thống kê điểm kiểm tra môn toán (thang điểm 10) của một nhóm gồm 6 học sinh ta có bảng số liệu sau:

Tên học sinh	Kim	Son	Ninh	Bình	Viết	Nam
Điểm	9	8	7	10	8	9

Tìm độ lệch chuẩn δ của bảng số liệu trên (làm tròn đến hàng phần trăm).

A. $\delta \approx 0,92$. **B.** $\delta \approx 0,95$. **C.** $\delta \approx 0,96$. **D.** $\delta \approx 0,91$.

Lời giải

Chọn C

♦ Ta có: $\bar{x} = \frac{9+8+7+10+8+9}{6} = \frac{51}{6} = 8,5$.

♦ Suy ra: $\delta^2 = \frac{1}{6} \left(2(9-8,5)^2 + 2(8-8,5)^2 + (7-8,5)^2 + (10-8,5)^2 \right) = \frac{11}{12}$.

♦ Do đó $\delta = \sqrt{\frac{11}{12}} \approx 0,96$.

Câu 17. Có 100 học sinh tham dự kì thi học sinh giỏi Toán (thang điểm 20). Kết quả cho trong bảng sau:

Điểm (x)	9	10	11	12	13	14	15	16	17	18	19
Tần số	1	1	3	5	8	13	19	24	14	10	2

Khi đó **độ lệch chuẩn** là

- A.** 1,98. **B.** 3,96. **C.** 15,23 **D.** 1,99.

Lời giải

Chọn D

Ta có:

$$\bar{x} = \frac{1.9 + 1.10 + 3.11 + 5.12 + 8.13 + 13.14 + 19.15 + 24.16 + 14.17 + 10.18 + 2.19}{100} = 15,23$$

$$\overline{x^2} = \frac{1.9^2 + 1.10^2 + 3.11^2 + 5.12^2 + 8.13^2 + 13.14^2 + 19.15^2 + 24.16^2 + 14.17^2 + 10.18^2 + 2.19^2}{100} = 235,91$$

Phương sai của bảng số liệu là: $s^2 = \overline{x^2} - (\bar{x})^2 = 235,91 - 15,23^2 = 3,9571$.

Độ lệch chuẩn là: $s = \sqrt{s^2} = \sqrt{3,9571} = 1,99$.

Câu 18. Điểm thi của lớp 10C của một trường Trung học Phổ Thông được trình bày ở bảng phân bố tần số sau:

Điểm thi	5	6	7	8	9	10	
Tần số	7	5	10	12	4	2	$n = 40$

Phương sai của bảng phân bố tần số đã cho là:

- A.** 0,94 **B.** 3,94. **C.** 2,94. **D.** 1,94.

Lời giải

Chọn D

Trong dãy số liệu về điểm thi của lớp 10C ta có:

$$\bar{x} = \frac{1}{n} \cdot (n_1x_1 + n_2x_2 + \dots + n_6x_6) = \frac{1}{40} \cdot (7.5 + 5.6 + 10.7 + 12.8 + 4.9 + 2.10) = 7,175$$

Phương sai:

$$\begin{aligned} s^2 &= \frac{1}{n} \cdot \left(n_1 \cdot (x_1 - \bar{x})^2 + n_2 \cdot (x_2 - \bar{x})^2 + \dots + n_6 \cdot (x_6 - \bar{x})^2 \right) \\ &= \frac{1}{40} \cdot \left(7 \cdot (5 - 7,175)^2 + 5 \cdot (6 - 7,175)^2 + 10 \cdot (7 - 7,175)^2 \right. \\ &\quad \left. + 12 \cdot (8 - 7,175)^2 + 4 \cdot (9 - 7,175)^2 + 2 \cdot (10 - 7,175)^2 \right) \\ &\approx 1,94 \end{aligned}$$

Câu 19. Theo dõi thời gian làm một bài toán (tính bằng phút) của 40 học sinh, giáo viên lập được bảng sau:

Thời gian (x)	4	5	6	7	8	9	10	11	12	
Tần số (n)	6	3	4	2	7	5	5	7	1	$N = 40$

Phương sai của mẫu số liệu trên gần với số nào nhất?

- A.** 6. **B.** 12. **C.** 40. **D.** 9.

Lời giải

Ta có giá trị trung bình của mẫu số liệu là $\bar{x} = \frac{x_1.n_1 + x_2.n_2 + \dots + x_k.n_k}{N} = \frac{317}{40}$.

Phương sai của mẫu số liệu là $s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{N} = 6$.

- Câu 20.** Cho dãy số liệu thống kê: 1; 2; 3; 4; 5; 6; 7. Phương sai của các số liệu thống kê là
A. 1. **B.** 2. **C.** 3. **D.** 4.

Lời giải

Giá trị trung bình của dãy số liệu thống kê đã cho là: $\bar{x} = \frac{1+2+3+4+5+6+7}{7} = 4$.

Phương sai của các số liệu thống kê là

$$S_x^2 = \frac{(\bar{x}-1)^2 + (\bar{x}-2)^2 + (\bar{x}-3)^2 + (\bar{x}-4)^2 + (\bar{x}-5)^2 + (\bar{x}-6)^2 + (\bar{x}-7)^2}{7}$$

$$= \frac{(4-1)^2 + (4-2)^2 + (4-3)^2 + (4-4)^2 + (4-5)^2 + (4-6)^2 + (4-7)^2}{7} = \frac{28}{7} = 4.$$

- Câu 21.** Cho dãy số liệu thống kê: 1; 2; 3; 4; 5; 6; 7. Khoảng biến thiên là
A. 1. **B.** 2. **C.** 3. **D.** 6.

Lời giải

Chọn D

$$7-1=6$$

- Câu 22.** Số liệu thống kê 100 học sinh tham gia kì thi học sinh giỏi toán (thang điểm 20). 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.

- A.** 2,01. **B.** 1,89. **C.** 1,98. **D.** 1,99.

Lời giải

Điểm số trung bình của các học sinh tham gia thi học sinh giỏi là

$$\bar{x} = \frac{1.9 + 1.10 + 3.11 + 5.12 + 8.13 + 13.14 + 19.15 + 24.16 + 14.17 + 10.18 + 2.19}{100} \approx 15,23.$$

Phương sai của số liệu thống kê là

$$S_x^2 = \frac{(\bar{x}-9)^2 + (\bar{x}-10)^2 + 3(\bar{x}-11)^2 + 5(\bar{x}-12)^2 + \dots + 2(\bar{x}-19)^2}{100} \approx 3,96.$$

Suy ra độ lệch chuẩn của bảng số liệu thống kê là $S_x = \sqrt{S_x^2} \approx 1,99$

- Câu 23.** Cho mẫu số liệu thống kê $\{1; 2; 3; 4; 5; 6; 7; 8; 9\}$. Tính (gần đúng) độ lệch chuẩn của mẫu số liệu trên?

- A.** 2,45. **B.** 2,58. **C.** 6,67. **D.** 6,0.

Lời giải

Ta có giá trị trung bình $\bar{x} = \frac{1+2+3+4+5+6+7+8+9}{9} = 5$.

Do đó độ lệch chuẩn

$$s = \sqrt{\frac{(1-5)^2 + (2-5)^2 + (3-5)^2 + (4-5)^2 + (5-5)^2 + (6-5)^2 + (7-5)^2 + (8-5)^2 + (9-5)^2}{9}}$$

$$s = \frac{2\sqrt{15}}{3} \approx 2,58.$$

- Câu 24.** Cho mẫu số liệu thống kê $\{1; 2; 3; 4; 5; 6; 7; 8; 9\}$. Tìm khoảng tứ phân vị của mẫu số liệu trên?

A. 2.

B. 5.

C. 3.

D. 4.

Lời giải

Ta có $Q_1 = 2,5, Q_3 = 7,5 \Rightarrow \Delta_Q = 5$

Câu 25. Một cửa hàng bán gạo, thống kê số kg gạo mà cửa hàng bán mỗi ngày trong 30 ngày, được bảng tần số:

Bảng tần số	
Số kg gạo	Tần số
100	7
120	4
130	2
160	8
180	3
200	2
250	4
Tổng	30

Phương sai của bảng số liệu gần đúng với giá trị nào dưới đây nhất?

A. 155.

B. 2318.

C. 3325.

D. 1234.

Lời giải

Ta có số trung bình của bảng số liệu là:

$$\bar{x} = \frac{7 \cdot 100 + 4 \cdot 120 + 2 \cdot 130 + 8 \cdot 160 + 3 \cdot 180 + 2 \cdot 200 + 4 \cdot 250}{30} \approx 155$$

Phương sai của bảng số liệu:

$$s_x^2 \approx \frac{7(100-155)^2 + 4(120-155)^2 + 2(130-155)^2 + 8(160-155)^2 + 3(180-155)^2 + 2(200-155)^2 + 4(250-155)^2}{30} \approx 2318.$$

Câu 26. Sản lượng lúa (tạ) của 40 thửa ruộng thí nghiệm có cùng diện tích được trình bày trong bảng phân bố tần số sau đây:

Sản lượng	20	21	22	23	24
Tần số	5	8	11	10	6

Phương sai của mẫu số liệu là:

A. $s_x^2 = 1,5$

B. $s_x^2 = 1,24$

C. 1,54

D. 22,1

Lời giải

Ta có sản lượng trung bình của 40 thửa ruộng là:

$$\bar{x} = \frac{1}{40}(5 \cdot 20 + 8 \cdot 21 + 11 \cdot 22 + 10 \cdot 23 + 6 \cdot 24) = 22,1 \text{ (tạ)}$$

Phương sai:

$$s_x^2 = \frac{1}{n} [n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + \dots + n_k(x_k - \bar{x})^2] = 1,54$$

Câu 27. Điểm kiểm tra giữa kỳ 2 của một học sinh lớp 10 như sau: 2, 4, 6, 8, 10. Phương sai của mẫu số liệu trên là bao nhiêu?

A. 6

B. 8

C. 10

D. 40

Lời giải

Chọn B

$$\bar{x} = \frac{2+4+6+8+10}{5} = 6.$$

$$s^2 = \frac{1}{5} \sum_{i=1}^5 (x_i - \bar{x})^2 = 8.$$

Câu 28. Điểm kiểm tra giữa kỳ 2 của một học sinh lớp 10 như sau: 2, 4, 6, 8, 10. Khoảng biến thiên của mẫu số liệu trên là bao nhiêu?

A. 6

B. 8

C. 10

D. 40

Lời giải

Chọn B

Câu 29. Cho thống kê điểm thi môn toán trong một kì thi của 400 em học sinh. Người ta thấy có 72 bài được điểm 5. Hỏi tần suất của giá trị $x_i = 5$ là bao nhiêu

A. 72%.

B. 36%.

C. 10%.

D. 18%.

Lời giải

Ta có tần số của giá trị x_i là $n_i = 72$, suy ra tần suất của giá trị x_i là: $f_i = \frac{n_i}{N} = \frac{72}{400} = 18\%$

Vậy $f_i = 18\%$.

Câu 30. Cho bảng số liệu điểm thi học kì 2 của 40 học sinh lớp 10A (thang điểm là 10):

Điểm	5	6	7	8	9	10	
Tần số	5	12	8	9	4	2	N=40

Tính phương sai S_x^2

A. $S_x^2 = 1,784$.B. $S_x^2 = 1,874$.C. $S_x^2 = 1,847$.D. $S_x^2 = 1,748$.

Lời giải

Ta có điểm trung bình của 40 em học sinh là:

$$\bar{x} = \frac{5 \times 5 + 12 \times 6 + 8 \times 7 + 9 \times 8 + 4 \times 9 + 2 \times 10}{40} = \frac{281}{40} = 7,025$$

$$S_x^2 = \frac{5(5-7,025)^2 + 12(6-7,025)^2 + 8(7-7,025)^2 + 9(8-7,025)^2 + 4(9-7,025)^2 + 2(10-7,025)^2}{40}$$

$$= 1,874$$

Câu 31. Điểm thi môn Toán lớp 10A₂ của một Trường trung học phổ thông được trình bày ở bảng phân bố tần số sau

Điểm thi	5	6	7	8	9	10	
Tần số	7	5	10	12	4	2	$n = 40$

Trong các giá trị dưới đây, giá trị nào gần nhất với phương sai của bảng phân bố tần số trên?

A. 0,94.

B. 3,94.

C. 2,94.

D. 1,94.

Lời giải

Trong dãy số liệu về điểm thi môn Toán lớp 10A₂ ta có

$$\bar{x} = \frac{1}{n} \cdot (n_1 x_1 + n_2 x_2 + \dots + n_6 x_6) = \frac{1}{40} \cdot (7 \cdot 5 + 5 \cdot 6 + 10 \cdot 7 + 12 \cdot 8 + 4 \cdot 9 + 2 \cdot 10) = 7,175.$$

Phương sai là:

$$\begin{aligned}s^2 &= \frac{1}{n} \cdot \left(n_1 \cdot (x_1 - \bar{x})^2 + n_2 \cdot (x_2 - \bar{x})^2 + \dots + n_6 \cdot (x_6 - \bar{x})^2 \right) \\&= \frac{1}{40} \cdot \left(7 \cdot (5 - 7,175)^2 + 5 \cdot (6 - 7,175)^2 + 10 \cdot (7 - 7,175)^2 \right. \\&\quad \left. + 12 \cdot (8 - 7,175)^2 + 4 \cdot (9 - 7,175)^2 + 2 \cdot (10 - 7,175)^2 \right) \\&\approx 1,94.\end{aligned}$$

Nguyễn Bảo Vương