CÁC DẠNG BÀI TẬP BÀI PHƯƠNG SAI VÀ ĐỘ LỆCH CHUẨN CỦA MẪU SỐ LIỆU GHÉP NHÓM

Phương pháp tính phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm

Xét mẫu số liệu ghép nhóm cho bởi công thức sau:

Nhóm	$[u_1;u_2)$	$[u_1;u_2)$	$\cdots$	$[u_k;u_{k+1})$
Giá trị đại diện	c_1	c_2	$\cdots$	c_k
Tần số	n_1	n_2	$\cdots$	n_k

Số trung bình cộng của mẫu số liệu ghép nhóm trên được tính bằng công thức:

$$\bar{x}=\frac{1}{n}\left(n_1c_1+n_2c_2+\ldots+n_kc_k\right)$$

Phương sai: Phương sai của mẫu số liệu ghép nhóm, kí hiệu S^2, được tính bởi công thức

$$S^2=\frac{1}{n}\left[n_1\left(c_1-\bar{x}\right)^2+n_2\left(c_2-\bar{x}\right)^2+\ldots+n_k\left(c_k-\bar{x}\right)^2\right]$$

Trong đó: $n=n_1+n_2+\ldots+n_k$ là cỡ mẫu; $\bar{x}=\frac{1}{n}\left(n_1c_1+n_2c_2+\ldots+n_kc_k\right)$

Độ lệch chuẩn: Độ lệch chuẩn của mẫu số liệu ghép nhóm, kí hiệu S, là căn bậc hai số học của phương sai, nghĩa là $S=\sqrt{S^2}$

A. BÀI TẬP TRẮC NGHIỆM

PHẦN I. Câu trắc nghiệm nhiều phương án lựa chọn. Mỗi câu hỏi thí sinh chỉ chọn một phương án.

Câu 1: Số đặc trưng nào không sử dụng thông tin của nhóm số liệu đầu tiên và nhóm số liệu cuối cùng?

A. Khoảng biến thiên. **B.** Khoảng tứ phân vị.

C. Phương sai. **D.** Độ lệch chuẩn.

Lời giải

Số đặc trưng không sử dụng thông tin của nhóm số liệu đầu tiên và nhóm số liệu cuối cùng là khoảng tứ phân vị.

Câu 2: Nếu thay tất cả các tần số trong mẫu số liệu ghép nhóm trên bằng 4 thì số đặc trưng nào sau đây không thay đổi?

A. Khoảng biến thiên. **B.** Khoảng tứ phân vị.

C. Phương sai. **D.** Độ lệch chuẩn.

Lời giải

Nếu thay tất cả các tần số trong mẫu số liệu ghép nhóm trên bằng 4 thì số đặc trưng không đổi là khoảng biến thiên.

Câu 3: Trong các khẳng định sau, khẳng định nào sai?

A. Phương sai luôn luôn là số không âm.

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

C. Phương sai càng lớn thì độ phân tác của các giá trị quanh số trung bình càng lớn.

D. Phương sai luôn luôn lớn hơn độ lệch chuẩn.

Lời giải

Khi $x \in (0;1)$ thì $s^2 < s$ nên khẳng định phương sai luôn lớn hơn độ lệch chuẩn là sai.

Câu 4: Doanh thu bán hàng trong 20 ngày được lựa chọn ngẫu nhiên của một của hàng được ghi lại ở bảng sau (đơn vị: triệu đồng):

Doanh thu	[5;7)	[7;9)	[9;11)	[11;13)	[13;15)
Số ngày	2	7	7	3	1

Số trung bình của mẫu số liệu trên thuộc khoảng nào trong các khoảng dưới đây?

A. $[7;9)$. **B.** $[9;11)$. **C.** $[11;13)$. **D.** $[13;15)$.

Lời giải

Bảng tần số ghép nhóm theo giá trị đại diện là:

Doanh thu	[5;7)	[7;9)	[9;11)	[11;13)	[13;15)
Giá trị đại diện	6	8	10	12	14
Số ngày	2	7	7	3	1

Số trung bình: $\overline{x} = \dfrac{2.6 + 7.8 + 7.10 + 3.12 + 1.14}{20} = 9,4$.

Câu 5: Một siêu thị thống kê số tiền (đơn vị: chục nghìn đồng) mà 44 khách hàng mua hàng ở siêu thị đó trong một ngày. Số liệu được ghi lại trong bảng dưới đây.

Nhóm	**Giá trị đại diện**	**Tần số**
[40;45)	42,5	4
[45;50)	47,5	14
[50;55)	52,5	8
[55;60)	57,5	10
[60;75)	62,5	6
[65;70)	67,5	2
		$n = 44$

Phương sai của mẫu số liệu ghép nhóm trên là:

A. 53,2. **B.** 46,1. **C.** 30. **D.** 11.

Lời giải

Số trung bình cộng của mẫu số liệu ghép nhóm là:

$$\overline{x} = \frac{4.42,5 + 14.47,5 + 8.52,5 + 10.57,5 + 6.62,5 + 2.67,5}{44} = \frac{585}{11}$$

Phương sai của mẫu số liệu ghép nhóm là:

$$s^2=\frac{4\left(42,5-\frac{585}{11}\right)^2+14\left(47,5-\frac{585}{11}\right)^2+8\left(52,5-\frac{585}{11}\right)^2+10\left(57,5-\frac{585}{11}\right)^2}{44}$$
$$+\frac{+6\left(62,5-\frac{585}{11}\right)^2+2.\left(67,5-\frac{585}{11}\right)^2}{44}\approx 46,12$$

Câu 6: Một siêu thị thống kê số tiền (đơn vị: chục nghìn đồng) mà 44 khách hàng mua hàng ở siêu thị đó trong một ngày. Số liệu được ghi lại trong bảng dưới đây.

Nhóm	Giá trị đại diện	Tần số
$[40;45)$	42,5	4
$[45;50)$	47,5	14
$[50;55)$	52,5	8
$[55;60)$	57,5	10
$[60;75)$	62,5	6
$[65;70)$	67,5	2
		$n=44$

Độ lệch chuẩn của mẫu số liệu ghép nhóm trên (làm tròn kết quả đến hàng đơn vị) là:

A. 6,8. **B.** 7,3. **C.** 3,3. **D.** 46,1.

Lời giải

Số trung bình cộng của mẫu số liệu ghép nhóm là:

$$\overline{x}=\frac{4.42,5+14.47,5+8.52,5+10.57,5+6.62,5+2.67,5}{44}=\frac{585}{11}$$

Phương sai của mẫu số liệu ghép nhóm là:

$$s^2=\frac{4\left(42,5-\frac{585}{11}\right)^2+14\left(47,5-\frac{585}{11}\right)^2+8\left(52,5-\frac{585}{11}\right)^2+10\left(57,5-\frac{585}{11}\right)^2}{44}$$
$$+\frac{6\left(62,5-\frac{585}{11}\right)^2+2.\left(67,5-\frac{585}{11}\right)^2}{44}\approx 46,12$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s=\sqrt{s^2}=\sqrt{46,12}\approx 6,8$

Câu 7: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khoẻ. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

Quãng đường (km)	$[2,7;3,0)$	$[3,0;3,3)$	$[3,3;3,6)$	$[3,6;3,9)$	$[3,9;4,2)$
Số ngày	3	6	5	4	2

Phương sai của mẫu số liệu ghép nhóm là

A. 3,39. **B.** 11,62. **C.** 0,1314. **D.** 0,36.

Lời giải

Giá trị đại diện	2,85	3,15	3,45	3,75	4,05
Số ngày	3	6	5	4	2

Số trung bình: $\overline{x} = \frac{3.2,85 + 6.3,15 + 5.3,45 + 4.3,75 + 2.4,05}{20} = 3,39$

Phương sai: $S^2 = \frac{3.2,85^2 + 6.3,15^2 + 5.3,45^2 + 4.3,75^2 + 2.4,05^2}{20} - 3,39^2 = 0,1314$

Câu 8: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khoẻ. Quãng đường đi bộ mỗi ngày (đơn vị: km) của bác Hương trong 20 ngày được thống kê lại ở bảng sau:

Quãng đường (km)	[2,7;3,0)	[3,0;3,3)	[3,3;3,6)	[3,6;3,9)	[3,9;4,2)
Số ngày	3	6	5	4	2

Độ lệch chuẩn của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

A. 3,41. **B.** 11,62. **C.** 0,017. **D.** 0,36.

Lời giải

Giá trị đại diện	2,85	3,15	3,45	3,75	4,05
Số ngày	3	6	5	4	2

Số trung bình: $\overline{x} = \frac{3.2,85 + 6.3,15 + 5.3,45 + 4.3,75 + 2.4,05}{20} = 3,39$

Phương sai: $S^2 = \frac{3.2,85^2 + 6.3,15^2 + 5.3,45^2 + 4.3,75^2 + 2.4,05^2}{20} - 3,39^2 = 0,1314$

Độ lệch chuẩn: $\sigma = \sqrt{0,1314} \approx 0,36$

Câu 9: Bạn Chi rất thích nhảy hiện đại. Thời gian tập nhảy mỗi ngày trong thời gian gần đây của bạn Chi được thống kê lại ở bảng sau:

Thời gian (phút)	[20;25)	[25;30)	[30;35)	[35;40)	[40;45)
Số ngày	6	6	4	1	1

Phương sai của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

A. 31,77. **B.** 32. **C.** 31. **D.** 31,44.

Lời giải

Số trung bình: $\overline{x} = \frac{6.22,5 + 6.27,5 + 4.32,5 + 37,5 + 42,5}{18} \approx 28,33$

Phương sai: $S^2 = \frac{6.22,5^2 + 6.27,5^2 + 4.32,5^2 + 37,5^2 + 42,5^2}{18} - 28,33^2 = 31,25$

Câu 10: Dũng là học sinh rất giỏi chơi rubik, bạn có thể giải nhiều loại khối rubik khác nhau. Trong một lần tập luyện giải khối rubik 3×3, bạn Dũng đã tự thống kê lại thời gian giải rubik trong 25 lần giải liên tiếp ở bảng sau:

Thời gian giải rubik (giây)	[8;10)	[10;12)	[12;14)	[14;16)	[16;18)
Số lần	4	6	8	4	3

Độ lệch chuẩn của mẫu số liệu ghép nhóm có giá trị gần nhất với giá trị nào dưới đây?

A. 5,98. **B.** 6. **C.** 2,44. **D.** 2,5.

Lời giải

Giá trị đại diện	9	11	13	15	17
Số lần	4	6	8	4	3

Số trung bình: $\overline{x}=\dfrac{4.9+6.11+8.13+4.15+3.17}{25}=12,68$

Độ lệch chuẩn: $\sigma=\sqrt{\dfrac{4.9^2+6.11^2+8.13^2+4.15^2+3.17^2}{25}-12,68^2}\approx 5,98$.

Câu 11: Để đánh giá chất lượng của một loại pin điện thoại mới, người ta ghi lại thời gian nghe nhạc liên tục của điện thoại được sạc đầy pin cho đến khi hết pin cho kết quả sau:

Thời gian (giờ)	[5;5,5)	[5,5;6)	[6;6,5)	[6,5;7)	[7;7,5)
Số chiếc điện thoại (tần số)	2	8	15	10	5

Tính độ lệch chuẩn của mẫu số liệu ghép nhóm trên (làm tròn đến 4 chữ số thập phân)

A. $0,4252$. **B.** $0,5314$. **C.** $0,6214$. **D.** $0,5268$.

Lời giải

Thời gian (giờ)	[5;5,5)	[5,5;6)	[6;6,5)	[6,5;7)	[7;7,5)
Giá trị đại diện	5,25	5,75	6,25	6,75	7,25
Số chiếc điện thoại (tần số)	2	8	15	10	5

Số trung bình của mẫu số liệu:

$$\overline{x}=\frac{m_1.x_1+\ldots+m_k.x_k}{n}=\frac{2.5,25+8.5,75+15.6,25++10.6,75+5.7,25}{40}=6,35$$

Phương sai của mẫu số liệu ghép nhóm:

$$s^2=\frac{1}{40}\left(2.5,25^2+8.5,75^2+15.6,25^2+10.6,75^2+5.7,25^2\right)-6,35^2=0,2775$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s=\sqrt{s^2}=\sqrt{0,2775}\approx 0,5268$

Câu 12: Người ta ghi lại tiền lãi (đơn vị: triệu đồng) của một số nhà đầu tư (với số tiền đầu tư như nhau), khi đầu tư vào hai lĩnh vực A,B cho kết quả như sau:

Tiền lãi	[5;10)	[10;15)	[15;20)	[20;25)	[25;30)
Số nhà đầu tư vào lĩnh vực A	2	5	8	6	4
Số nhà đầu tư vào lĩnh vực B	8	4	2	5	6

Tính độ lệch chuẩn cho các mẫu số liệu về tiền lãi của các nhà đầu tư ở lĩnh vực B

A. $8,039$. **B.** $8,5314$. **C.** $8,6214$. **D.** $8,5268$.

Lời giải

Ta có mẫu số liệu ghép nhóm với giá tri đại diện là:

Tiền lãi	[5;10)	[10;15)	[15;20)	[20;25)	[25;30)
Giá trị đại diện	7,5	12,5	17,5	22,5	27,5
Số nhà đầu tư vào lĩnh vực A	2	5	8	6	4
Số nhà đầu tư vào lĩnh vực B	8	4	2	5	6

Phương sai của mẫu số liệu về tiền lãi khi đầu tư vào lĩnh vực B:

$$s_B^2 = \frac{1}{25}\left(7,5^2.8+12,5^2.4+17,5^2.2+22,5^2.5+27,5^2.6\right)-16,9^2 = 64,64$$

Độ lệch chuẩn của mẫu số liệu về tiền lãi khi đầu tư vào lĩnh vực B: $s_B = \sqrt{64,64} \approx 8,039$ (triệu đồng)

PHẦN II. Câu trắc nghiệm đúng sai. Trong mỗi ý a), b), c), d) ở mỗi câu, thí sinh chọn đúng hoặc sai.

Câu 1: Kết quả 40 lần nhảy xa của hai vận động viên Dũng và Huy được lần lượt thống kê trong bảng bên dưới (đơn vị: mét)

Nhóm	Dũng	Huy
$[6,22;6,46)$	3	2
$[6,46;6,70)$	7	5
$[6,70;6,94)$	5	8
$[6,94;7,18)$	20	19
$[7,18;7,42)$	5	6
	$n=40$	$n=40$

Xét tính đúng sai của các khẳng định sau:

a) Số trung bình cộng của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Dũng cho bởi Bảng 15 (làm tròn kết quả đến hàng phần trăm) là $6,92$ m.

b) Độ lệch chuẩn của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Dũng cho bởi Bảng 15 (làm tròn kết quả đến hàng phần trăm) là $0,26$ m.

c) Phương sai của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Huy cho bởi Bảng 16 (làm tròn kết quả đến hàng phần trăm) là $0,16$.

d) Kết quả nhảy xa của vận động viên Dũng đồng đều hơn kết quả nhảy xa của vận động viên Huy.

Lời giải

a) Đúng: Số trung bình cộng của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Dũng là:

$$\overline{x}_D = \frac{3.6,34 + 7.6,58 + 5.6,82 + 20.7,06 + 5.7,30}{40} = \frac{276,88}{40} \approx 6,92 \text{ (m)}$$

b) Đúng: Phương sai của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Dũng (làm tròn kết quả đến hàng phần trăm) là:

$$s_D^2 = \frac{1}{40}\Big[3.(6,34-6,92)^2 + 7.(6,58-6,92)^2 + 5.(6,82-6,92)^2$$

$$+20.(7,06-6,92)^2 + 5.(7,30-6,92)^2\Big] = \frac{2,9824}{40} \approx 0,07.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm trên là: $s_D \approx \sqrt{0,07} \approx 0,26$ (m)

c) Sai: Số trung bình cộng của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Huy là:

$$\overline{x}_H = \frac{2.6,34 + 5.6,58 + 8.6,82 + 19.7,06 + 6.7,30}{40} = \frac{278,08}{40} \approx 6,95 \text{ (m)}$$

Vậy phương sai của mẫu số liệu ghép nhóm biểu diễn kết quả 40 lần nhảy xa của vận động viên Huy (làm tròn kết quả đến hàng phần trăm) là:

$$s_H^2 = \frac{1}{40}\Big[2.(6,34-6,95)^2 + 5.(6,58-6,95)^2 + 8.(6,82-6,95)^2 + 19.(7,06-6,95)^2$$

$$+6.(7,30-6,95)^2\Big] = \frac{2,5288}{40} \approx 0,06$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm trên là: $s_H \approx \sqrt{0,06} \approx 0,24$ (m)

d) Sai: Do $s_H \approx 0,24 < s_D \approx 0,26$ nên kết quả nhảy xa của vận động viên Huy đồng đều hơn kết quả nhảy xa của vận động viên Dũng.

Câu 2: Mẫu số liệu ghép nhóm thống kê mức lương của hai công ty A, B (đơn vị: triệu đồng) được thể hiện như bảng dưới đây

Nhóm	Công ty A	Công ty B
$[10;15)$	15	25
$[15;20)$	18	15
$[20;25)$	10	7
$[25;30)$	10	5
$[30;35)$	5	5
$[35;40)$	2	3
	$n=60$	$n=60$

Xét tính đúng sai của các khẳng định sau:

a) Số trung bình cộng của mẫu số liệu ghép nhóm của công ty A là: $\frac{62}{3}$

b) Độ lệch chuẩn của mẫu số liệu ghép nhóm của công ty A là: 5

c) Phương sai của mẫu số liệu ghép nhóm của công ty B (làm tròn kết quả đến hàng phần trăm) là $52,91$.

b) Nhận thấy độ lệch chuẩn của công ty A nhỏ hơn công ty B nên mức lương của công ty A đồng đều hơn.

Lời giải

a) Đúng: Số trung bình cộng của mẫu số liệu ghép nhóm của công ty A là:

$$\overline{x_A}=\frac{15.12,5+18.17,5+10.22,5+10.27,5+5.32,5+2.37,5}{60}=\frac{62}{3}$$

b) Sai: Phương sai của mẫu số liệu ghép nhóm của công ty A là:

$$s_A^2=\frac{15\cdot\left(12,5-\frac{62}{3}\right)^2+18\cdot\left(17,5-\frac{62}{3}\right)^2+10\cdot\left(22,5-\frac{62}{3}\right)^2+10\cdot\left(27,5-\frac{62}{3}\right)^2}{60}$$

$$+\frac{5\cdot\left(32,5-\frac{62}{3}\right)^2+2\cdot\left(37,5-\frac{62}{3}\right)^2}{60}\approx 49,14$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm của công ty A là: $s_A=\sqrt{s_A^2}=\sqrt{49,13}\approx 7$

c) Sai: Số trung bình cộng của mẫu số liệu ghép nhóm của công ty B là:

$$\overline{x_B}=\frac{25.12,5+15.17,5+7.22,5+5.27,5+5.32,5+3.37,5}{60}=\frac{229}{12}$$

Phương sai của mẫu số liệu ghép nhóm của công ty B là:

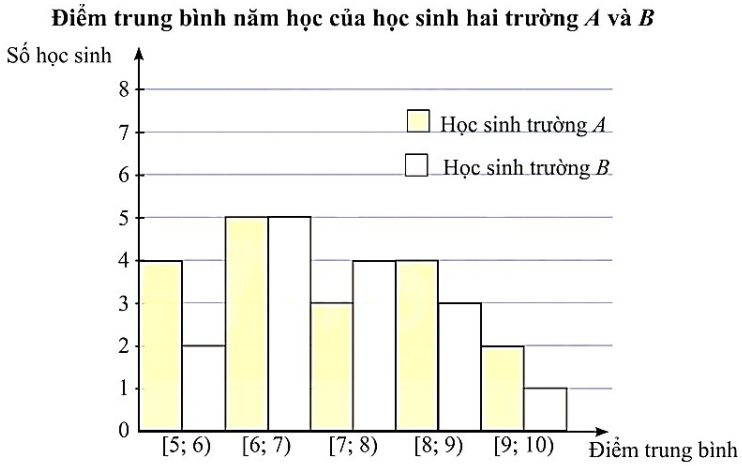
$$s_B^2 = \frac{25.\left(12,5-\frac{229}{12}\right)^2 + 15.\left(17,5-\frac{229}{12}\right)^2 + 7.\left(22,5-\frac{229}{12}\right)^2 + 5.\left(27,5-\frac{229}{12}\right)^2}{60}$$

$$+\frac{5.\left(32,5-\frac{229}{12}\right)^2 + 3.\left(37,5-\frac{229}{12}\right)^2}{60} \approx 57,91$$

d) Đúng: Độ lệch chuẩn của mẫu số liệu ghép nhóm của công ty B là: $s_B = \sqrt{s_B^2} = \sqrt{57,91} \approx 7,61$

Nhận thấy độ lệch chuẩn của công ty A nhỏ hơn công ty B nên mức lương của công ty A đồng đều hơn.

Câu 3: Biểu đồ sau mô tả kết quả điều tra về điểm trung bình năm học của học sinh hai trường A và B



Người ta lập được bảng tần số ghép nhóm cho mẫu số liệu trên như sau:

Điểm trung bình	[5;6)	[6;7)	[7;8)	[8;9)	[9;10)
Giá trị đại diện	5,5	6,5	7,5	8,5	9,5
Học sinh trường A	4	5	3	4	2
Học sinh trường B	2	5	4	3	1

Xét tính đúng sai của các khẳng định sau:

a) Tứ phân vị thứ nhất của mẫu số liệu ghép nhóm của học sinh trường A là: $6,1$

b) Khoảng tứ phân vị của mẫu số liệu ghép nhóm của học sinh trường B là: $1,73$

c) Nếu so sánh theo khoảng tứ phân vị của mẫu số liệu ghép nhóm thì học sinh trường B có điểm trung bình đồng đều hơn.

d) Nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì học sinh trường A có điểm trung bình đồng đều hơn.

Lời giải

a) Đúng: Cỡ mẫu: $n_A = 18$

Gọi $x_1;x_2;\ldots;x_{18}$ là mẫu số liệu gốc về điểm trung bình năm học của học sinh hai trường A được xếp theo thứ tự không giảm.

Ta có: $x_1;\ldots;x_4\in[5;6);x_5;\ldots;x_9\in[6;7);x_{10};\ldots;x_{12}\in[7;8);x_{12};\ldots;x_{16}\in[8;9);x_{17};x_{18}\in[9;10)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $x_5\in[6;7)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là: $Q_1=6+\dfrac{\dfrac{18}{4}-4}{5}(7-6)=6,1$

Tứ phân vị thứ ba của mẫu số liệu gốc là $x_{14}\in[8;9)$.

Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là: $Q_3=8+\dfrac{\dfrac{3.18}{4}-(4+5+3)}{4}(9-8)=8,375$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q=Q_3-Q_1=2,275$

b) Đúng: Cỡ mẫu: $n_B=15$

Gọi $y_1;y_2;\ldots;y_{15}$ là mẫu số liệu gốc về điểm trung bình năm học của học sinh hai trường B được xếp theo thứ tự không giảm.

Ta có: $y_1;y_2\in[5;6);y_3;\ldots;y_7\in[6;7);y_8;\ldots;y_{11}\in[7;8);y_{12};\ldots;y_{14}\in[8;9);y_{15}\in[9;10)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $y_4\in[6;7)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là: $Q_1'=6+\dfrac{\dfrac{15}{4}-2}{5}(7-6)=6,35$

Tứ phân vị thứ ba của mẫu số liệu gốc là $y_{12}\in[8;9)$.

Do đó tứ phân vị thứ ba của mẫu số liệu ghép nhóm là: $Q_3'=8+\dfrac{\dfrac{3.15}{4}-(2+5+4)}{3}(9-8)=8,08$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q'=Q_3'-Q_1'=1,73$

c) Đúng: Vậy nếu so sánh theo khoảng tứ phân vị của mẫu số liệu ghép nhóm thì học sinh trường B có điểm trung bình đồng đều hơn.

d) Sai: Xét số liệu của trường A:

Số trung bình: $\overline{x_A}=\dfrac{4.5,5+5.6,5+3.7,5+4.8,5+2.9,5}{18}=7,22$

Độ lệch chuẩn: $\sigma_A = \sqrt{\frac{4.5,5^2 + 5.6,5^2 + 3.7,5^2 + 4.8,5^2 + 2.9,5^2}{18} - 7,22^2} \approx 1,79$

Xét số liệu của trường B :

Số trung bình: $\overline{x_B} = \frac{2.5,5 + 5.6,5 + 4.7,5 + 3.8,5 + 1.9,5}{15} = 7,23$

Độ lệch chuẩn: $\sigma_B = \sqrt{\frac{2.5,5^2 + 5.6,5^2 + 4.7,5^2 + 3.8,5^2 + 1.9,5^2}{15} - 7,23^2} \approx 1,31$

Vậy nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì học sinh trường B có điểm trung bình đồng đều hơn.

Câu 4: Bảng dưới đây biểu diễn mẫu số liệu ghép nhóm về nhiệt độ không khí trung bình các tháng năm 2021 tại Hà Nội và Huế (đơn vị: độ C).

Nhóm	Hà Nội	Huế
$[16,8;19,8)$	2	1
$[19,8;22,8)$	3	2
$[22,8;25,8)$	2	3
$[25,8;28,8)$	1	2
$[28,8;31,8)$	4	4
	$n = 12$	$n = 12$

Xét tính đúng sai của các khẳng định sau:

a) Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên số liệu ở Hà Nội là: $8,75$

b) Độ lệch chuẩn của mẫu số liệu ghép nhóm trên số liệu ở Hà Nội (làm tròn kết quả đến hàng phần trăm) là: $3,56$

c) Phương sai của mẫu số liệu ghép nhóm trên số liệu ở Huế là: $12,75$

d) Huế có nhiệt độ không khí trung bình tháng đồng đều hơn vì độ lệch chuẩn nhỏ hơn

Lời giải

a) Đúng: Xét số liệu ở Hà Nội:

Khoảng biến thiên: $R = 31,8 - 16,8 = 15$

Số phần tử của mẫu là $n = 12$

Tần số tích lũy của các nhóm lần lượt là $cf_1 = 2, cf_2 = 5, cf_3 = 7, cf_4 = 8, cf_5 = 12$

Ta có: $\frac{n}{4}=\frac{12}{4}=3$ mà $2<3<5$ suy ra nhóm 2 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 3.

Xét nhóm 2 là nhóm $[19,8;22,8)$ có $s=19,8, h=3, n_2=3$ và nhóm 1 là nhóm $[16,8;19,8)$ có $cf_1=2$. Ta có tứ phân vị thứ nhất là: $Q_1=s+\left(\frac{3-cf_1}{n_2}\right).h=19,8+\left(\frac{3-2}{3}\right).3=20,8$

Ta có: $\frac{3n}{4}=\frac{3.12}{4}=9$ mà $8<9<12$ suy ra nhóm 5 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 9.

Xét nhóm 5 là nhóm $[28,8;31,8)$ có $t=28,8, l=3, n_5=4$ và nhóm 4 là nhóm $[25,8;28,8)$ có $cf_4=8$

Ta có tứ phân vị thứ ba là: $Q_3=t+\left(\frac{9-cf_4}{n_5}\right).l=28,8+\left(\frac{9-8}{4}\right).3=29,55$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $Q_3-Q_1=29,55-20,8=8,75$

b) Sai: Số trung bình cộng của mẫu số liệu ghép nhóm là:

$$\overline{x_1}=\frac{2.18,3+3.21,3+2.24,3+27,3+4.30,3}{12}=24,8$$

Phương sai của mẫu số liệu ghép nhóm là:

$$s_1^2=\frac{2(18,3-24,8)^2+3(21,3-24,8)^2+2(24,3-24,8)^2+(27,3-24,8)^2+4(30,3-24,8)^2}{12}=20,75$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s_1=\sqrt{s_1^2}=\sqrt{20,75}\approx 4,56$

c) Sai: Xét số liệu ở Huế:

Khoảng biến thiên: $R=31,8-16,8=15$

Số phần tử của mẫu là $n=12$

Tần số tích lũy của các nhóm lần lượt là $cf_1=1, cf_2=3, cf_3=6, cf_4=8, cf_5=12$

Ta có: $\frac{n}{4}=\frac{12}{4}=3$ suy ra nhóm 2 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 3. Xét nhóm 2 là nhóm $[19,8;22,8)$ có $s=19,8, h=3, n_2=2$ và nhóm 1 là nhóm $[16,8;19,8)$ có $cf_1=1$

Ta có tứ phân vị thứ nhất là: $Q_1=s+\left(\frac{3-cf_1}{n_2}\right).h=19,8+\left(\frac{3-1}{2}\right).3=22,8$

Ta có: $\frac{3n}{4}=\frac{3.12}{4}=9$ mà $8<9<12$ suy ra nhóm 5 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 9.

Xét nhóm 5 là nhóm $[28,8;31,8)$ có $t=28,8, l=3, n_5=4$ và nhóm 4 là nhóm $[25,8;28,8)$ có $cf_4=8$

Ta có tứ phân vị thứ ba là: $Q_3=t+\left(\frac{9-cf_4}{n_5}\right).l=28,8+\left(\frac{9-8}{4}\right).3=29,55$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $Q_3-Q_1=29,55-22,8=6,75$

Số trung bình cộng của mẫu số liệu ghép nhóm là:

$$\overline{x_2}=\frac{18,3+2.21,3+3.24,3+2.27,3+4.30,3}{12}=25,8$$

Phương sai của mẫu số liệu ghép nhóm là:

$$s_2^2=\frac{(18,3-25,8)^2+3(21,3-25,8)^2+3(24,3-25,8)^2+2(27,3-25,8)^2+4(30,3-25,8)^2}{12}=15,75$$

d) Đúng: Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s_2=\sqrt{s_2^2}=\sqrt{15,75}\approx 3,97$

Huế có nhiệt độ không khí trung bình tháng đồng đều hơn vì độ lệch chuẩn nhỏ hơn

Câu 5: Bảng dưới đây thống kê độ ẩm không khí trung bình các tháng năm 2021 tại Đà Lạt và Vũng Tàu (đơn vị: %).

Tháng	1	2	3	4	5	6	7	8	9	10	11	12
Đà Lạt	83	79	79	87	87	87	88	89	90	91	88	86
Vũng Tàu	75	77	78	77	79	79	81	79	81	83	80	77

Người ta lần lượt ghép các số liệu của Đà Lạt, Vũng Tàu thành mẫu số liệu ghép nhóm như sau:

Đà Lạt			Vũng Tàu		
Nhóm	Giá trị đại diện	Tần số	Nhóm	Giá trị đại diện	Tần số
[75;78,3)	76,65	0	[75;78,3)	76,65	5
[78,3;81.6)	79,95	2	[78,3;81.6)	79,95	6
[81,6;84,9)	83,25	1	[81,6;84,9)	83,25	1
[84,9;88,2)	86,55	6	[84,9;88,2)	86,55	0
[88,2;91,5)	89,85	3	[88,2;91,5)	89,85	0

Xét tính đúng sai của các khẳng định sau:

a) Xét số liệu ở Đà Lạt ta có khoảng biến thiên là $16,5$.

b) Xét số liệu ở Đà Lạt thì ta có độ lệch chuẩn của mẫu số liệu ghép nhóm (làm tròn kết quả đến hàng phần trăm) là $3,28$.

c) Xét số liệu ở Vũng Tàu thì phương sai của mẫu số liệu ghép nhóm là $4,235$.

d) Đà Lạt có nhiệt độ không khí trung bình tháng đồng đều hơn vì độ lệch chuẩn nhỏ hơn

Lời giải

a) Đúng: Xét số liệu ở Đà Lạt:

Khoảng biến thiên: $R = 91,5 - 75 = 16,5$

b) Đúng: Số phần tử của mẫu là $n = 12$

Tần số tích lũy của các nhóm lần lượt là $cf_1 = 0, cf_2 = 2, cf_3 = 3, cf_4 = 9, cf_5 = 12$

Ta có: $\frac{n}{4} = \frac{12}{4} = 3$ suy ra nhóm 3 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 3.

Xét nhóm 3 là nhóm $[81,6;84,9)$ có $s = 81,6, h = 3,3, n_3 = 1$ và nhóm 2 là nhóm $[78,3;81,6)$ có $cf_2 = 2$. Ta có tứ phân vị thứ nhất là: $Q_1 = s + \left(\frac{3 - cf_2}{n_3}\right) \cdot h = 81,6 + \left(\frac{3-2}{1}\right) \cdot 3,3 = 84,9$

Ta có: $\frac{3n}{4} = \frac{3.12}{4} = 9$ suy ra nhóm 4 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 9.

Xét nhóm 4 là nhóm $[84,9;88,2)$ có $t = 84,9, l = 3,3, n_4 = 6$ và nhóm 3 là nhóm $[81,6;84,9)$ có $cf_3 = 3$

Ta có tứ phân vị thứ ba là: $Q_3 = t + \left(\frac{9 - cf_3}{n_4}\right).l = 84,9 + \left(\frac{9-3}{6}\right).3,3 = 88,2$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $Q_3 - Q_1 = 88,2 - 84,9 = 3,3$

Số trung bình cộng của mẫu số liệu ghép nhóm là:

$$\overline{x_1} = \frac{0.76,65 + 2.79,95 + 83,25 + 6.86,55 + 3.89,85}{12} = 86$$

Phương sai của mẫu số liệu ghép nhóm là:

$$s_1^2 = \frac{0(76,65-86)^2 + 2(79,95-86)^2 + (83,25-86)^2 + 6(86,55-86)^2 + 3(89,95-86)^2}{12} = 10,7825$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s_1 = \sqrt{s_1^2} = \sqrt{10,7825} \approx 3,28$

c) Xét số liệu ở Vũng Tàu:

Khoảng biến thiên: $R = 91,5 - 75 = 16,5$

Số phần tử của mẫu là $n=12$

Tần số tích lũy của các nhóm lần lượt là $cf_1=5, cf_2=11, cf_3=12, cf_4=12, cf_5=12$

Ta có: $\frac{n}{4}=\frac{12}{4}=3$ suy ra nhóm 1 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 3.

Xét nhóm 1 là nhóm $[75;78,3)$ có $s=75, h=3,3, n_1=5$

Ta có tứ phân vị thứ nhất là: $Q_1=s+\left(\frac{3-cf_0}{n_1}\right)\cdot h=75+\left(\frac{3-0}{5}\right).3,3=76,98$

Ta có: $\frac{3n}{4}=\frac{3.12}{4}=9$ suy ra nhóm 2 là nhóm đầu tiên có tần số tích lũy lớn hơn hoặc bằng 9.

Xét nhóm 2 là nhóm $[78,3;81,6)$ có $t=78,3, l=3,3, n_2=6$ và nhóm 1 là nhóm $[75;78,3)$ có $cf_1=5$. Ta có tứ phân vị thứ ba là: $Q_3=t+\left(\frac{9-cf_1}{n_2}\right).l=75+\left(\frac{9-5}{6}\right).3,3=77,2$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là: $Q_3-Q_1=77,2-76,98=0,22$

Số trung bình cộng của mẫu số liệu ghép nhóm là: $\overline{x_2}=\frac{5.76,65+6.79,95+83,25}{12}=78,85$

Phương sai của mẫu số liệu ghép nhóm là:

$$s_2^2=\frac{5(76,65-78,85)^2+6(79,95-78,85)^2+(83,25-78,85)^2}{12}=4,235$$

d) Sai: Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $s_2=\sqrt{s_2^2}=\sqrt{4,235}\approx 2,06$

Vũng Tàu có nhiệt độ không khí trung bình tháng đồng đều hơn vì độ lệch chuẩn nhỏ hơn.

Câu 6: Bảng sau thống kê lại tổng số giờ nắng trong tháng 6 của các năm từ 2002 đến 2021 tại hai trạm quan trắc đặt ở Nha Trang và Quy Nhơn.

Số giờ nắng	[130;160)	[160;190)	[190;220)	[220;250)	[250;280)	[280;310)
Số năm ở Nha Trang	1	1	1	8	7	2
Số năm ở Quy Nhơn	0	1	2	4	10	3

Xét tính đúng sai của các khẳng định sau:

a) Xét số liệu ở Nha Trang thì khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $32,64$

b) Nếu so sánh theo khoảng tứ phân vị thì số giờ nắng trong tháng 6 của Quy Nhơn đồng đều hơn

c) Xét số liệu của Quy Nhơn ta có độ lệch chuẩn của mẫu số liệu ghép nhóm (làm tròn kết quả đến hàng phần trăm) là: $30,59$

d) Nếu so sánh theo độ lệch chuẩn thì số giờ nắng trong tháng 6 của Nha Trang đồng đều hơn

a) Sai: Cỡ mẫu: $n = 20$

Gọi $x_1; x_2; \ldots; x_{20}$ là mẫu số liệu gốc về số giờ nắng trong tháng 6 trong 20 năm của Nha Trang được xếp theo thứ tự không giảm.

Ta có: $x_1 \in [130;160); x_2 \in [160;190); x_3 \in [190;220); x_4; \ldots; x_{11} \in [220;250); x_{12}; \ldots;$

$x_{18} \in [250;280); x_{19}; x_{20} \in [280;310)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(x_5 + x_6) \in [220;250)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là:

$$Q_1 = 220 + \frac{\frac{20}{4} - (1+1+1)}{8}(250 - 220) = 227,5$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(x_{15} + x_{16}) \in [250;280)$.

Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là:

$$Q_3 = 250 + \frac{\frac{3.20}{4} - (1+1+1+8)}{7}(280 - 250) = \frac{1870}{7}$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q = Q_3 - Q_1 = 39,64$

b) Đúng: Gọi $y_1; y_2; \ldots; y_{50}$ là mẫu số liệu gốc về số giờ nắng trong tháng 6 trong 20 năm của Quy Nhơn được xếp theo thứ tự không giảm.

Ta có:

$y_1; \in [160;190); y_2; y_3 \in [190;220); y_4; \ldots; y_7 \in [220;250); y_8; \ldots; y_{17} \in [250;280);$
$y_{4=18}; \ldots; y_{20} \in [280;310)$

Tứ phân vị thứ nhất của mẫu số liệu gốc là $\frac{1}{2}(y_5 + y_6) \in [220;250)$.

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là:

$$Q_1' = 220 + \frac{\frac{20}{4} - (1+2)}{4}(250 - 220) = 235$$

Tứ phân vị thứ ba của mẫu số liệu gốc là $\frac{1}{2}(y_{15} + y_{16}) \in [250;280)$.

Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là:

$$Q_3' = 250 + \frac{\frac{3.20}{4} - (1+2+4)}{10}(280-250) = 274$$

Khoảng tứ phân vị của mẫu số liệu ghép nhóm là: $\Delta_Q' = Q_3' - Q_1' = 39$

Vậy nếu so sánh theo khoảng tứ phân vị thì số giờ nắng trong tháng 6 của Quy Nhơn đồng đều hơn

Xét số liệu của Nha Trang:

Số trung bình: $\overline{x_X} = \frac{1.145+1.175+1.205+8.235+7.265+2.295}{20} = 242,5$

Độ lệch chuẩn: $\sigma_X = \sqrt{\frac{1.145^2+1.175^2+1.205^2+8.235^2+7.265^2+2.295^2}{20} - 242,5^2} \approx 35,34$

c) Đúng: Xét số liệu của Quy Nhơn:

Số trung bình: $\overline{x_Y} = \frac{1.175+2.205+4.235+10.265+3.295}{20} = 253$

Độ lệch chuẩn: $\sigma_Y = \sqrt{\frac{1.175^2+2.205^2+4.235^2+10.265^2+3.295^2}{20} - 253^2} \approx 30,59$

d) Sai: Vậy nếu so sánh theo độ lệch chuẩn thì số giờ nắng trong tháng 6 của Quy Nhơn đồng đều hơn.

Câu 7: Thống kê tổng số giờ nắng trong tháng 9 tại một trạm quan trắc đặt ở Cà Mau trong các năm từ 2002 đến 2021 được thống kê như sau:

111,6	134,9	130,3	134,2	140,9	109,3	152,4	156,3	116,1	96,7
105,2	80,8	80,8	110	109	139	145	161	126	114

Người ta lập được bảng tần số ghép nhóm như sau:

Số giờ nắng	[80;98)	[98;116)	[116;134)	[134;152)	[152;170)
Giá trị đại diện	89	107	125	143	161
Số năm	3	6	3	5	3

Xét tính đúng sai của các khẳng định sau:

a) Số trung bình của mẫu số liệu ghép nhóm là $124,1$.

b) Phương sai của mẫu số liệu ghép nhóm là $566,19$.

c) Độ lệch chuẩn của mẫu số liệu ghép nhóm (kết quả các phép tính làm tròn đến hàng phần nghìn) là $23,795$.

d) Sai số tương đối của độ lệch chuẩn của mẫu số liệu ghép nhóm so với độ lệch chuẩn của mẫu số liệu gốc (kết quả các phép tính làm tròn đến hàng phần nghìn) là $4,805\%$.

Lời giải

a) Đúng: Xét theo mẫu số liệu gốc. Cỡ mẫu là $n = 20$.

Số trung bình của mẫu số liệu trên là: $\overline{x_1} = \dfrac{111,6 + 134,9 + \ldots + 114}{20} = 122,755.$

Phương sai của mẫu số liệu trên là: $S_1^2 = \dfrac{1}{20}\left(111,6^2 + 134,9^2 + \ldots + 114^2\right) - 122,755^2 \approx 515,453.$

Độ lệch chuẩn của mẫu số liệu trên là: $S_1 \approx \sqrt{515,453} \approx 22,704.$

Xét theo mẫu số liệu ghép nhóm

Số trung bình của mẫu số liệu ghép nhóm là $\overline{x_2} = \dfrac{3.89 + 6.107 + 3.125 + 5.143 + 3.161}{20} = 124,1.$

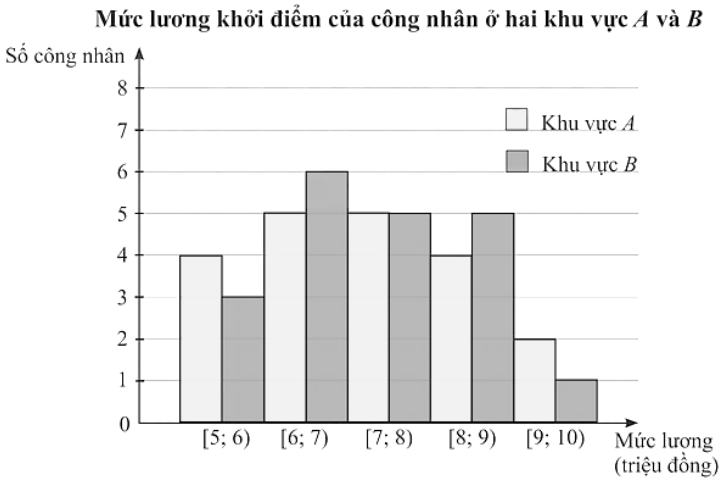
b) Đúng: Phương sai của mẫu số liệu ghép nhóm là

$$S_2^2 = \frac{1}{20}\left(3.89^2 + 6.107^2 + 3.125^2 + 5.143^2 + 3.161^2\right) - 124,1^2 = 566,19.$$

c) Đúng: Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S_2 = \sqrt{566,19} \approx 23,795.$ 77

d) Đúng: Sai số tương đối của độ lệch chuẩn của mẫu số liệu ghép nhóm so với độ lệch chuẩn của mẫu số liệu gốc là: $\dfrac{|S_2 - S_1|}{S_1} = \dfrac{|23,795 - 22,704|}{22,704} \cdot 100\% \approx 4,805\%.$

Câu 8: Biểu đồ dưới đây mô tả kết quả điều tra về mức lương khởi điểm (đơn vị: triệu đồng) của một số công nhân ở hai khu vực A và B.



Người ta lập được bảng tần số ghép nhóm cho mẫu số liệu như sau:

Mức lương	[5;6)	[6;7)	[7;8)	[8;9)	[9;10)
Mức lương đại diện (triệu đồng)	5,5	6,5	7,5	8,5	9,5
Khu vực A	4	5	5	4	2
Khu vực B	3	6	5	5	1

Xét tính đúng sai của các khẳng định sau:

a) Xét mẫu số liệu của khu vực A ta có số trung bình của mẫu số liệu ghép nhóm là $6,25.$

b) Xét mẫu số liệu của khu vực A ta có độ lệch chuẩn của mẫu số liệu ghép nhóm là $\sqrt{1,5875}$

c) Xét mẫu số liệu của khu vực B ta có phương sai của mẫu số liệu ghép nhóm là $1,3875.$

d) Nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì mức lương khởi điểm của công nhân khu vực B đồng đều hơn của công nhân khu vực A.

Lời giải

a) Sai: Xét mẫu số liệu của khu vực A. Cỡ mẫu là $n_A = 4+5+5+4+2 = 20$.

Số trung bình của mẫu số liệu ghép nhóm là $\overline{x}_A = \dfrac{4.5,5+5.6,5+5.7,5+4.8,5+2.9,5}{20} = 7,25.$

Phương sai của mẫu số liệu ghép nhóm là

$$S_A^2 = \frac{1}{20}\left(4.5,5^2+5.6,5^2+5.7,5^2+4.8,5^2+2.9,5^2\right)-(7,25)^2 = 1,5875$$

b) Đúng: Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S_A = \sqrt{1,5875}$

Xét mẫu số liệu của khu vực B. Cỡ mẫu là $n_B = 3+6+5+5+1 = 20$.

Số trung bình của mẫu số liệu ghép nhóm là: $\overline{x}_B = \dfrac{3.5,5+6.6,5+5.7,5+5.8,5+1.9,5}{20} = 7,25.$

c) Sai: Phương sai của mẫu số liệu ghép nhóm là

$$S_B^2 = \frac{1}{20}\left(3\cdot 5,5^2+6\cdot 6,5^2+5\cdot 7,5^2+5.8,5^2+1.9,5^2\right)-(7,25)^2 = 1,2875.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S_B = \sqrt{1,2875}$

d) Đúng: Do $S_A > S_B$ nên nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì mức lương khởi điểm của công nhân khu vực B đồng đều hơn của công nhân khu vực A.

Câu 9: Giá đóng cửa của một cổ phiếu là giá của cổ phiếu đó cuối một phiên giao dịch. Bảng sau thống kê giá đóng cửa (đơn vị: nghìn đồng) của hai mã cổ phiếu A và B trong 50 ngày giao dịch liên tiếp.

Giá đóng cửa	$[120;122)$	$[122;124)$	$[124;126)$	$[126;128)$	$[128;130)$
Cổ phiếu A	8	9	12	10	11
Cổ phiếu B	16	4	3	6	21

Xét tính đúng sai của các khẳng định sau:

a) Xét mẫu số liệu của cổ phiếu A ta có phương sai của mẫu số liệu ghép nhóm là $7,5216$

b) Xét mẫu số liệu của cổ phiếu B ta có số trung bình của mẫu số liệu ghép nhóm là $115,28.$

c) Xét mẫu số liệu của cổ phiếu B ta có độ lệch chuẩn của mẫu số liệu ghép nhóm là $\sqrt{15,4096}$

d) Người ta có thể dùng phương sai và độ lệch chuẩn để so sánh mức độ rủi ro của các loại cổ phiếu có giá trị trung bình gần bằng nhau. Cổ phiếu nào có phương sai, độ lệch chuẩn cao hơn thì được coi là có độ rủi ro lớn hơn. Theo quan điểm trên, thì cổ phiếu A có độ rủi ro thấp hơn cổ phiếu B.

Lời giải

Ta có bảng thống kê giá đóng cửa theo giá trị đại diện:

Giá đóng cửa	121	123	125	127	129
Cổ phiếu A	8	9	12	10	11
Cổ phiếu B	16	4	3	6	21

a) Đúng: Xét mẫu số liệu của cổ phiếu A:

Số trung bình của mẫu số liệu ghép nhóm là

$$\overline{x_1} = \frac{8.121 + 9.123 + 12.125 + 10.127 + 11.129}{50} = 125,28.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_1^2 = \frac{1}{50}\left(8.121^2 + 9.123^2 + 12.125^2 + 10.127^2 + 11.129^2\right) - (125,28)^2 = 7,5216$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_1 = \sqrt{S_1^2} = \sqrt{7,5216}$.

b) Sai: Xét mẫu số liệu của cổ phiếu B:

Số trung bình của mẫu số liệu ghép nhóm là

$$\overline{x_2} = \frac{16.121 + 4.123 + 3.125 + 6.127 + 21.129}{50} = 125,28.$$

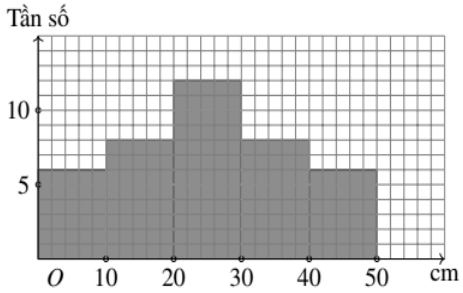
c) Sai: Phương sai của mẫu số liệu ghép nhóm là

$$S_2^2 = \frac{1}{50}\left(16.121^2 + 4.123^2 + 3.125^2 + 6.127^2 + 21.129^2\right) - \left(125,48\right)^2 = 12,4096.$$

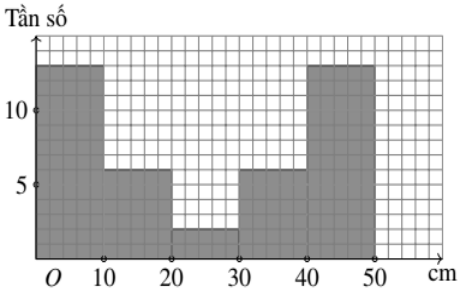
Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_2 = \sqrt{S_2^2} = \sqrt{12,4096}$.

d) Vậy nếu đánh giá độ rủi ro theo phương sai và độ lệch chuẩn thì cổ phiếu A có độ rủi ro thấp hơn cổ phiếu B.

Câu 10: Một công ty giống cây trồng đã thử nghiệm hai phương pháp chăm sóc khác nhau cho cây hướng dương. Sau hai tuần, người ta thấy cây được chăm sóc theo cả hai phương pháp đều thấp hơn 50 cm.



Chiều cao của cây chăm sóc theo phương pháp A



Chiều cao của cây chăm sóc theo phương pháp B

a) Khoảng biến thiên của chiều cao các cây được chăm sóc theo mỗi phương pháp A và B bằng nhau.

b) Trung bình của chiều cao các cây được chăm sóc theo mỗi phương pháp A và B bằng nhau.

c) Độ lệch chuẩn của chiều cao các cây được chăm sóc theo phương án A là $12,65$ (cm).

d) Dựa vào độ lệch chuẩn thì chiều cao của các loại cây được chăm sóc theo phương án B ít bị chênh lệch hơn so với phương án A .

Lời giải

a) Đúng: Khoảng biến thiên của chiều cao các cây được chăm sóc theo mỗi phương pháp A và B bằng nhau và cùng bằng 50.

b) Đúng: Ước tính số trung bình và độ lệch chuẩn của chiều cao các cây được chăm sóc theo mỗi phương pháp. Cỡ mẫu của hai mẫu số liệu thống kê là $N = 40$. Ta có bảng tần số ghép nhóm về chiều cao của cây được chăm sóc theo phương pháp A như sau:

Chiều cao (cm)	$[0;10)$	$[10;20)$	$[20;30)$	$[30;40)$	$[40;50)$
Giá trị đại diện	5	15	25	35	45
Tần số	6	8	12	8	6

Chiều cao trung bình của các cây được chăm sóc theo phương án A là:

$$\overline{x_A} = \frac{5.6 + 18.5 + 25.12 + 35.8 + 45.6}{40} = 25$$

Ta có bảng tần số ghép nhóm về chiều cao của cây được chăm sóc theo phương pháp B như sau:

Chiều cao (cm)	$[0;10)$	$[10;20)$	$[20;30)$	$[30;40)$	$[40;50)$
Giá trị đại diện	5	15	25	35	45
Tần số	13	6	2	6	13

Chiều cao trung bình của các cây được chăm sóc theo phương án B là:

$$\overline{x_B} = \frac{5.13 + 15.6 + 25.2 + 35.6 + 45.13}{40} = 25 \text{ cm.}$$

c) Độ lệch chuẩn của chiều cao các cây được chăm sóc theo phương án A là:

$$s_A = \sqrt{\frac{5^2.6 + 15^2.8 + 25^2.12 + 35^2.8 + 45^2.6}{40} - 25^2} \approx 12,65$$

d) Độ lệch chuẩn của chiều cao các cây được chăm sóc theo phương án B là:

$$s_B=\sqrt{\frac{5^2.13+15^2.6+25^2.2+35^2.6+45^2.13}{40}-25^2}\approx 17,03\,\text{cm}^2$$

PHẦN III. Câu trắc nghiệm trả lời ngắn

Câu 1: Hãy tính độ lệch chuẩn của mẫu số liệu ghép nhóm (kết quả được làm tròn đến hàng phần trăm)

Chiều cao (cm)	[160;164)	[164;168)	[168;172)	[172;176)	[176;180)
Số học sinh	3	5	8	4	1

Lời giải

Cỡ mẫu: $n=21$. Giá trị trung bình của mẫu số liệu mới:

$$\overline{x}=\frac{1}{n}\left(n_1c_1+n_2c_2+\ldots+n_kc_k\right)=\frac{1}{21}\left(3.162+5.166+8.170+4.174+1.178\right)=\frac{3550}{21}$$

Phương sai của mẫu số liệu mới: $S^2=\frac{1}{n}\left[n_1\left(c_1-\overline{x}\right)^2+n_2\left(c_2-\overline{x}\right)^2+\ldots+n_k\left(c_k-\overline{x}\right)^2\right]$

$$=\frac{1}{21}\left[3\left(162-\frac{3550}{21}\right)^2+5\left(166-\frac{3550}{21}\right)^2+\ldots+1\left(178-\frac{3550}{21}\right)^2\right]=\frac{8000}{441}$$

Độ lệch chuẩn của mẫu số liệu mới: $\sigma=\sqrt{S^2}=\sqrt{\frac{8000}{441}}=\frac{40\sqrt{5}}{21}\approx 4,26$

Câu 2: Để đánh giá chất lượng một loại pin điện thoại mới, người ta ghi lại thời gian nghe nhạc liên tục của điện thoại được sạc đầy pin cho đến khi hết pin cho kết quả như sau:

Thời gian (giờ)	[5;5,5)	[5,5;6)	[6;6,5)	[6,5;7)	[7;7,5)
Số chiếc điện thoại (tần số)	2	8	15	10	5

Tính độ lệch chuẩn của mẫu số liệu ghép nhóm trên (kết quả được làm tròn đến hàng phần trăm)

Lời giải

Chọn giá trị đại diện cho các nhóm số liệu, ta có:

Thời gian (giờ)	[5;5,5)	[5,5;6)	[6;6,5)	[6,5;7)	[7;7,5)
Giá trị đại diện	5,25	5,75	6,25	6,75	7,25
Số chiếc điện thoại (tần số)	2	8	15	10	5

Thời gian trung bình nghe nhạc liên tục của điện thoại là:

$$\overline{x}=\frac{1}{40}\left(5,25.2+5,75.8+6,25.15+6,75.10+7,25.5\right)=6,35$$

Phương sai của mẫu số liệu là:

$$s^2=\frac{1}{40}\left(5,25^2.2+5,75^2.8+6,25^2.15+6,75^2.10+7,25^2.5\right)-6,35^2=0,2775$$

Độ lệch chuẩn của mẫu số liệu là: $\sqrt{0,2775}=\frac{\sqrt{111}}{20}\approx 0,53$

Câu 3: Bảng dưới đây thống kê cự li ném tạ của một vận động viên.

Cự li (m)	[19;19,5)	[19,5;20)	[20;20,5)	[20,5;21)	[21;21,5)
Tần số	13	45	24	12	6

Hãy tính độ lệch chuẩn của mẫu số liệu ghép nhóm trên (kết quả được làm tròn đến hàng phần trăm)

Lời giải

Giá trị đại diện	19,25	19,75	2025	20,75	21,25
Tần số	13	45	24	12	6

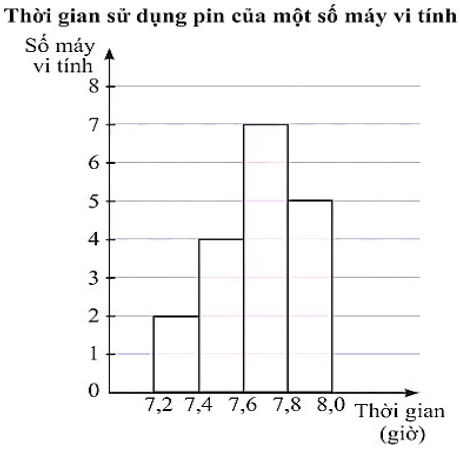
Cỡ mẫu: $n=100$

Số trung bình: $\overline{x}=\frac{13.19,25+45.19,75+24.20,25+12.20,75+6.21,25}{100}=20,015$

Phương sai: $S^2=\frac{13.19,25^2+45.19,75^2+24.20,25^2+12.20,75^2+6.21,25^2}{100}-20,015^2\approx 0,28$

Độ lệch chuẩn: $\sigma=\sqrt{0,28}\approx 0,53$

Câu 4: Kết quả khảo sát thời gian sử dụng liên tục (đơn vị: giờ) từ lúc sạc đầy cho đến khi hết của pin một số máy vi tính cùng loại được mô tả bằng biểu đồ bên.



Hãy xác định độ lệch chuẩn của thời gian sử dụng pin (kết quả được làm tròn đến hàng phần trăm)

Lời giải

Giá trị đại diện	7,3	7,5	7,7	7,9
Số máy	2	4	7	5

Cỡ mẫu: $n = 18$

Số trung bình: $\overline{x} = \dfrac{2.7,3 + 4.7,5 + 7.7,7 + 5.7,9}{18} \approx 7,67$

Phương sai: $S^2 = \dfrac{2.7,3^2 + 4.7,5^2 + 7.7,7^2 + 5.7,9^2}{18} - 7,67^2 \approx 0,04$

Độ lệch chuẩn: $\sigma = \sqrt{0,04} \approx 0,19$

Câu 5: Thành tích môn nhảy cao của các vận động viên tại một giải điền kinh dành cho học sinh trung học phổ thông như sau:

Mức xà (cm)	[170;172)	[172;174)	[174;176)	[176;180)
Số vận động viên	3	10	6	1

Hãy xác định độ lệch chuẩn của thời gian sử dụng pin (kết quả được làm tròn đến hàng phần trăm)

Lời giải

Mẫu số liệu với giá trị đại diện

Mức xà (cm)	[170;172)	[172;174)	[174;176)	[176;178)
Giá trị đại diện	171	173	175	177
Sộ vận động viên	3	10	6	1

Giá trị trung bình: $\overline{x} = \dfrac{1}{20}(171.3 + 173.10 + 175.6 + 177.1) = 173,5$ (cm)

Phương sai của mẫu số liệu: $s^2 = \dfrac{1}{20}\left(171^2.3 + 173^2.10 + 175^2.6 + 177^2.1\right) - 173,5^2 = 2,35$

Độ lệch chuẩn của mẫu số liệu: $s = \sqrt{2,35} = \dfrac{\sqrt{235}}{10} \approx 1,53$ (cm)

Câu 6: Chiều dài của 40 bé trai sơ sinh 12 ngày tuổi chọn ngẫu nhiên ở một bệnh viện được nhà nghiên cứu thống kê trong bảng dưới đây:

Chiều dài (cm)	[44;46)	[46;48)	[48;50)	[50;52)	[52;54)	[54;56)
Số trẻ	3	3	10	15	7	2

Tính độ lệch chuẩn của chiều dài nhóm 40 bé trai sơ sinh (làm tròn kết quả đến hàng phần nghìn).

Lời giải

Bổ sung thêm các giá trị đại diện vào bảng trên ta lập được bảng sau. Chiều dài của 40 bé trai sơ sinh (đơn vị: cm)

Nhóm	c_i	n_i
[44;46)	45	3
[46;48)	47	3
[48;50)	49	10
[50;52)	51	15
[52;54)	53	7
[54;56)	55	2
		$N = 40$

Từ mẫu số liệu đã cho, ta tính được số trung bình là:

$$\overline{x} = \frac{3.45 + 3.47 + 10.49 + 15.51 + 7.53 + 2.55}{40} = \frac{2012}{40} = 50,3.$$

$\overline{x}$ không phải là số nguyên nên để tính phương sai ta tính:

$$\overline{x^2} = \frac{3.45^2 + 3.47^2 + 10.49^2 + 15.51^2 + 7.53^2 + 2.55^2}{40} = 2536.$$

Do đó: $s^2 = \overline{x^2} - (\overline{x})^2 = 2536 - 50,3^2 = 2536 - 2530,09 = 5,91$.

Vậy mẫu số liệu về chiều dài của 40 trẻ sơ sinh có độ lệch chuẩn là $s = \sqrt{5,91} \approx 2,43$.

Câu 7: Một bác tài xế thống kê lại độ dài quãng đường (đơn vị: km) bác đã lái xe mỗi ngày trong một tháng ở bảng sau:

Độ dài quãng đường (km)	[50;100)	[100;150)	[150;200)	[200;250)	[250;300)
Số ngày	5	10	9	9	2

Hãy xác định độ lệch chuẩn của mẫu số liệu trên

Lời giải

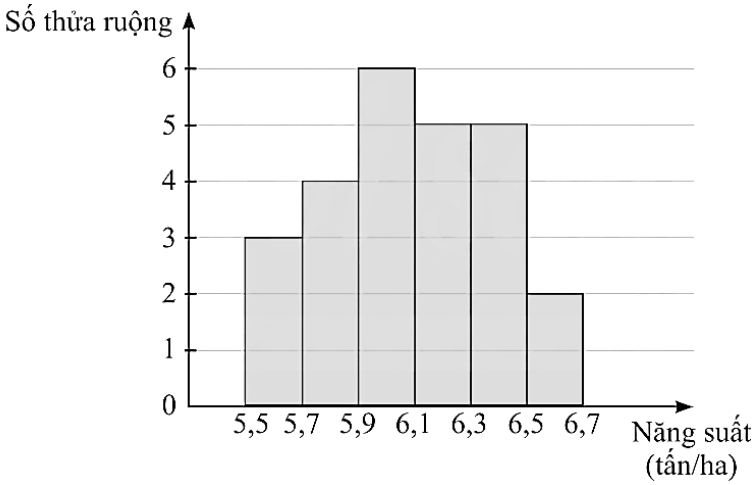
Độ dài quãng đường (km)	75	125	175	225	275
Số ngày	5	10	9	9	2

Số trung bình: $\overline{x} = \frac{5.75 + 10.125 + 9.175 + 4.225 + 2.275}{30} = 155$

Độ lệch chuẩn: $\sigma = \sqrt{\frac{5.75^2 + 10.125^2 + 9.175^2 + 4.225^2 + 2.275^2}{30} - 155^2} = 3100$

Câu 8: Kết quả khảo sát năng suất (đơn vị: tấn/ha) của một số thửa ruộng được minh hoạ ở biểu đồ sau.

Năng suất lúa của một số thửa ruộng



Người ta lập được bảng tần số ghép nhóm và tần số tương đối ghép nhóm tương ứng của mẫu số liệu trên như sau:

Năng suất	[5,5;5,7)	[5,7;5,9)	[5,9;6,1)	[6,1;6,3)	[6,3;6,5)	[6,5;6,7)
Số thửa ruộng	3	4	6	5	5	2
Giá trị đại diện	5,6	5,8	6,0	6,2	6,4	6,6
Tần số tương đối	3	4	6	5	5	2

Hãy xác định độ lệch chuẩn của mẫu số liệu trên (làm tròn kết quả đến hàng phần nghìn).

Lời giải

Số trung bình: $\overline{x}=\dfrac{3.5,6+4.5,8+6.6,0+5.6,2+5.6,4+2.6,6}{25}=6,088$

Độ lệch chuẩn: $\sigma=\sqrt{\dfrac{5.75^2+10.125^2+9.175^2+4.225^2+2.275^2}{30}-155^2}\approx 0,29$

Câu 9: Một vận động viên luyện tập chạy cự li 100 m đã ghi lại kết quả luyện tập như sau:

Thời gian (giây)	[10,2;10,4)	[10,4;10,6)	[10,6;10,8)	[10,8;11)
Số vận động viên	3	7	8	2

Hãy xác định phương sai của mẫu số liệu trên?

Lời giải

Chọn giá trị đại diện cho các nhóm số liệu, ta có:

Thời gian (giây)	[10,2;10,4)	[10,4;10,6)	[10,6;10,8)	[10,8;11)
Giá trị đại diện	10,3	10,5	10,7	10,9
Số vận động viên	3	7	8	2

Tổng số vận động viên là: $3+7+8+2=20$

Thời gian chạy trung bình của các vận động viên là:

$$\overline{x} = \frac{1}{20}(10,3.3 + 10,5.7 + 10,7.8 + 10,9.2) = 10,59 \text{ (giây)}$$

Phương sai của mẫu số liệu là:

$$s^2 = \frac{1}{20}\left(10,3^2.3 + 10,5^2.7 + 10,7^2.8 + 10,9^2.2\right) - 10,59^2 = 0,0299$$

Câu 10: Cân nặng của một số quả mít trong một khu vườn được thống kê ở bảng sau:

Cân nặng (kg)	[4;6)	[6;8)	[8;10)	[10;12)	[12;14)
Số quả mít	6	12	19	9	4

Hãy tính phương sai của mẫu số liệu ghép nhóm trên. (Kết quả các phép tính làm tròn đến hàng phần trăm)?

Lời giải

Ta có bảng thống kê cân nặng của các quả mít theo giá trị đại diện:

Cân nặng đại diện (kg)	5	7	9	11	13
Tần số	6	12	19	9	4

Cỡ mẫu $n = 6 + 12 + 19 + 9 + 4 = 50$.

Số trung bình của mẫu số liệu ghép nhóm là: $\overline{x} = \frac{6.5 + 12.7 + 19.9 + 9.11 + 4.13}{50} = 8,72.$

Phương sai của mẫu số liệu ghép nhóm là:

$$S^2 = \frac{1}{50}\left(6.5^2 + 12.7^2 + 19.9^2 + 9.11^2 + 4.13^2\right) - 8,72^2 \approx 4,80.$$

B. BÀI TẬP TỰ LUẬN

Bài tập 1: Tìm cân nặng trung bình của học sinh lớp 11D cho trong bảng sau

Cân nặng	[40,5;45,5)	[45,5;50,5)	[50,5;55,5)	[55,5;60,5)	[60,5;65,5)	[65,5;70,5)
Số học sinh	10	7	16	4	2	3

Lời giải

Trong mỗi khoảng cân nặng, giá trị đại diện là trung bình cộng của hai giá trị đầu mút nên ta có bảng sau:

Cân nặng (kg)	43	48	53	58	63	68
Số học sinh	10	7	16	4	2	3

Tổng số học sinh là $n = 42$. Cân nặng trung bình cú lớp 11D là:

$$\bar{x} = \frac{10.43 + 7.48 + 16.53 + 4.58 + 2.63 + 3.68}{42} \approx 51,81 \text{ (kg)}$$

Bài tập 2: Kết quả khảo sát cân nặng của 25 quả cam ở mỗi lô hàng A và B được cho bởi bảng sau:

Cân nặng (g)	[150; 155)	[155; 160)	[160; 165)	[165; 170)	[170; 175)
Số quả cam ở lô hàng A	2	6	12	4	1
Số quả cam ở lô hàng B	1	3	7	10	4

a) Hãy ước lượng cân nặng trung bình của mỗi quả cam ở lô hàng A và B

b) Nếu so sánh theo số trung bình thì cam ở lô hàng nào nặng hơn?

Lời giải

Ta có bảng thống kê số lượng cam theo giá trị đại diện:

Cân nặng (g)	152,5	157,5	162,5	167,5	172,5
Số quả cam ở lô hàng A	2	6	12	4	1
Số quả cam ở lô hàng B	1	3	7	10	4

a) Cân nặng trung bình của một quả cam ở lô hàng A là xấp xỉ bằng:

$$\frac{(2.152,5 + 6.157,5 + 12.162,5 + 4.167,5 + 1.172,5)}{25} = 161,7 \text{ (g)}$$

Cân nặng trung bình của một quả cam ở lô hàng B là xấp xỉ bằng:

$$\frac{(1.152,2 + 3.157,5 + 7.162,5 + 10.167,5 + 4.172,5)}{25} = 165,1 \text{ (g)}$$

b) Nếu so sánh theo số trung bình thì cam ở lô hàng B nặng hơn cam ở lô hàng A.

Bài tập 3: Cân nặng của một số quả mít trong một khu vườn được thống kê ở bảng sau:

Cân nặng (kg)	[4;6)	[6;8)	[8;10)	[10;12)	[12;14)
Số quả mít	6	12	19	9	4

Hãy tính phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm trên. (Kết quả các phép tính làm tròn đến hàng phần trăm.)

Lời giải

Ta có bảng thống kê cân nặng của các quả mít theo giá trị đại diện:

Cân nặng đại diện (kg)	5	7	9	11	13
Số quả mít	6	12	19	9	4

Cỡ mẫu $n = 6 + 12 + 19 + 9 + 4 = 50$.

Số trung bình của mẫu số liệu ghép nhóm là: $\bar{x} = \frac{6.5 + 12.7 + 19.9 + 9.11 + 4.13}{50} = 8,72.$

Phương sai của mẫu số liệu ghép nhóm là

$$S^2=\frac{1}{50}\left(6\cdot 5^2+12\cdot 7^2+19.9^2+9\cdot 11^2+4.13^2\right)-8,72^2\approx 4,80$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S\approx\sqrt{4,80}\approx 2,19.$

Bài tập 4: Thống kê tổng số giờ nắng trong tháng 9 tại một trạm quan trắc đặt ở Cà Mau trong các năm từ 2002 đến 2021 được thống kê như sau:

111,6	134,9	130,3	134,2	140,9	109,3	154,4	156,3	116,1	96,7
105,2	80,8	80,8	110	109	139	145	161	126	114

a) Hãy tính phương sai và độ lệch chuẩn của mẫu số liệu trên.

b) Hãy lập bảng tần số ghép nhóm với nhóm đầu tiên là $[80;98)$ và độ dài mỗi nhóm bằng 18. Tính phương sai, độ lệch chuẩn của mẫu số liệu ghép nhóm.

c) Hãy tính sai số tương đối của độ lệch chuẩn của mẫu số liệu ghép nhóm so với độ lệch chuẩn của mẫu số liệu gốc.

(Kết quả các phép tính làm tròn đến hàng phần nghìn.)

Lời giải

a) Cỡ mẫu là $n=20$.

Số trung bình của mẫu số liệu trên là: $\overline{x_1}=\frac{111,6+134,9+\ldots+114}{20}=122,755.$

Phương sai của mẫu số liệu trên là: $S_1^2=\frac{1}{20}\left(111,6^2+134,9^2+\ldots+114^2\right)-122,755^2\approx 515,453.$

Độ lệch chuẩn của mẫu số liệu trên là: $S_1\approx\sqrt{515,453}\approx 22,704.$

b) Ta có bảng sau:

Số giờ nắng	$[80;98)$	$[98;116)$	$[116;134)$	$[134;152)$	$[152;170)$
Giá trị đại diện	89	107	125	143	161
Số năm	3	6	3	5	3

Số trung bình của mẫu số liệu ghép nhóm là: $\overline{x_2}=\frac{3.89+6.107+3.125+5.143+3.161}{20}=124,1.$

Phương sai của mẫu số liệu ghép nhóm là

$$S_2^2=\frac{1}{20}\left(3.89^2+6.107^2+3.125^2+5.143^2+3.161^2\right)-124,1^2=566,19.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S_2=\sqrt{566,19}\approx 23,795.$

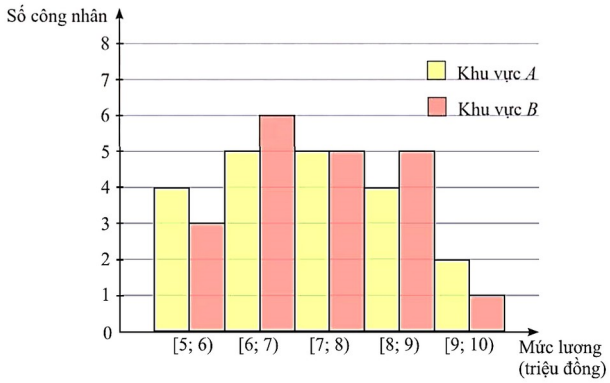
c) Sai số tương đối của độ lệch chuẩn của mẫu số liệu ghép nhóm so với độ lệch chuẩn của mẫu số liệu gốc là: $\frac{|S_2-S_1|}{S_1}=\frac{|23,795-22,704|}{22,704}\cdot 100\%\approx 4,805\%.$

Bài tập 5: Biểu đồ dưới đây mô tả kết quả điều tra về mức lương khởi điểm (đơn vị: triệu đồng) của một số công nhân ở hai khu vực A và B.

a) Hãy xác định giá trị đại diện cho mỗi nhóm và lập bảng tần số ghép nhóm cho mẫu số liệu đó.

b) Nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì công nhân ở khu vực nào có mức lương khởi điểm đồng đều hơn?

Mức lương khởi điểm của công nhân ở hai khu vực A và B



Lời giải

a) Ta có bảng sau:

Mức lương	$[5;6)$	$[6;7)$	$[7;8)$	$[8;9)$	$[9;10)$
Mức lương đại diện (triệu đồng)	5,5	6,5	7,5	8,5	9,5
Khu vực A	4	5	5	4	2
Khu vực B	3	6	5	5	1

b) Xét mẫu số liệu của khu vực A :

Cỡ mẫu là $n_A = 4+5+5+4+2 = 20$.

Số trung bình của mẫu số liệu ghép nhóm là: $\overline{x}_A = \dfrac{4\cdot 5,5+5.6,5+5\cdot 7,5+4.8,5+2.9,5}{20} = 7,25.$

Phương sai của mẫu số liệu ghép nhóm là

$$S_A^2 = \frac{1}{20}\left(4\cdot 5,5^2+5\cdot 6,5^2+5\cdot 7,5^2+4\cdot 8,5^2+2\cdot 9,5^2\right)-(7,25)^2 = 1,5875.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S_A = \sqrt{1,5875}.$

Xét mẫu số liệu của khu vực B :

Cỡ mẫu là $n_B = 3+6+5+5+1 = 20$.

Số trung bình của mẫu số liệu ghép nhóm là: $\overline{x}_B = \dfrac{3\cdot 5,5+6.6,5+5\cdot 7,5+5.8,5+1.9,5}{20} = 7,25.$

Phương sai của mẫu số liệu ghép nhóm là

$$S_B^2 = \frac{1}{20}\left(3\cdot 5,5^2+6\cdot 6,5^2+5\cdot 7,5^2+5\cdot 8,5^2+1\cdot 9,5^2\right)-(7,25)^2 = 1,2875.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là: $S_B = \sqrt{1,2875}$

Do $S_A > S_B$ nên nếu so sánh theo độ lệch chuẩn của mẫu số liệu ghép nhóm thì mức lương khởi điểm của công nhân khu vực B đồng đều hơn của công nhân khu vực A.

Bài tập 6: Giá đóng cửa của một cổ phiếu là giá của cổ phiếu đó cuối một phiên giao dịch. Bảng sau thống kê giá đóng cửa (đơn vị: nghìn đồng) của hai mã cổ phiếu A và B trong 50 ngày giao dịch liên tiếp.

Giá đóng cửa	[120;122)	[122;124)	[124;126)	[126;128)	[128;130)
Cổ phiếu A	8	9	12	10	11
Cổ phiếu B	16	4	3	6	21

Người ta có thể dùng phương sai và độ lệch chuẩn để so sánh mức độ rủi ro của các loại cổ phiếu có giá trị trung bình gần bằng nhau. Cổ phiếu nào có phương sai, độ lệch chuẩn cao hơn thì được coi là có độ rủi ro lớn hơn.

Theo quan điểm trên, hãy so sánh độ rủi ro của cổ phiếu A và cổ phiếu B.

Lời giải

Ta có bảng thống kê giá đóng cửa theo giá trị đại diện:

Giá đóng cửa	121	123	125	127	129
Cổ phiếu A	8	9	12	10	11
Cổ phiếu B	16	4	3	6	21

Xét mẫu số liệu của cổ phiếu A:

Số trung bình của mẫu số liệu ghép nhóm là

$$\overline{x_1} = \frac{8.121 + 9.123 + 12.125 + 10.127 + 11.129}{50} = 125,28.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_1^2 = \frac{1}{50}\left(8.121^2 + 9.123^2 + 12.125^2 + 10.127^2 + 11.129^2\right) - (125,28)^2 = 7,5216$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_1 = \sqrt{S_1^2} = \sqrt{7,5216}$.

Xét mẫu số liệu của cổ phiếu B:

Số trung bình của mẫu số liệu ghép nhóm là

$$\overline{x_2} = \frac{16.121 + 4.123 + 3.125 + 6.127 + 21.129}{50} = 125,28.$$

Phương sai của mẫu số liệu ghép nhóm là

$$S_2^2 = \frac{1}{50}\left(16.121^2 + 4.123^2 + 3.125^2 + 6.127^2 + 21.129^2\right) - (125,48)^2 = 12,4096.$$

Độ lệch chuẩn của mẫu số liệu ghép nhóm là $S_2 = \sqrt{S_2^2} = \sqrt{12,4096}$.

Vậy nếu đánh giá độ rủi ro theo phương sai và độ lệch chuẩn thì cổ phiếu A có độ rủi ro thấp hơn cổ phiếu B.

Bài tập 7: Mai và Ngọc cùng sử dụng vòng đeo tay thông minh để ghi lại số bước chân hai bạn đi mỗi ngày trong một tháng. Kết quả được ghi lại ở bảng sau:

Giá trị đại diện	4	6	8	10	12
Mai	6	7	6	6	5
Ngọc	2	5	10	8	2

a) Hãy tính số trung bình và độ lệch chuẩn của mẫu số liệu ghép nhóm trên.

b) Nếu so sánh theo độ lệch chuẩn thì bạn nào có số lượng bước chân đi mỗi ngày đều đặn hơn?

Lời giải

a) Cỡ mẫu: $n = 30$

Xét mẫu số liệu của Mai:

Số trung bình: $\overline{x_1} = \frac{6.4 + 7.6 + 6.8 + 6.10 + 5.12}{30} = 7,8$

Phương sai: $S_1^2 = \frac{\left(6.4^2 + 7.6^2 + 6.8^2 + 6.10^2 + 5.12^2\right)}{30} - 7,8^2 = 7,56$

Độ lệch chuẩn: $\sigma_1 = \sqrt{7,56} \approx 2,75$

Xét mẫu số liệu của Ngọc:

Số trung bình: $\overline{x_2} = \frac{2.4 + 5.6 + 13.8 + 8.10 + 2.12}{30} = 8,2$

Phương sai: $S_2^2 = \frac{\left(2.4^2 + 5.6^2 + 13.8^2 + 8.10^2 + 2.12^2\right)}{30} - 8,2^2 \approx 3,83$

Độ lệch chuẩn: $\sigma_2 = \sqrt{3,83} \approx 1,96$

b) Nếu so sánh theo độ lệch chuẩn thì bạn Ngọc có số lượng bước chân đi mỗi ngày đều đặn hơn.

Bài tập 8: Bảng dưới đây thống kê cự li ném tạ của một vận động viên.

Cự li (m)	[19;19,5)	[19,5;20)	[20;20,5)	[20,5;21)	[21;21,5)
Tần số	13	45	24	12	6

Hãy tính phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm trên.

Lời giải

Giá trị đại diện	19,25	19,75	20,25	20,75	21,25
Tần số	13	45	24	12	6

Cỡ mẫu: $n = 100$

Số trung bình: $\overline{x} = \frac{13.19,25 + 45.19,75 + 24.20,25 + 12.20,75 + 6.21,25}{100} = 20,015$

Phương sai: $S^2 = \frac{13.19,25^2 + 45.19,75^2 + 24.20,25^2 + 12.20,75^2 + 6.21,25^2}{100} - 20,015^2 \approx 0,28$

Độ lệch chuẩn: $\sigma = \sqrt{0,28} \approx 0,53$