

CHƯƠNG



CÁC SỐ ĐẶC TRƯNG ĐO MỨC ĐỘ PHÂN TÁN CỦA MẪU SỐ LIỆU GHÉP NHÓM



HỆ THỐNG BÀI TẬP TOÁN THỰC TẾ.

Câu 1: Bảng 22, Bảng 23 lần lượt 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: $^{\circ}\text{C}$)

Nhóm	Giá trị đại diện	Tần số
[16,8; 19,8)	18,3	2
[19,8; 22,8)	21,3	3
[22,8; 25,8)	24,3	2
[25,8; 28,8)	27,3	1
[28,8; 31,8)	30,3	4
		$n = 12$

Bảng 22

Nhóm	Giá trị đại diện	Tần số
[16,8; 19,8)	18,3	1
[19,8; 22,8)	21,3	2
[22,8; 25,8)	24,3	3
[25,8; 28,8)	27,3	2
[28,8; 31,8)	30,3	4
		$n = 12$

Bảng 23

(Nguồn: Niên giám thống kê 2021, NXB Thống kê, 2022)

- Tính khoảng biến thiên, khoảng tứ phân vị, phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm của Hà Nội và Huế.
- Trong hai thành phố Hà Nội và Huế, thành phố nào có nhiệt độ không khí trung bình tháng đồng đều hơn?

Câu 2: Bảng 24 thống kê độ ẩm không khí trung bình các tháng 2021 tại Đà Lạt và Vũng Tàu (đơn vị: %)

Tháng	1	2	3	4	5	6	7	8	9	10	11	12
Độ ẩm												
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

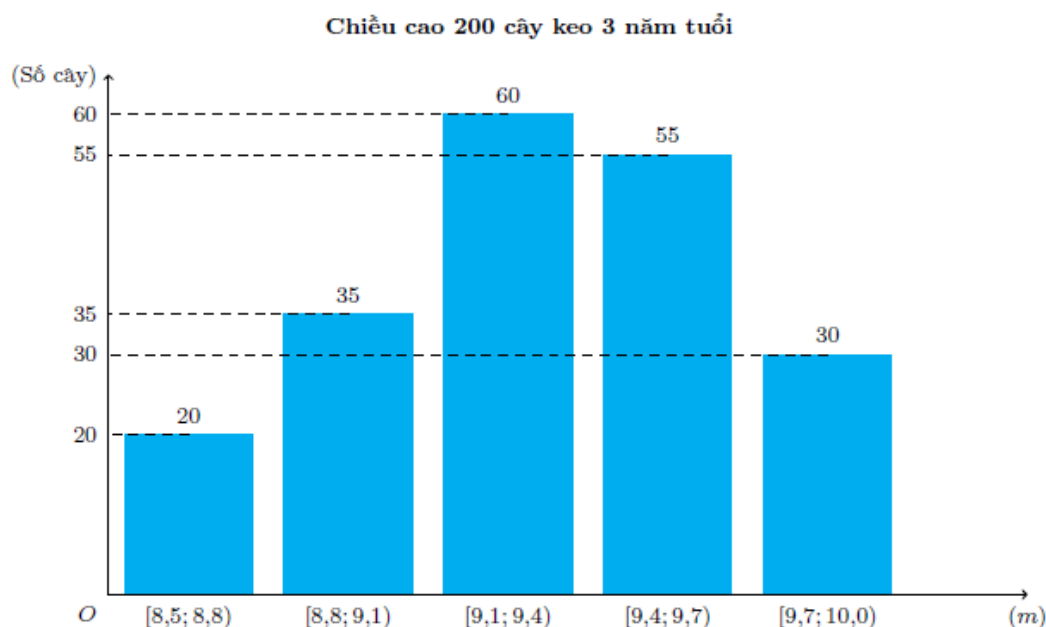
- Hãy lần lượt ghép lớp các số liệu của Đà Lạt, Vũng Tàu thành năm nhóm sau:
 $[75; 78,3), [78,3; 81,6), [81,6; 84,9), [84,9; 88,2), [88,2; 91,5)$.
- Tính khoảng biến thiên, khoảng tứ phân vị, phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm của Đà Lạt và Vũng Tàu.
- Trong hai thành phố Đà Lạt và Vũng Tàu, thành phố nào có độ ẩm không khí trung bình tháng đồng đều hơn?

Câu 3: Bảng sau cho ta bảng tần số ghép nhóm về số liệu thống kê chiều cao (đơn vị: mét) của 40 núi cao nhất Đông Nam Á. Độ lệch chuẩn của mẫu số liệu đó bằng bao nhiêu (làm tròn kết quả đến hàng đơn vị)?

Nhóm	Tần số
$[3500; 4000)$	10
$[4000; 4500)$	7
$[4500; 5000)$	16
$[5000; 5500)$	4
$[5500; 6000)$	3
	$n = 40$

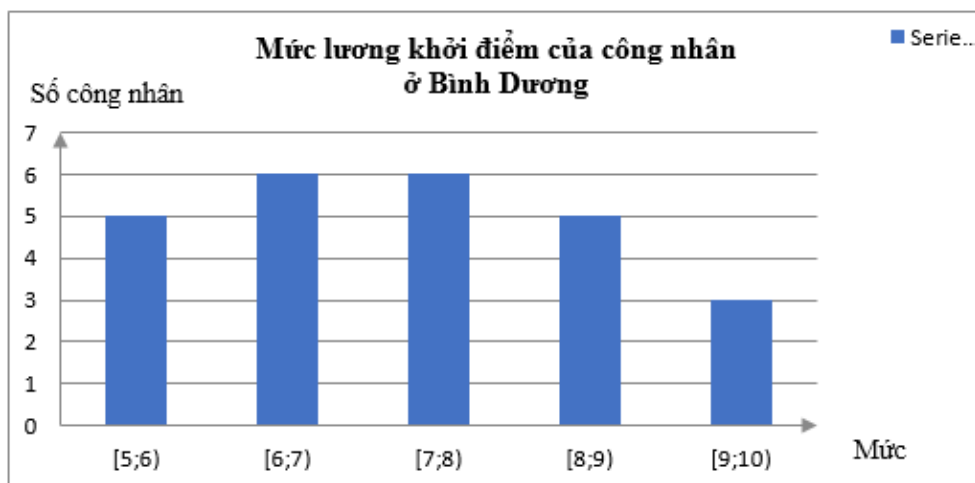
(Nguồn: <https://vi.wikipedia.org>)

Câu 4: Kết quả đo chiều cao của 200 cây keo 3 năm tuổi ở một nông trường được biểu diễn ở biểu đồ dưới đây



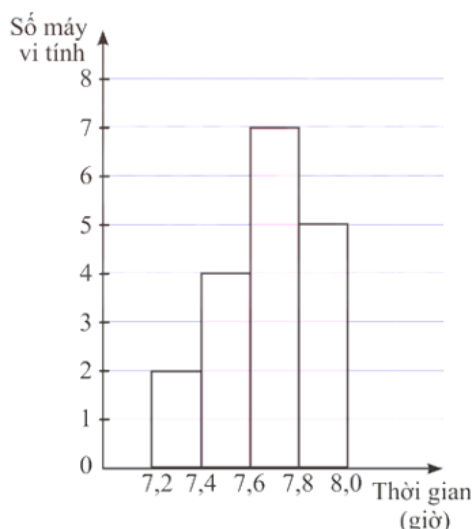
Hãy tìm khoảng tứ phân vị của mẫu số liệu ghép nhóm cho bởi biểu đồ trên.

Câu 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 Bình Dương.



Tính phương sai của mẫu số liệu ghép nhóm.

Câu 6: 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)

Câu 7: Một giống cây xoan đào được trồng tại hai địa điểm A và B. Người ta thống kê đường kính thân của một số cây xoan đào 5 năm tuổi ở bảng sau:

Đường kính (cm)	[30; 32)	[32; 34)	[34; 36)	[36; 38)	[38; 40)
Số cây trồng ở địa điểm A	25	38	20	10	7
Số cây trồng ở địa điểm B	22	27	19	18	14

a) Hãy so sánh đường kính trung bình của thân cây xoan đào trồng tại địa điểm A và địa điểm B.
b) Nếu so sánh theo độ lệch chuẩn thì cây trồng tại địa điểm nào có đường kính đồng đều hơn?

Câu 8: 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)

Câu 9: Tổng lượng mưa trong tháng 8 đo được tại một trạm quan trắc đặt tại Vũng Tàu từ năm 2002 đến năm 2020 được ghi lại như dưới đây (đơn vị: mm):

121,8 158,3 334,9 200,9 165,6 161,5 194,3 220,7 189,8 234,2
165,9 165,9 134 173 169 189 254 168 255

(Nguồn: Tổng cục thống kê)

a) Hoàn thiện bảng tần số ghép nhóm theo mẫu sau:

Tổng lượng mưa trong tháng 8 (mm)	[120; 175)	[175; 230)	[230; 285)	[285; 340)
Số năm	?	?	?	?

b) Hãy ước lượng số trung bình, tứ phân vị và một của mẫu số liệu ở bảng tần số ghép nhóm trên.

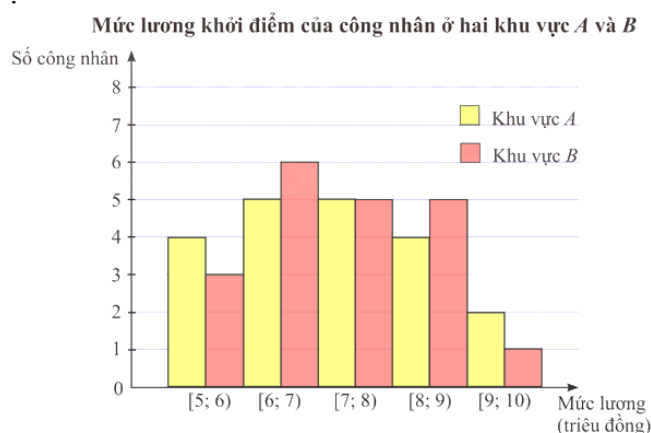
Câu 10: Bảng sau thống kê lại tổng số giờ nắng trong tháng 6 của các năm từ 2002 đến 2021 tại hai trạm quan trắc đặt ở Nha Trang và Quy Nhơn.

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

(Nguồn: Tổng cục Thống kê)

- Xét số liệu ở Nha Trang thì khoảng tứ phân vị của mẫu số liệu ghép nhóm là: 32,64
- 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 Quy Nhơn ta có độ lệch chuẩn của mẫu số liệu ghép nhóm (làm tròn kết quả đến hàng phần trăm) là: 30,59
- Nếu so sánh theo độ lệch chuẩn thì số giờ nắng trong tháng 6 của Nha Trang đồng đều hơn

Câu 11: 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

- Khoảng biến thiên của mẫu số liệu ghép nhóm là 4.
- Xét mẫu số liệu của khu vực A ta có khoảng tứ phân vị của mẫu số liệu ghép nhóm là 2,05.
- Xét mẫu số liệu của khu vực B ta có phương sai của mẫu số liệu ghép nhóm là 1,5875.
- 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 12: Bảng thống kê thời gian (đơn vị: phút) tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình và bạn An

Thời gian (phút)	[15; 20)	[20; 25)	[25; 30)	[30; 35)	[35; 40)
Số ngày tập của bạn Bình	5	10	10	2	1
Số ngày tập của bạn An	5	5	15	3	0

- Khoảng biến thiên của mẫu số liệu về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn An là 20.
- Khoảng tứ phân vị của mẫu số liệu về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình là 28.
- Khoảng tứ phân vị của mẫu số liệu về thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn An là 22.
- Dựa vào khoảng tứ phân vị của hai mẫu số liệu trên thì thời gian tập thể dục buổi sáng mỗi ngày trong tháng 2 năm 2023 của bạn Bình phân tán hơn bạn An.

Câu 13: Một công ty cung cấp nước sạch thống kê lượng nước các hộ gia đình trong một khu vực tiêu thụ trong một tháng ở bảng sau:

Lượng nước tiêu thụ (m^3)	[3; 6)	[6; 9)	[9; 12)	[12; 15)	[15; 18)
Số hộ gia đình	24	57	42	29	8

- Lượng nước tiêu thụ trung bình trong tháng của một hộ gia đình trong khu vực nói trên xấp xỉ bằng $9,4 (m^3)$.
- Khoảng biến thiên của mẫu số liệu trên là $R = 15 (m^3)$.
- Công ty muốn gửi một thông báo khuyến nghị tiết kiệm nước đến 25% các hộ gia đình có lượng nước tiêu thụ cao nhất. Vậy công ty nên gửi thông báo tiết kiệm nước đến các hộ gia đình có lượng nước tiêu thụ từ $11 (m^3)$ nước trở lên.
- Độ lệch chuẩn của mẫu số liệu trên bằng $3 (m^3)$.

Câu 14: 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

- Khoảng tứ phân vị của mẫu số liệu ghép nhóm trạm quan trắc ở Nha Trang bằng 45.
- Nếu so sánh theo khoảng tứ phân vị thì số giờ nắng trong tháng 6 của Quy Nhơn đồng đều hơn Nha Trang.
- Số trung bình của mẫu số liệu ghép nhóm trạm quan trắc ở Quy Nhơn bằng 242,5.
- Nếu so sánh theo độ lệch chuẩn thì số giờ nắng trong tháng 6 của Quy Nhơn đồng đều hơn Nha Trang.

Câu 15: Thống kê điểm thi khảo sát đầu năm môn Toán của hai lớp 12A và 12B, ta thu được kết quả sau:

Điểm thi	[5; 6)	[6; 7)	[7; 8)	[8; 9)	[9; 10]
Số học sinh lớp 12A	0	2	6	12	10
Số học sinh lớp 12B	2	12	10	6	0

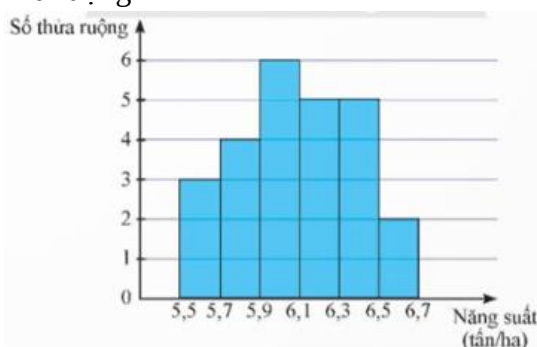
- Từ khoảng biến thiên của điểm thi của học sinh hai lớp 12A và 12B, điểm thi khảo sát môn Toán của lớp 12A phân tán hơn của lớp 12B.
- Số điểm trung bình môn Toán trong bài khảo sát đầu năm của lớp 12B lớn hơn của lớp 12A.
- Khoảng tứ phân của lớp 12A lớn hơn 1.
- Khoảng tứ phân vị của lớp 12A lớn hơn so với lớp 12B.

Câu 16: Thống kê lương mới từ 01/07/2024 của giáo viên hạng III của một trường THPT thu được kết quả sau:

Lương (triệu)	[4; 6)	[6; 8)	[8; 10)	[10; 12]
Số lượng GV	6	20	30	5

- Trong thống kê số lượng giáo viên có mức lương cao nhất có số lượng thấp nhất.
- Lương trung bình của giáo viên hạng III trong thống kê là 10.
- Nhóm tứ phân vị thứ hai của thống kê là nhóm [6; 8).
- Khoảng tứ phân vị thống kê là nhỏ hơn 1

Câu 17: Kết quả khảo sát năng suất (đơn vị: tấn/ha) của một số thửa ruộng được minh họa ở biểu đồ sau: Năng suất lúa của một số thửa ruộng



- Có 25 thửa ruộng đã được khảo sát.
- Khoảng biến thiên của mẫu số liệu trên là 1,2 (tấn/ha).
- Khoảng tứ phân vị của mẫu số liệu ghép nhóm trên là 0,4675.
- Số trung bình của mẫu số liệu ghép nhóm trên là 6,088.

Câu 18: 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

- Khoảng biến thiên của hai mẫu số liệu trên là 180.
- Số trung bình của mẫu số liệu ghép nhóm trạm quan trắc ở Quy Nhơn bằng 242,5.
- Khoảng tứ phân vị của mẫu số liệu ghép nhóm trạm quan trắc ở Nha Trang bằng 45.
- Nếu so sánh theo khoảng tứ phân vị thì số giờ nắng trong tháng 6 của Quy Nhơn đồng đều hơn Nha Trang.

Câu 19: Mỗi ngày bác Hương đều đi bộ để rèn luyện sức khỏe. 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

- Khoảng biến thiên của mẫu số liệu ghép nhóm là 0,9.
- Khoảng tứ phân vị của mẫu số liệu ghép nhóm là 0,575.
- Phương sai của mẫu số liệu ghép nhóm là 0,36.
- Độ lệch chuẩn của mẫu số liệu ghép nhóm có giá trị gần bằng 0,36.

Câu 20: Một bác tài xế taxi 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)	[60; 80)	[80; 100)	[100; 120)	[120; 140)	[140; 160)
Số ngày	3	9	10	6	2

- Khoảng biến thiên của mẫu số liệu ghép nhóm là 100 (km).
- Khoảng tứ phân vị của mẫu số liệu ghép nhóm gần bằng $33,67$.
- Số trung bình của mẫu số liệu ghép nhóm là $101,27$.
- Độ lệch chuẩn của mẫu số liệu ghép nhóm gần bằng $30,68$.

Câu 21: Một giống cây gỗ Trắc được trồng tại hai quả đồi. Người ta thống kê đường kính thân của một số gỗ Trắc 5 năm tuổi ở bảng sau:

Đường kính (cm)	[30; 32)	[32; 34)	[34; 36)	[36; 38)	[38; 40)
Số cây trồng ở quả đồi thứ nhất	25	38	20	10	7
Số cây trồng ở quả đồi thứ hai	22	27	19	18	14

- Đường kính trung bình của các cây Trắc ở quả đồi thứ nhất sau 5 năm trồng là $\bar{x}_1 = 32,82cm$.
- Đường kính trung bình của các cây Trắc ở quả đồi thứ hai sau 5 năm trồng là $\bar{x}_2 = 34,5m$
- Đường kính trung bình của các cây Trắc ở quả đồi thứ nhất sau 5 năm trồng cao hơn ở quả đồi thứ hai.
- Cây gỗ Trắc trồng trên quả đồi thứ nhất có đường kính đồng đều hơn trên quả đồi thứ hai.

Câu 22: 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

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

- Số trung bình của mẫu số liệu ghép nhóm là 124,1.
- Phương sai của mẫu số liệu ghép nhóm là 566,19.
- Độ lệch chuẩn của mẫu số liệu ghép nhóm (kết quả các phép tính làm tròn đến hàng phần nghìn) là 23,795.
- 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%.

Câu 23: 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

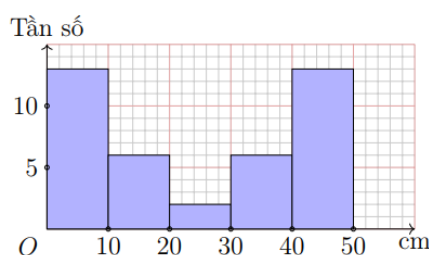
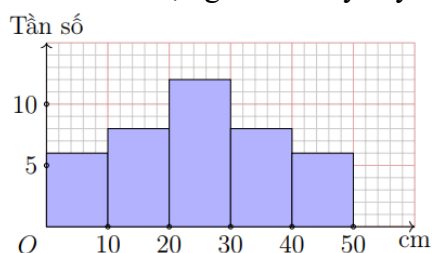
- Xét mẫu số liệu của cổ phiếu A ta có phương sai của mẫu số liệu ghép nhóm là $7,5216$.
- Xét mẫu số liệu của cổ phiếu B ta có số trung bình của mẫu số liệu ghép nhóm là $115,28$.
- Xét mẫu số liệu của cổ phiếu B ta có độ lệch chuẩn của mẫu số liệu ghép nhóm là $\sqrt{15,4096}$.
- Người ta có thể dùng 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.

Câu 24: Một giáo viên ghi lại điểm kiểm tra học kì II môn Toán của hai lớp 12A và 12B có kết quả ghi lại ở bảng sau:

Điểm kiểm tra	[2,5;4)	[4;5;5)	[5;5;7)	[7;8;5)	[8;5;10]
Số học sinh lớp 12A	3	16	9	6	4
Số học sinh lớp 12B	0	15	13	4	6

- Sĩ số học sinh hai lớp 12A và 12B bằng nhau.
- Điểm kiểm tra trung bình của hai lớp chênh lệch nhau không quá 0,5 điểm.
- Phương sai của mẫu số liệu của lớp 12B là 2,54 (kết quả làm tròn đến hàng phần trăm).
- Nếu xét theo độ lệch chuẩn thì điểm kiểm tra của lớp 12B ít phân tán hơn điểm kiểm tra của lớp 12A.

Câu 25: 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 ở cả hai phương pháp như sau:



Chiều cao của cây được chăm sóc theo phương pháp A Chiều cao của cây được chăm sóc theo phương pháp B

- Chiều cao trung bình của các cây được chăm sóc theo phương án A là 24,8 cm.
- Độ lệch chuẩn của chiều cao các cây được chăm sóc theo phương án A lớn hơn so với các cây được chăm sóc theo phương án B.
- Chiều cao trung bình của cây hướng dương ở cả hai cách chăm sóc là như nhau.
- Chiều cao của các cây được chăm sóc theo phương án A ít bị chênh lệch hơn so với chiều cao của các cây được chăm sóc theo phương án B.