§4. PHƯƠNG SAI VÀ ĐỘ LỆCH CHUẨN

A. KIẾN THỨC CẦN NHỚ

Khi hai dãy số liệu thống kê có cùng đơn vị đo và có số trung bình bằng nhau (hay xấp xỉ nhau) thì người ta thường quan tâm đến mức độ đồng đều của các số liệu so với số trung bình và để làm điều đó người ta tính và so sánh *độ lệch chuẩn* (hay phương sai) của dãy số liệu. Dãy số liệu nào có *độ lệch chuẩn* (phương sai) *càng nhỏ thì mức độ phân tán càng nhỏ, do đó càng có độ đồng đều cao.*

Độ lệch chuẩn s và phương sai của dãy số liệu được cho bởi các công thức định nghĩa sau :

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{k} n_i(x_i - \overline{x})^2} \quad (1) \text{ - với bảng phân bố tần số ;}$$

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{k} n_i(c_i - \overline{x})^2} \quad (2) \text{ - với bảng phân bố tần số ghép lớp ;}$$

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{k} f_i(x_i - \overline{x})^2} \quad (1') \text{ - với bảng phân bố tần suất ;}$$

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{k} f_i(c_i - \overline{x})^2} \quad (2') \text{ - với bảng phân bố tần suất ghép lớp ;}$$

hoặc $s = \sqrt{\overline{x^2} - (\overline{x})^2}$. (3)

Phương sai s^2 của dãy số liệu là bình phương của độ lệch chuẩn s.

Trong thực hành, nếu tính toán thủ công không dùng máy tính cầm tay thì ta thường sử dụng các công thức trên để tính phương sai trước (biểu thức trong dấu căn) rồi khai căn bậc hai kết quả để tính độ lệch chuẩn. Ta cũng có thể dùng máy tính cầm tay để tính nhanh chóng độ lệch chuẩn, phương sai của dãy số liệu thống kê. Chẳng hạn, nếu sử dụng máy Casio fx-570ES thì ta làm như sau :

Sau khi nhập các số liệu thống kê (các giá trị và tần số hay tần suất), ta ấn liên tiếp các phím SHIFT, 1 (STAT), 5 (Var), 3 ($\chi\sigma n$), =, ta được kết quả là s. (Nếu muốn tính s^2 ta ấn tiếp phím x^2, = .)

B. BÀI TẬP MẪU

a) Sử dụng các công thức (1'), (2') và (3) để tính phương sai và độ lệch chuẩn (chính xác đến hàng phần trăm) của các số liệu thống kê cho trong bảng 1 (Bài tập mẫu, tiết 1).

b) Sử dụng máy tính cầm tay kiểm tra lại các kết quả tìm được ở câu a).

c) So sánh thành tích chạy 50 m của lớp 10A (kết quả cho trong bảng 1) với thành tích chạy 50 m của lớp 10B cùng trường, biết rằng bảng số liệu thống kê kết quả chạy của lớp 10B có giá trị trung bình là 7,5 giây và có độ lệch chuẩn là 0,71.

Giải

a) Sử dụng kết quả trong bài tập mẫu tiết 3, ta có $\overline{x} = 7{,}507575758 \approx 7{,}51$.

Sử dụng bảng phân bố tần số ghép lớp (bảng 2) và công thức (1') ta có :

$$s^2 = \frac{1}{33}[2.(6{,}25 - 7{,}51)^2 + 5.(6{,}75 - 7{,}51)^2 + 10.(7{,}25 - 7{,}51)^2 +$$

$$+ 9.(7{,}75 - 7{,}51)^2 + 4.(8{,}25 - 7{,}51)^2 + 3.(8{,}75 - 7{,}51)^2]$$

$$= 0{,}426084848 \approx 0{,}43.$$

Sử dụng bảng phân bố tần suất ghép lớp (bảng 3) và công thức (2') ta có :

$$s^2 = \frac{1}{100}[6{,}06.(6{,}25 - 7{,}51)^2 + 15{,}15.(6{,}75 - 7{,}51)^2 + 30{,}30.(7{,}25 - 7{,}51)^2$$

$$+ 27{,}27.(7{,}75 - 7{,}51)^2 + 12{,}12.(8{,}25 - 7{,}51)^2 + 9{,}10.(8{,}75 - 7{,}51)^2$$

$$= 0{,}42619600 \approx 0{,}43.$$

Sử dụng công thức (3) ta có :

$$\overline{x^2} = \frac{1}{33}(2.6{,}25^2 + 5.6{,}75^2 + 10.7{,}25^2 + 9.7{,}75^2 + 4.8{,}25^2 + 3.8{,}75^2)$$

$$= 56{,}78977273 \approx 56{,}79$$

và $\overline{x} = 7{,}507575 \Rightarrow (\overline{x})^2 \approx 56{,}36 \Rightarrow s^2 = \overline{x^2} - (\overline{x})^2 = 0{,}4260903466 \approx 0{,}43.$

Từ đó, độ lệch chuẩn là $s = \sqrt{s^2} = \sqrt{0{,}4260903466} = 0{,}6527559625 \approx 0{,}65$ (giây)

b) Sử dụng máy Casio fx-570ES và bảng 2 ta ấn liên tiếp các phím : MODE ; 3 ; 1 ; SHIFT ; MODE ; 4 ; 1 ; SHIFT ; MODE ; 2 ; 6,25 ; = ; 6,75 ; = ; 7,25 ; = ; 7,75 ; = ; 8,25 ; = ; 8,75 ; = ∇ ; ▷ ; 2 ; = ; 5 ; = ; 10 ; = ; 9 ; = ; 4 ; = ; 3 ; = ; AC

SHIFT ; 1 (STAT) ; 5 (Var) ; 2 ; = ta được kết quả $\overline{x} \approx 7{,}51$

SHIFT ; 1 (STAT) ; 5 (Var) ; 3 ; = ta được $s \approx 0{,}65$.

Nhận xét : Nếu làm tròn kết quả ngay từ các phép tính trung gian thì kết quả có thể bị sai lệch. Chẳng hạn, nếu làm tròn $s^2 = 0{,}43$ thì ta sẽ tính được $s = \sqrt{0{,}43} \approx 0{,}66$.

Nếu không sử dụng các bảng phân bố ghép lớp mà dùng trực tiếp bảng dữ liệu đầu tiên (bảng 1) thì kết quả cũng có độ chính xác cao hơn.

➢ **Chú ý.** Nếu không có máy tính cầm tay và phải tính thủ công thì nên sử dụng công thức (3).

c) Thành tích chạy trung bình của học sinh hai lớp xấp xỉ nhau, tuy nhiên lớp 10B có độ lệch chuẩn lớn hơn nên thành tích của các học sinh lớp 10A đồng đều hơn.

C. BÀI TẬP

14. a) Tính phương sai và độ lệch chuẩn của dãy số liệu về chiều cao của các học sinh nam và các học sinh nữ cho ở bảng 5 ;

b) Giả sử trường Trung học phổ thông *M* còn có một nhóm học sinh nam lớp 10 chuyên toán (kí hiệu là nhóm *T*) có chiều cao trung bình là $\overline{x} = 163$ cm, có độ lệch chuẩn là $s = 13$. So sánh chiều cao của ba nhóm học sinh đã cho (nhóm nam, nhóm nữ, nhóm *T*).

15. Hai xạ thủ cùng tập bắn, mỗi người đã bắn 30 viên đạn vào bia. Kết quả được ghi lại ở các bảng sau

Điểm số của xạ thủ A

8	9	10	9	9	10	8	7	6	8
10	7	10	9	8	10	8	9	8	6
10	9	7	9	9	9	6	8	6	8

Bảng 13

Điểm số của xạ thủ B

9	9	10	6	9	10	8	8	5	9
9	10	6	10	7	8	10	9	10	9
9	10	7	7	8	9	8	7	8	8

Bảng 14

a) Tính số trung bình, phương sai và độ lệch chuẩn của các số liệu thống kê cho ở bảng 13, bảng 14.

b) Xét xem trong lần tập bắn này, xạ thủ nào bắn chụm hơn ?

16. Người ta điều tra sản phẩm của hai tổ đóng gói các túi đường (có khối lượng quy định là 2 kg). Kết quả điều tra cho các số liệu thống kê ghi ở hai bảng sau

Khối lượng của 40 túi đường được đóng gói bởi tổ A (đơn vị là kg)

1,95	2,09	1,91	1,99	1,93	2,07	2,15	1,96	1,93	1,94
1,94	2,05	2,02	1,97	1,91	1,95	2,05	2,04	2,03	2,00
2,02	1,94	1,92	1,97	2,00	2,02	2,04	2,05	2,02	2,02
1,94	2,01	1,99	1,95	2,03	2,06	1,91	2,14	1,90	2,25

Bảng 15

Khối lượng của 40 túi đường được đóng gói bởi tổ B (đơn vị là kg)

1,77	1,79	1,80	1,69	1,76	1,69	1,69	1,93	1,94	1,98
2,07	1,98	1,96	1,97	2,06	1,96	1,96	1,91	1,93	2,06
1,97	2,07	2,06	2,08	1,91	1,95	2,05	1,93	1,94	2,02
2,22	2,31	1,80	2,30	2,30	2,23	2,31	2,25	2,24	2,23

Bảng 16

a) Lập bảng phân bố tần số và tần suất ghép lớp theo sản phẩm của tổ *A*, với các lớp

[1,90 ; 1,98) ; [1,98 ; 2,06) ; [2,06 ; 2,14) ; [2,14 ; 2,22) ; [2,22 ; 2,30].

b) Lập bảng phân bố tần số và tấn suất ghép lớp theo sản phẩm của tổ *B*, với các lớp

[1,5 ; 1,7) ; [1,7 ; 1,9) ; [1,9 ; 2,1) ; [2,1 ; 2,3) ; [2,3 ; 2,5].

c) Tính số trung bình, phương sai và độ lệch chuẩn của các số liệu thống kê cho ở bảng 15, bảng 16. Từ đó, xét xem trong lần điều tra này, sản phẩm của tổ nào có khối lượng đồng đều hơn ?