

PHẦN A. LÝ THUYẾT VÀ VÍ DỤ

1. PHƯƠNG SAI VÀ ĐỘ LỆCH CHUẨN

Cho mẫu số liệu ghép nhóm:

Nhóm	$[a_1; a_2)$	...	$[a_i; a_{i+1})$	...	$[a_k; a_{k+1})$
Tần số	m_1		m_i		m_k

- Phương sai của mẫu số liệu ghép nhóm, kí hiệu là s^2 , là một số được tính theo công thức sau:

$$s^2 = \frac{m_1(x_1 - \bar{x})^2 + \dots + m_k(x_k - \bar{x})^2}{n} \quad \text{trong đó,} \quad n = m_1 + \dots + m_k; x_i = \frac{a_i + a_{i+1}}{2} \quad \text{với } i = 1, 2, \dots, k \text{ là}$$

giá trị đại diện cho nhóm $[a_i; a_{i+1})$ và $\bar{x} = \frac{m_1 x_1 + \dots + m_k x_k}{n}$ là số trung bình của mẫu số liệu ghép nhóm.

- Độ lệch chuẩn của mẫu số liệu ghép nhóm, kí hiệu là s , là căn bậc hai số học của phương sai của mẫu số liệu ghép nhóm, tức là $s = \sqrt{s^2}$.

Nhận xét. Ta có thể tính phương sai theo công thức:

$$s^2 = \frac{1}{n} (m_1 \cdot x_1^2 + \dots + m_k \cdot x_k^2) - (\bar{x})^2$$

Độ lệch chuẩn có cùng đơn vị với đơn vị của mẫu số liệu.

Ý nghĩa. Phương sai, độ lệch chuẩn của mẫu số liệu ghép nhóm là các xấp xỉ cho phương sai, độ lệch chuẩn của mẫu số liệu gốc. Chúng được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm xung quanh số trung bình của mẫu số liệu đó. Phương sai, độ lệch chuẩn càng lớn thì mẫu số liệu càng phân tán.

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

$$\hat{s}^2 = \frac{m_1(x_1 - \bar{x})^2 + \dots + m_k(x_k - \bar{x})^2}{n - 1}, \hat{s} = \sqrt{\hat{s}^2}.$$

Ví dụ 1: Người ta theo dõi sự thay đổi cân nặng, được tính bằng hiệu cân nặng trước và sau ba tháng áp dụng chế độ ăn kiêng của một số người cho kết quả như sau:

Thay đổi cân nặng (kg)	$[-1; 0)$	$[0; 1)$	$[1; 2)$	$[2; 3)$	$[3; 4)$
Số người nam	2	3	5	3	2
Số người nữ	2	7	12	7	2

Tính số trung bình, phương sai, độ lệch chuẩn và nhận xét về sự thay đổi cân nặng của người nam, người nữ sau ba tháng áp dụng chế độ ăn kiêng.

Giải

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

Giá trị đại diện	-0,5	0,5	1,5	2,5	3,5
Số người nam	2	3	5	3	2
Số người nữ	2	7	12	7	2

Tổng số người nam là: $n_1 = 2 + 3 + 5 + 3 + 2 = 15$.

Tổng số người nữ là: $n_2 = 2 + 7 + 12 + 7 + 2 = 30$.

Thay đổi cân nặng trung bình của người nam là:

$$\bar{x}_1 = \frac{1}{15} [2 \cdot (-0,5) + 3 \cdot 0,5 + 5 \cdot 1,5 + 3 \cdot 2,5 + 2 \cdot 3,5] = 1,5 (kg).$$

Thay đổi cân nặng trung bình của người nữ là:

$$\bar{x}_2 = \frac{1}{30} [2 \cdot (-0,5) + 7 \cdot 0,5 + 12 \cdot 1,5 + 7 \cdot 2,5 + 2 \cdot 3,5] = 1,5 (kg).$$

Phương sai và độ lệch chuẩn của mẫu số liệu về thay đổi cân nặng của người nam là:

$$s_1^2 = \frac{1}{15} [2 \cdot (-0,5)^2 + 3 \cdot 0,5^2 + 5 \cdot 1,5^2 + 3 \cdot 2,5^2 + 2 \cdot 3,5^2] - 1,5^2 \approx 1,21^2; s_1 \approx 1,21.$$

Phương sai và độ lệch chuẩn của mẫu số liệu về thay đổi cân nặng của người nữ là:

$$s_2^2 = \frac{1}{30} [2 \cdot (-0,5)^2 + 7 \cdot 0,5^2 + 12 \cdot 1,5^2 + 7 \cdot 2,5^2 + 2 \cdot 3,5^2] - 1,5^2 \approx 2,06^2; s_2 \approx 2,06.$$

Như vậy, sau ba tháng áp dụng chế độ ăn kiêng này, về trung bình sự thay đổi cân nặng của nam và nữ là như nhau. Tuy nhiên, sự biến động về thay đổi cân nặng của nữ nhiều hơn so với của nam.

2. SỬ DỤNG PHƯƠNG SAI, ĐỘ LỆCH CHUẨN ĐO ĐỘ RỦI RO

Trong tài chính, người ta có nhiều cách để đo độ rủi ro của một phương án đầu tư. Một trong các cách đó là sử dụng độ lệch chuẩn của lợi nhuận thu được theo phương án đầu tư. Độ lệch chuẩn càng lớn thì phương án đầu tư càng rủi ro.

Ví dụ 2: Anh An đầu tư số tiền bằng nhau vào hai lĩnh vực kinh doanh A, B . Anh An thống kê số tiền thu được mỗi tháng trong vòng 60 tháng theo mỗi lĩnh vực cho kết quả như sau:

Số tiền (triệu đồng)	[5;10)	[10;15)	[15;20)	[20;25)	[25;30)
Số tháng đầu tư vào lĩnh vực A	5	10	30	10	5
Số tháng đầu tư vào lĩnh vực B	20	5	10	5	20

So sánh giá trị trung bình và độ lệch chuẩn của số tiền thu được mỗi tháng khi đầu tư vào mỗi lĩnh vực A, B. Đầu tư vào lĩnh vực nào "rủi ro" hơn?

Giải

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

Giá trị đại diện	7,5	12,5	17,5	22,5	27,5
Số tháng đầu tư vào lĩnh vực A	5	10	30	10	5
Số tháng đầu tư vào lĩnh vực B	20	5	10	5	20

Số tiền trung bình thu được khi đầu tư vào các lĩnh vực A, B tương ứng là:

$$\bar{x}_A = \frac{1}{60}(5 \cdot 7,5 + \dots + 5 \cdot 27,5) = 17,5$$

$$\bar{x}_B = \frac{1}{60}(20 \cdot 7,5 + \dots + 20 \cdot 27,5) = 17,5$$

Như vậy, về trung bình đầu tư vào các lĩnh vực A, B số tiền thu được hàng tháng như nhau. Độ lệch chuẩn của số tiền thu được hàng tháng khi đầu tư vào các lĩnh vực A, B tương ứng là:

$$s_A = \sqrt{\frac{1}{60}(5 \cdot 7,5^2 + \dots + 5 \cdot 27,5^2) - (17,5)^2} = 5$$

$$s_B = \sqrt{\frac{1}{60}(20 \cdot 7,5^2 + \dots + 20 \cdot 27,5^2) - (17,5)^2} \approx 8,42.$$

Như vậy, độ lệch chuẩn của mẫu số liệu về số tiền thu được hàng tháng khi đầu tư vào lĩnh vực B cao hơn khi đầu tư vào lĩnh vực A. Người ta nói rằng, đầu tư vào lĩnh vực B là "rủi ro" hơn.

Ví dụ sau cho thấy không phải lúc nào ta cũng có thể dùng độ lệch chuẩn của lợi nhuận thu được để so sánh độ rủi ro của các phương án đầu tư.

Ví dụ 3: Thống kê lợi nhuận hàng tháng (đơn vị: triệu đồng) trong 20 tháng của hai nhà đầu tư được cho như sau:

Lợi nhuận	[10; 20)	[20; 30)	[30; 40)	[40; 50)	[50; 60)
Số tháng	2	4	8	4	2

Bảng 3.2. Lợi nhuận theo tháng của nhà đầu tư nhỏ

Lợi nhuận	[510; 520)	[520; 530)	[530; 540)	[540; 550)	[550; 560)
Số tháng	4	3	6	3	4

Bảng 3.3. Lợi nhuận theo tháng của nhà đầu tư lớn

Tính độ lệch chuẩn của hai mẫu số liệu ghép nhóm trên. Có nên dựa vào độ lệch chuẩn để so sánh độ rủi ro của hai nhà đầu tư này không?

Giải

Chọn điểm đại diện cho các nhóm số liệu ta tính được các số đặc trưng như sau:

Lợi nhuận trung bình một tháng của các nhà đầu tư tương ứng là:

$$\bar{x}_A = \frac{1}{20}(2 \cdot 15 + \dots + 2 \cdot 55) = 35 \quad (\text{triệu đồng}); \quad \bar{x}_B = \frac{1}{20}(4 \cdot 515 + \dots + 4 \cdot 555) = 535 \quad (\text{triệu đồng}).$$

Độ lệch chuẩn của lợi nhuận hàng tháng của hai nhà đầu tư tương ứng là:

$$s_A = \sqrt{\frac{1}{20}(2 \cdot 15^2 + \dots + 2 \cdot 55^2) - (35)^2} \approx 10,95$$

$$s_B = \sqrt{\frac{1}{20}(4 \cdot 515^2 + \dots + 4 \cdot 555^2) - (535)^2} \approx 13,78.$$

Độ lệch chuẩn cho lợi nhuận hàng tháng của nhà đầu tư lớn cao hơn của nhà đầu tư nhỏ. Lợi nhuận trung bình của hai nhà đầu tư khác nhau rất nhiều, do đó ta không nên dùng độ lệch chuẩn để so sánh mức độ rủi ro của hai nhà đầu tư này.

Nhận xét. Ta không nên dùng phương sai hay độ lệch chuẩn để so sánh độ rủi ro của hai phương án đầu tư khi lợi nhuận trung bình của hai phương án đầu tư này khác nhau rất nhiều.