CÁC THAM SỐ ĐẶC TRƯNG CỦA BIẾN NGẪU NHIÊN

§3. THAM SỐ ĐẶC TRƯNG CỦA BIẾN NGẪU NHIÊN

Những thông tin cô đọng phản ánh từng phần về biến ngẫu nhiên giúp ta so sánh giữa các đại lượng với nhau được gọi là ***các đặc trưng số***. Có 3 loại đặc trưng số là

- Các đặc trưng số cho xu hướng trung tâm của BNN:

  Trung vị, ***Mode***, ***Kỳ vọng***,…

- Các đặc trưng số cho độ phân tán của BNN:

  Phương sai, ***Độ lệch chuẩn***,…

- Các đặc trưng số cho dạng phân phối xác suất.

3.1. TRUNG VỊ và MODE

3.1.1. Trung vị (tham khảo)

Trung vị (*median*) của BNN X, ký hiệu $MedX$, là số thực m thỏa:

$$\boxed{P(X \le m) = P(X \ge m).}$$

VD 1. Cho BNN X có bảng phân phối xác suất:

X	0	1	2	3
P	0,125	0,375	0,375	0,125

Ta có: $MedX = m$ thỏa $1 < m < 2$.

VD 2. Cho BNN X có hàm mật độ:

$$f(x) = \frac{2}{9}\begin{cases} 3x - x^2, & x \in [0;\ 3] \\ 0, & x \notin [0;\ 3]. \end{cases}$$

Ta có: $P(X \le m) = P(X \ge m) \Rightarrow MedX = m \in [0;\ 3]$

$$\Rightarrow \frac{1}{2} = \frac{2}{9}\int_0^m (3x - x^2)dx \Rightarrow m = \frac{3}{2} \in [0;\ 3].$$

3.1.2. MODE

Mode của biến ngẫu nhiên X, ký hiệu $ModX$, là giá trị $x_0 \in X$ thỏa:

- $P(X = x_0)$ max nếu X là rời rạc, và
- $f(x_0)$ max nếu X liên tục có hàm mật độ $f(x)$.

Chú ý

- $ModX$ còn được gọi là ***giá trị tin chắc nhất*** của X.
- Biến ngẫu nhiên X có thể có nhiều $ModX$.

VD 3. Cho BNN X có bảng phân phối xác suất:

X	0	1	**2**	4	5	8
P	0,10	0,20	**0,30**	0,05	0,25	0,10

Ta có: $ModX = 2$.

VD 4. Tìm $ModX$, biết X có bảng phân phối xác suất:

X	1	2	4	5	8
P	$1-3p$	0,18	0,07	0,25	p

Giải. Ta có:

$$1 - 3p + 0,18 + 0,07 + 0,25 + p = 1 \Rightarrow p = 0,25.$$

Vậy $ModX = 1;\ 5;\ 8$.

VD 5. Tìm $ModX$, biết X có hàm mật độ xác suất:

$$f(x) = \begin{cases} \dfrac{3}{64}x^2(4-x), & x \in [0;\ 4] \\ 0, & x \notin [0;\ 4]. \end{cases}$$

Giải. Với $x \in [0;\ 4]$, ta có:

$$f'(x) = 0 \Leftrightarrow x = 0,\ x = \frac{8}{3}.$$

$$f(0) = f(4) = 0,\ f\left(\frac{8}{3}\right) = \frac{4}{9}$$

$$\Rightarrow \max_{\mathbb{R}} f(x) = f\left(\frac{8}{3}\right) = \frac{4}{9}.$$ Vậy $ModX = \dfrac{8}{3}$.

3.2. KỲ VỌNG

3.2.1. Định nghĩa

Kỳ vọng (*Expectation*) của biến ngẫu nhiên X, ký hiệu EX hay $M(X)$, là một số thực được xác định như sau:

- Nếu X là rời rạc với xác suất $P(X = x_i) = p_i$ thì:

$$\boxed{EX = \sum_i x_i p_i.}$$

- Nếu X là liên tục có hàm mật độ $f(x)$ thì:

$$\boxed{EX = \int_{-\infty}^{+\infty} x.f(x)dx.}$$

Đặc biệt

Nếu biến ngẫu nhiên rời rạc $X = \{x_1;\, x_2;\ldots;\, x_n\}$ với xác suất tương ứng là $p_1,\ p_2,\ldots,\ p_n$ thì:

$$\boxed{EX = x_1p_1 + x_2p_2 + \ldots + x_np_n.}$$

VD 6. Cho BNN X có bảng phân phối xác suất:

X	-1	0	2	3
P	0,1	0,2	0,4	0,3

Tính kỳ vọng của X?

Giải. Ta có:

$$EX = -1\times 0{,}1 + 0\times 0{,}2 + 2\times 0{,}4 + 3\times 0{,}3.$$

Vậy $EX = 1{,}6$.

VD 7. Một lô hàng gồm 10 sản phẩm tốt và 2 phế phẩm. Lấy ngẫu nhiên 4 sản phẩm từ lô hàng đó, gọi X là số sản phẩm tốt trong 4 sản phẩm lấy ra.
Tìm phân phối xác suất và tính kỳ vọng của X?

Giải. Ta có: $X = \{2;\ 3;\ 4\}$.

Phân phối xác suất của X là:

$$p_k = P(X = k) = \frac{C_{10}^k C_2^{4-k}}{C_{12}^4},\ k = 2;\ 3;\ 4.$$

$$\text{Vậy } EX = 2.p_2 + 3.p_3 + 4.p_4 = \frac{10}{3}.$$

VD 8. Tìm kỳ vọng của BNN X có hàm mật độ:

$$f(x)=\begin{cases}\frac{3}{4}(x^2+2x), & x\in[0;\ 1]\\ 0, & x\notin[0;\ 1].\end{cases}$$

Giải. $EX=\int\limits_{-\infty}^{+\infty}xf(x)dx=\frac{3}{4}\int\limits_{0}^{1}(x^3+2x^2)dx$

$$=\frac{3}{4}\left(\frac{x^4}{4}+\frac{2x^3}{3}\right)\Bigg|_0^1=\frac{11}{16}.$$

Chú ý

- Nếu X là BNN liên tục trên $[a;\, b]$ thì $EX \in [a;\, b]$.
- Nếu $X = \{x_1, \ldots, x_n\}$ thì:

$$EX \in [\min\{x_1, \ldots, x_n\};\ \ \max\{x_1, \ldots, x_n\}].$$

VD 9. Cho BNN X có bảng phân phối xác suất:

X	1	2	4	5	7
P	a	0,2	b	0,2	0,1

Tìm giá trị của tham số a và b để $EX = 3,5$?

Giải. Ta có:

$$\begin{cases} a + 0,2 + b + 0,2 + 0,1 = 1 \\ 1.a + 2.0,2 + 4.b + 5.0,2 + 7.0,1 = 3,5 \end{cases}$$

$$\Rightarrow \begin{cases} a + b = 0,5 \\ a + 4b = 1,4 \end{cases} \Rightarrow \begin{cases} a = 0,2 \\ b = 0,3. \end{cases}$$

<u>VD 10.</u> Cho biến ngẫu nhiên X có hàm mật độ:

$$f(x) = \begin{cases} ax + bx^2, & x \in [0;\, 1] \\ 0, & x \notin [0;\, 1]. \end{cases}$$

Cho biết $EX = 0,6$. Hãy tính $P(X < 0,5)$?

<u>Giải.</u> Ta có:

$$EX = 0,6 \Rightarrow \int\limits_0^1 (ax^2 + bx^3)dx = \frac{3}{5} \Rightarrow \frac{a}{3} + \frac{b}{4} = \frac{3}{5} \ \ (1)$$

$$\int\limits_{-\infty}^{+\infty} f(x)dx = 1 \Rightarrow \int\limits_0^1 (ax + bx^2)dx = 1 \Rightarrow \frac{a}{2} + \frac{b}{3} = 1 \ \ (2).$$

Từ (1) và (2) $\Rightarrow a = \dfrac{18}{5},\ b = -\dfrac{12}{5}.$

$$P\left(X < 0,5\right) = \int_{-\infty}^{0,5} f(x)dx = \frac{1}{5}\int_{0}^{0,5} (18x - 12x^2)dx$$

$$= \frac{1}{5}\left(9x^2 - 4x^3\right)\Big|_{0}^{0,5} = 0,35.$$

3.2.2. Ý nghĩa của Kỳ vọng

- Kỳ vọng của biến ngẫu nhiên X là ***giá trị trung bình*** (tính theo xác suất) mà X nhận được, nó phản ánh giá trị trung tâm phân phối xác suất của X.

- Trong thực tế sản xuất hay kinh doanh, khi cần chọn phương án cho ***năng suất*** hay ***lợi nhuận*** cao, người ta thường chọn phương án sao cho ***kỳ vọng năng suất*** hay ***kỳ vọng lợi nhuận*** cao.

VD 11. Một thống kê cho biết tỉ lệ tai nạn xe máy ở thành phố H là 0,001. Công ty bảo hiểm A đề nghị bán loại bảo hiểm tai nạn xe máy cho ông B ở thành phố H trong 1 năm với số tiền chi trả là 10 (triệu đồng), phí bảo hiểm là 0,1 (triệu đồng). Hỏi trung bình công ty A lãi bao nhiêu khi bán bảo hiểm cho ông B?

Giải. Gọi X (triệu đồng) là tiền lãi của công ty A sau khi bán bảo hiểm cho ông B.

- Khi ông B không bị tai nạn thì:

$$X = 0{,}1 \text{ và xác suất } p = 0{,}999.$$

- Khi ông B bị tai nạn thì:

$$X = 0{,}1 - 10 = -9{,}9;\ p = 0{,}001.$$

Vậy trung bình tiền lãi của công ty A là:

$$EX = 0,1 \times 0,999 - 9,9 \times 0,001 = 0,09 \text{ (triệu đồng)}.$$

VD 12. Ông A tham gia một trò chơi đỏ, đen như sau: Trong một hộp có 4 bi đỏ và 6 bi đen. Mỗi lần ông A lấy ra 1 bi: nếu là đỏ thì được thưởng 100 (ngàn đồng), nếu là đen thì bị mất 70 (ngàn đồng). Hỏi trung bình mỗi lần lấy bi ông A nhận được bao nhiêu tiền?

Giải. Gọi X (ngàn đồng) là số tiền ông A nhận được sau 1 lần lấy bi.

Ta có bảng phân phối xác suất:

X	100	–70
P	0,4	0,6

$$EX = 100 \times 0,4 - 70 \times 0,6 = -2.$$

Vậy trung bình mỗi lần lấy bi ông A bị mất 2.000 đồng.

VD 13. Người thợ chép tranh mỗi tuần chép hai bức tranh độc lập A và B với xác suất hỏng tương ứng là 0,03 và 0,05. Nếu thành công thì người thợ sẽ kiếm lời từ bức tranh A là 1,3 triệu đồng và B là 0,9 triệu đồng, nhưng nếu hỏng thì bị lỗ do bức tranh A là 0,8 triệu đồng và do B là 0,6 triệu đồng. Hỏi trung bình người thợ nhận được bao nhiêu tiền chép tranh mỗi tuần?

A. 2,185 triệu đồng; B. 2,148 triệu đồng.

C. 2,116 triệu đồng; D. 2,062 triệu đồng.

Giải. Gọi X (triệu đồng) là số tiền người thợ nhận được từ việc chép tranh trong 1 tuần.

- Khi cả A và B đều thành công thì:
  $X = 1,3 + 0,9 = 2,2$ và $p = 0,97.0,95 = 0,9215.$
- Khi A thành công và B hỏng thì:
  $X = 1,3 - 0,6 = 0,7$ và $p = 0,97.0,05 = 0,0485.$
- Khi A hỏng và B thành công thì:
  $X = -0,8 + 0,9 = 0,1$ và $p = 0,03.0,95 = 0,0285.$
- Khi cả A và B đều hỏng thì:
  $X = -0,8 - 0,6 = -1,4$ và $p = 0,03.0,05 = 0,0015.$

Số tiền trung bình mỗi tuần người thợ nhận được là:

$$EX = 2,062 \text{ (triệu đồng)}.$$

VD 14. Một dự án xây dựng được viện C thiết kế cho cả 2 bên A và B xét duyệt một cách độc lập. Xác suất (khả năng) để A và B chấp nhận dự án này khi xét duyệt thiết kế là 70% và 80%. Nếu chấp nhận dự án thì bên A phải trả cho C là 400 triệu đồng, còn ngược lại thì phải trả 100 triệu đồng. Nếu chấp nhận dự án thì bên B phải trả cho C là 1 tỉ đồng, còn ngược lại thì phải trả 300 triệu đồng. Biết chi phí cho thiết kế của C là 1 tỉ đồng và 10% thuế doanh thu. Hỏi trung bình viện C có lãi bao nhiêu khi nhận thiết kế trên?

Hướng dẫn

Gọi X (triệu đồng) là tiền lãi (đã trừ thuế) của C.

Tính tương tự VD 13, ta được $EX = 53$ (triệu đồng).

3.2.3. Kỳ vọng của hàm của biến ngẫu nhiên

Giả sử $Y = \varphi(X)$ là hàm của biến ngẫu nhiên X.

- Nếu X là biến ngẫu nhiên rời rạc thì:

$$\boxed{EY = \sum_i y_i.p_i = \sum_i \varphi(x_i).p_i}$$

- Nếu X là biến ngẫu nhiên liên tục thì:

$$\boxed{EY = \int_{-\infty}^{+\infty} y.f(x)dx = \int_{-\infty}^{+\infty} \varphi(x).f(x)dx}$$

Chú ý

Khi biến ngẫu nhiên X là rời rạc thì ta nên lập bảng phân phối xác suất của Y, rồi tính EY.

VD 15. Cho BNN X có bảng phân phối xác suất:

X	-1	0	1	2
P	0,1	0,3	0,35	0,25

Tính EY với $Y = X^2 - 3$?

Giải. Ta có bảng phân phối xác suất của Y là:

Y	-3	-2	1
P	0,3	0,45	0,25

Vậy $EY = -3.0,3 - 2.0,45 + 1.0,25 = -1,55.$

VD 16. Cho BNN X có hàm mật độ xác suất:

$$f(x) = \begin{cases} \dfrac{2}{x^2}, & x \in [1;\ 2] \\ 0, & x \notin [1;\ 2]. \end{cases}$$

Tính EY với $Y = X^5 - \dfrac{2}{X}$?

Giải. Áp dụng công thức, ta có:

$$EY = \int\limits_{-\infty}^{+\infty} y.f(x)dx$$

$$= \int\limits_{1}^{2} \left(x^5 - \frac{2}{x}\right).\frac{2}{x^2}\,dx = \left(\frac{x^4}{2} + \frac{2}{x^2}\right)\Bigg|_1^2 = 6.$$

3.3. PHƯƠNG SAI

3.3.1. Định nghĩa

Phương sai (*Variance* hay *Dispersion*) của biến ngẫu nhiên X, ký hiệu $VarX$ hay $D(X)$, là một số thực không âm được xác định bởi:

$$\boxed{VarX = E(X - EX)^2 = E(X^2) - (EX)^2.}$$

- Nếu BNN X là rời rạc và $P(X = x_i) = p_i$ thì:

$$\boxed{VarX = \sum_i x_i^2.p_i - \left(\sum_i x_i.p_i\right)^2.}$$

- Nếu BNN X là liên tục và có hàm mật độ $f(x)$ thì:

$$\boxed{VarX = \int_{-\infty}^{+\infty} x^2.f(x)dx - \left(\int_{-\infty}^{+\infty} x.f(x)dx\right)^2.}$$

VD 17. Cho BNN X có bảng phân phối xác suất:

X	1	2	3
P	0,2	0,7	0,1

Ta có: $VarX = (1^2.0,2 + 2^2.0,7 + 3^2.0,1)$

$$-(1.0,2 + 2.0,7 + 3.0,1)^2 = 0,29.$$

VD 18. Tính phương sai của X, biết hàm mật độ:

$$f(x) = \begin{cases} \frac{3}{4}(x^2 + 2x), & x \in [0;\ 1] \\ 0, & x \notin [0;\ 1]. \end{cases}$$

Giải. Áp dụng công thức, ta có:

$$VarX = \int_0^1 x^2.\frac{3}{4}(x^2 + 2x)dx - \left[\int_0^1 x.\frac{3}{4}(x^2 + 2x)dx\right]^2$$

$$= \frac{3}{4}\left(\frac{x^5}{5} + \frac{x^4}{2}\right)\Bigg|_0^1 - \left[\frac{3}{4}\left(\frac{x^4}{4} + \frac{2x^3}{3}\right)\Bigg|_0^1\right]^2 = \frac{67}{1280}.$$

VD 19. Cho BNN X có hàm mật độ xác suất:

$$f(x)=\begin{cases}\dfrac{3}{4}(1-x^2), & |x|\le 1\\ 0, & |x|>1.\end{cases}$$

Tính phương sai của Y, cho biết $Y=2X^2$.

Giải. $VarY = \int\limits_{-1}^{1} y^2.f(x)dx - \left(\int\limits_{-1}^{1} y.f(x)dx\right)^2$

$$= \int\limits_{-1}^{1} (2x^2)^2.\frac{3}{4}(1-x^2)dx - \left[\int\limits_{-1}^{1} 2x^2.\frac{3}{4}(1-x^2)dx\right]^2$$

$$= 3\left(\frac{x^5}{5} - \frac{x^7}{7}\right)\Bigg|_{-1}^{1} - \left[\frac{3}{2}\left(\frac{x^3}{3} - \frac{x^5}{5}\right)\Bigg|_{-1}^{1}\right]^2 = \frac{32}{175}.$$

3.3.2. Ý nghĩa của Phương sai

- $(X - EX)^2$ là bình phương sai biệt giữa giá trị của X so với trung bình của nó. Và phương sai là trung bình của sai biệt này, nên phương sai cho ta hình ảnh về ***sự phân tán*** của các số liệu: phương sai càng nhỏ thì số liệu càng tập trung xung quanh trung bình của chúng.
- Trong kỹ thuật, phương sai đặc trưng cho độ sai số của thiết bị. Trong kinh doanh, phương sai đặc trưng cho độ rủi ro đầu tư.

- Do đơn vị đo của $VarX$ bằng bình phương đơn vị đo của X nên để so sánh được với các đặc trưng khác, người ta đưa vào khái niệm ***độ lệch tiêu chuẩn*** (*standard deviation*) là:

$$\boxed{\sigma = \sqrt{VarX}.}$$

VD 20. Năng suất (sản phẩm/phút) của hai máy tương ứng là các BNN X và Y, có bảng phân phối xác suất:

X	1	2	3	4
P	0,3	0,1	0,5	0,1

Y	2	3	4	5
P	0,1	0,4	0,4	0,1

Từ bảng phân phối xác suất, ta tính được:

$$EX = 2,4;\ VarX = 1,04;\ EY = 3,5;\ VarY = 0,65.$$

Vì $EX < EY$, $VarX > VarY$ nên nếu phải chọn mua một trong hai loại máy này thì ta chọn mua máy Y.

Chú ý

Trong trường hợp $\begin{cases} EX < EY \\ VarX < VarY \end{cases}$ hay $\begin{cases} EX > EY \\ VarX > VarY \end{cases}$

thì ta không thể so sánh được. Để giải quyết vấn đề này, trong thực tế người ta dùng ***tỉ số tương đối*** $\frac{\sigma}{\mu}.100\%$ (μ là trung bình) để so sánh sự ổn định của các BNN X và Y. Tỉ số tương đối ***càng nhỏ*** thì ***độ ổn định càng cao***.

VD 21. Điểm thi hết môn XSTK của lớp A và B tương ứng là các BNN X và Y. Từ bảng kết quả điểm thi người ta tính được:

$$EX = 6,25;\ VarX = 1,25;\ EY = 5,75;\ VarY = 0,75.$$

Ta có: $\dfrac{\sigma_x}{EX}.100\% = 17,89\%;\ \dfrac{\sigma_y}{EY}.100\% = 15,06\%.$

Vậy lớp B học đều (ổn định) hơn lớp A.

3.4. Một số đặc trưng khác (tham khảo)

Xét BNN X có kỳ vọng, phương sai là μ và σ^2.

a) Hệ số đối xứng của *X*

$$\boxed{\gamma_1(X) = \frac{E(X-\mu)^3}{\sigma^3}.}$$

Khi $\gamma_1(X) = 0$ thì phân phối của X là đối xứng; lệch phải khi $\gamma_1(X) > 0$ và lệch trái khi $\gamma_1(X) < 0$.

b) Hệ số nhọn của *X*

$$\boxed{\gamma_2(X) = \frac{E(X-\mu)^4}{\sigma^4}.}$$

Khi $\gamma_2(X)$ càng lớn thì phân phối của X càng nhọn.

……………………………………………………………………