§1 BIẾN NGẪU NHIÊN RỜI RẠC

1.1 Khái niệm biến số ngẫu nhiên

*** Khái niệm**

Biến ngẫu nhiên là biến số phụ thuộc vào yếu tố ngẫu nhiên nào đó. Ta sẽ ký hiệu biến ngẫu nhiên là $X, Y, \ldots, X_1, X_2, \ldots$

Trong toán học người ta đưa ra định nghĩa

$$X = X(\omega) : \Omega \longmapsto \mathbb{R}.$$

*** <u>Thí dụ 1</u>.** Gieo một con xúc xắc (phép thử) và gọi X là số chấm xuất hiện. Rõ ràng biến ngẫu nhiên X phụ thuộc vào kết cục của phép thử và lấy giá trị trong tập $\{1, 2, \ldots, 6\}$.

*** <u>Thí dụ 2</u>.** Chiều cao của một sinh viên là một biến ngẫu nhiên. Phụ thuộc vào độ chính xác của phép đo, ta có giá trị của nó nằm trong một khoảng số nào đó, chẳng hạn $(1{,}69; 1{,}71m)$.

Biến ngẫu nhiên
rời rạc

Biến ngẫu nhiên
liên tục

*** Phân loại**

(i) Biến ngẫu nhiên được gọi là *rời rạc*, nếu tập giá trị của nó là một tập hữu hạn $\{x_1, x_2, ..., x_n\}$ hoặc vô hạn đếm được các phần tử (*tập đếm được*).

(ii) Biến ngẫu nhiên được gọi là *liên tục*, nếu tập giá trị của nó lấp kín một khoảng (đoạn) hoặc một số khoảng trên trục số hoặc cũng có thể là cả trục số (*tập không đếm được*).

1.2 Bảng phân phối xác suất (Bảng PPXS)

Ta ký hiệu $p_i = P(X = x_i), i = 1, 2, \dots, n, \dots$

*** Định nghĩa.** *Bảng PPXS của một biến ngẫu nhiên rời rạc X* được xác định như sau

$X = x$	x_1	x_2	...	x_n	...
$p(x)$	p_1	p_2	...	p_n	...

Chú ý tập giá trị của X được sắp xếp theo thứ tự tăng dần, còn hàm $p(x) = P(X = x)$.

*** <u>Thí dụ 3</u>**. Gọi X là số chấm xuất hiện khi gieo một con xúc xắc (xem thí dụ 1). Khi đó bảng PPXS của X sẽ là

$X = x$	1	2	3	...	6
$p(x)$	1/6	1/6	1/6	...	1/6

Để ý các XS $p(x)$ bằng nhau trên tập giá trị {1, 2, ..., 6} của X (đồng khả năng). Ta nói rằng X có phân phối đều trên tập {1, 2, ..., 6} và ký hiệu $X \sim \mathcal{U}_{\{1,2,\dots,6\}}$ (*uniform distribution of probability*). Người ta cũng coi XS để X nhận các giá trị không có trong bảng bằng 0.

*** <u>Thí dụ 4</u>**. Một xạ thủ có 3 viên đạn. Anh ta được yêu cầu bắn từng viên cho đến khi trúng thì dừng bắn. Tìm PPXS của số đạn cần bắn, biết rằng XS bắn trúng mỗi lần bắn là 0,4.

Giải: Rõ ràng số đạn cần bắn, ký hiệu là X, là biến ngẫu nhiên rời rạc và có 3 giá trị 1, 2, 3.

$X = 1 \iff$ sự kiện phát đạn đầu trúng, $p_1 = P(X=1) = 0{,}4$;

$X = 2 \iff$ phát đầu trượt, phát thứ hai trúng và dễ dàng tính được $p_2 = 0{,}6.0{,}4 = 0{,}24$;

$X = 3 \iff$ hai phát đạn đầu trượt, $p_3 = 0{,}36$. Từ đó

$X = x$	1	2	3
$p(x)$	0,4	0,24	0,36

*** <u>Thí dụ 5</u>**. Lấy lại thí dụ 4 nhưng xạ thủ được yêu cầu bắn hết đạn.

Tìm PPXS của số đạn trúng.

Giải: Nếu gọi X là số đạn trúng thì tập giá trị của nó {0, 1, 2, 3}, ở đây ta có lược đồ Bernoulli với $n = 3$, $p = 0{,}4$ và $k = \overline{0{,}3}$. Sử dụng công thức Bernoulli $p(k) = C_n^k\, p^k q^{n-k}$

$X = x$	0	1	2	3
$p(x)$	0,064	0,288	0,432	0,216

Trong trường hợp tổng quát (n, p tuỳ ý) ta gọi PPXS của X là *phân phối nhị thức* (*binomial distribution*) và ký hiệu là $X \sim \mathcal{B}(n, p)$. Ở đây $X \sim \mathcal{B}(3; 0{,}4)$.

Hàm số $p(x) = P(X = x)$, $x \in$ tập giá trị của X, được gọi là *hàm xác suất* (*probability function*) của X; nó có hai tính chất quan trọng:

*** Tính chất**

(i) $p(x) > 0 \;\; \forall x$;

(ii) $\sum_{\forall x} p(x) = 1$.

*** <u>Thí dụ 6</u>**. Biến ngẫu nhiên X được coi là tuân theo *luật PPXS Poisson*, ký hiệu là $X \sim \mathcal{P}(\lambda)$, λ là tham số nào đó, nếu hàm xác suất của nó được xác định như sau

$$p(x) = \frac{\lambda^x}{x!} e^{-\lambda}, x = 0, 1, 2, \ldots$$

Chú ý hàm của các biến ngẫu nhiên rời rạc tiếp tục là biến ngẫu nhiên rời rạc. Việc xác định PPXS của biến hàm trong trường hợp này tương đối đơn giản.

*** Thí dụ 7**. Cho hai biến X và Y có bảng PPXS tương ứng:

x	-1	0	1
$p(x)$	0,3	0,4	0,3

y	1	2
$p(y)$	0,3	0,7

Lập bảng PPXS của: a) X^2; b) $X + Y$.

Giải:

a) Biến $Z = X^2$ chỉ có hai giá trị 0 và 1. Từ đó ta có bảng PPXS

z	0	1
$p(z)$	0,4	0,6

b) Biến $Z = X + Y$ có các giá trị 0, 1, 2 và 3. Để ý rằng

$$P(Z = z_k) = \sum_{\forall i,j:\, z_k = x_i + y_j} P(\{X = x_i\}.\{Y = y_j\}).$$

Tích trong tổng được hiểu là sự kiện xảy ra đồng thời $\{X = x_i\}$ và $\{Y = y_j\}$. Nếu thêm giả thiết hai sự kiện trên độc lập thì

$$P(\{X = x_i\}.\{Y = y_j\}) = P(X = x_i)P(Y = y_j).$$

Từ đó bảng PPXS của Z

z	0	1	2	3
$p(z)$	0,09	0,33	0,37	0,21

(chẳng hạn $P(Z = 2) = P(0;2) + P(1;1) = 0{,}4.0{,}7 + 0{,}3.0{,}3$).

1.3 Hàm phân phối xác suất (*hàm phân phối*)

Có một cách khác để đặc trưng cho phân phối XS.

*** Định nghĩa.** *Hàm phân phối xác suất* (*probability distribution function*) của một biến ngẫu nhiên X, ký hiệu là $F(x)$, được xác định như sau

$$F(x) = P(X < x), x \in \mathbb{R}.$$

Ý nghĩa: $F(x)$ phản ánh độ tập trung XS ở bên trái điểm x

$\Rightarrow$ *hàm phân phối tích luỹ (xác suất tích luỹ)*

$\Rightarrow$ *hàm phân phối tích phân.*

*** <u>Thí dụ 8</u>**. Từ bảng PPXS của thí dụ 4 ta có $F(x) = \sum_{x_i < x} p(x_i) \Rightarrow$

$$F(x) = \begin{cases} 0; & x < 1, \\ 0{,}4; & 1 < x \leq 2, \\ 0{,}64; & 2 < x \leq 3, \\ 1; & x > 3. \end{cases}$$

Nếu vẽ đồ thị, ta có dạng bậc thang. Để ý là tại mỗi nơi chuyển bậc thì ta có điểm gián đoạn loại 2. Hàm phân phối cho biến liên tục có vai trò quan trọng trong nghiên cứu xác suất.

*** Tính chất**

(i) $1 \geq F(x) \geq 0$;

(ii) $F(x)$ là hàm không giảm $\Leftrightarrow \forall\, x_2 > x_1 \Rightarrow F(x_2) \geq F(x_1)$;

(iii) $P(\alpha \leq X < \beta) = F(\beta) - F(\alpha)$;

(iv) $F(+\infty) = 1$; $F(-\infty) = 0$.

Việc chứng minh các tính chất này khá đơn giản dùng định nghĩa hàm phân phối.

Hạn chế: + khó biết nơi nào tập trung nhiều XS hơn,

+ sau này ta thấy không thể biểu diễn $F(x)$ dưới dạng hàm sơ cấp trong nhiều trường hợp.

1.4 Các số đặc trưng cơ bản

A. Kỳ vọng

Ta xét *phép tính kỳ vọng* trong trường hợp biến rời rạc. Giả thiết rằng X có hàm XS $p(x)$. Khi đó kỳ vọng (*Expectation value*) của $Z = g(X)$ được định nghĩa như sau

*** Định nghĩa 1.** Kỳ vọng của $Z = g(X)$ với $p(x)$ là hàm XS đã cho của biến X được tính theo

$$EZ = \sum_{\forall x} g(x)p(x). \tag{1}$$

Trường hợp riêng khi $Z = X$, ta có

*** Định nghĩa 2.** *Kỳ vọng của X*, với $p(x)$ là hàm XS đã cho của biến X, được tính theo

$$EX = \sum_{\forall x} xp(x) = \sum_{\forall i} x_i p_i. \tag{2}$$

*** Tính chất**

(i) $E(c) = c$, $c = const$;

(ii) $E(cX) = cEX$;

(iii) $E(X + Y) = EX + EY$;

(iv) $E(XY) = EX.EY$, X và Y độc lập.

Chú ý: tính chất (ii) và (iii) được gọi là *tính tuyến tính* của phép toán, vốn quen thuộc trong phép tính đạo hàm hay tích phân.

Ý nghĩa thực tế: kỳ vọng là *giá trị trung bình* (*mean value*) của biến ngẫu nhiên hoặc đóng vai trò *định vị* của biến.

*** <u>Thí dụ 9</u>**. Xét lại thí dụ 3 và cần tính kỳ vọng của X.

Giải: Do $X \sim \mathcal{U}_{\{1,2,\ldots,6\}}$ nên sử dụng (2)

$$EX = \frac{1}{6}(1 + 2 + \ldots + 6) = 3,5.$$

Nhận xét rằng nếu X có phân phối đều thì EX nhận giá trị chính giữa tập giá trị, ở đây $EX = (1 + 6)/2$.

Nếu muốn tìm kỳ vọng của tổng số chấm xuất hiện khi gieo 3 con xúc xắc, sủ dụng tính chất (iii) ta có kỳ vọng bằng $3.3,5 = 10,5$.

*** Thí dụ 10**. Một nhóm có 6 nam và 3 nữ. Chọn ngẫu nhiên ra 3 người. Hỏi trung bình chọn được bao nhiêu nam trong 3 người?

Giải: Gọi X – số nam được chọn, dễ thấy với $k = 0, 1, 2, 3$

$$P(X = k) = \frac{C_6^k C_3^{3-k}}{C_9^3}.$$

Thay vào (2) và thực hiện tính toán

$p(0) = 1/84; p(1) = 3/14; p(2) = 15/28; p(3) = 5/21$

$$\Rightarrow \qquad EX = 0.\frac{1}{84} + 1.\frac{3}{14} + 2.\frac{15}{28} + 3.\frac{5}{21} = 2.$$

*** Thí dụ 11**. Một trạm y tế có 1 bác sỹ, 2 y sỹ và 6 điều dưỡng. Chọn ngẫu nhiên từng người một cho đến khi được 1 điều dưỡng thì dừng chọn. Tính (theo cách chọn trên):

a) Số người trung bình được chọn.

b) Số y sỹ trung bình được chọn.

Giải:

a) Gọi X – số người được chọn theo cách của đầu bài và X nhận các giá trị 1, 2, 3 và 4. Từ đó, nếu đặt A_i - chọn được người không phải là điều dưỡng ở lần chọn thứ i, i =1, 2, 3, 4:

$$P(X = 1) = \mathrm{P}(\bar{A}_1) = 6/9 = 2/3;$$

$$P(X = 2) = P(A_1\bar{A}_2) = 3.6/(9.8) = 1/4;$$

$$P(X = 3) = P(A_1A_2\bar{A}_3) = 3.2.6/(9.8.7) = 1/14;$$

$$P(X = 4) = P(A_1A_2A_3\bar{A}_4) = 3.2.1.6/(9.8.7.6) = 1/84.$$

Có thể kiểm tra dễ dàng $\sum_{\forall x} p(x) = 1$. Áp dụng (2) ta có

$$EX = \sum_{\forall x} xp(x) = 1.\frac{2}{3} + 2.\frac{1}{4} + 3.\frac{1}{14} + 4.\frac{1}{84} = \frac{10}{7} \approx 1{,}4286.$$

b) Gọi Y là số y sỹ được chọn và ta cần tính EY. Việc lập bảng PPXS cho Y khá phức tạp. Tuy nhiên nếu sử dụng tính chất của kỳ vọng thì đơn giản hơn nhiều. Gọi Z – số bác sỹ được chọn theo cách của đầu bài. Ta có:

$$Y + Z + 1 = X \quad \Rightarrow \quad \boldsymbol{E}Y + EZ + 1 = EX = 10/7$$

$$\Rightarrow \quad \boldsymbol{E}Y + EZ = 3/7;$$

mặt khác theo quy tắc tam xuất

$$EY/EZ = 2/1 \quad \Rightarrow \quad \boldsymbol{E}Y = 2/7 \approx 0{,}2857.$$

B. Phương sai

Dùng phép tính kỳ vọng ở mục trước ta có thể đưa ra khái niệm *phương sai* (*Variance value*).

*** Định nghĩa.** Phương sai của biến ngẫu nhiên X, ký hiệu là VX, được xác định như sau

$$VX = E[(X - EX)^2]. \tag{3}$$

Nhận xét: $X - EX$ là độ lệch của X so với trung bình của nó

⇒ Phương sai – trung bình của bình phương độ lệch

⇒ Đặc trưng cho độ phân tán của X quanh trung bình

độ bất định; độ rủi ro ...

Chú ý: Công thức (3) có dạng tương đương

$$VX = E(X^2) - (EX)^2.$$

Cách tính: $VX = \sum_{\forall x}(x - EX)^2\, p(x) = \sum_{\forall i}(x_i - EX)^2 p_i$

$$= \sum_{\forall x} x^2\, p(x) - (EX)^2 = \sum_{\forall i} {x_i}^2 p_i - (EX)^2.$$

*** Tính chất**

(i) $V(c) = 0$, $c = const$;

(ii) $V(cX) = c^2 VX$;

(iii) $V(X + Y) = VX + VY$, với X và Y độc lập.

* **Thí dụ 12**. Tính phương sai của $X \sim \mathcal{U}_{\{1,2,\ldots,6\}}$ (xem thí dụ 3).

Giải: Ta đã có trong thí dụ 3 $EX = 3{,}5$. Bây giờ ta phải tính

$$E(X^2) = \frac{1}{6}(1 + 2^2 + 3^2 + \ldots + 6^2) = \frac{91}{6}$$

$$\Rightarrow VX = E(X^2) - (EX)^2 = \frac{91}{6} - 3{,}5^2 = \frac{35}{12} \approx 2{,}9167.$$

* **Thí dụ 13. Hai xạ thủ A và B bắn loạt 100 viên mỗi người vào bia có các vòng tròn từ 5 đến 10 điểm. Tần suất bắn đạt các điểm của mỗi xạ thủ được cho trong bảng sau:**

Điểm X	5	6	7	8	9	10
A	0,01	0,03	0,05	0,12	0,39	0,40
B	0,02	0,04	0,07	0,10	0,25	0,52

Bạn có thể đánh giá gì về kết quả bắn của hai xạ thủ?

Giải: Có thể coi tần suất là xác suất xấp xỉ. Từ đó tính được nếu gọi A và B tương ứng là điểm của A và B:

$$EA = 9{,}05;\ EB = 9{,}08;\ E(A^2) = 83{,}05;\ E(B^2) = 84{,}02$$

$$\Rightarrow \qquad VA = 1{,}1475;\ VB = 1{,}5736.$$

Ta thấy điểm trung bình của A ($EA = 9{,}05$) bé hơn của B ($EB =$

9,08), tuy nhiên phương sai của A $VA = 1{,}1475$ nhỏ hơn $VB = 1{,}5736$ hay kết quả của A ổn định hơn B.

BÀI TẬP

1. Có 4 xạ thủ cùng bắn mỗi người một phát vào bia, XS trúng bia của các xạ thủ tương ứng là 0,9; 0,8; 0,7 và 0,6.
   a) Tìm PPXS của số đạn trúng. Sau đó lập hàm phân phối tương ứng và vẽ đồ thị của hàm.
   b) Tính xác suất để có không quá 2 người bắn trúng.

2. Một lô gồm 100 sản phẩm có 10 phế phẩm. Lấy ngẫu nhiên ra 5 sản phẩm (không hoàn lại). Lập bảng PPXS của số phế phẩm trong 5 sản phẩm đó.

3. Biến ngẫu nhiên X có bảng PPXS sau, λ là tham số,

x	1	2	3	4	5
$p(x)$	2λ	2λ	3λ	λ	2λ

   a) Xác định tham số λ.
   b) Lập hàm phân phối của X và vẽ đồ thị.
   c) Tìm giá trị k nhỏ nhất sao cho $P(X \leq k) > 0{,}5$.

4. Cho biến ngẫu nhiên X có bảng PPXS

x	1	3	5	7	9
$p(x)$	0,1	0,2	0,3	0,3	0,1

   Tìm PPXS của $Y = \min\{X, 6\}$.

5. Trọng lượng (g) của một chất trong 200 gói cho kết quả

Trọng lượng	47	48	49	50	51
Số gói	17	36	98	35	14

Tìm PPXS của trọng lượng trên, sau đó tính kỳ vọng và phương sai của trọng lượng đó.

6. Chữa bệnh cho 40 người, người ta thống kê được thời gian khỏi bệnh như sau:
    - có 12 người khỏi sau 7 ngày điều trị,
    - có 15 người khỏi sau 8 ngày điều trị,
    - có 13 người khỏi sau 9 ngày điều trị.

   Tính kỳ vọng và phương sai của thời gian điều trị khỏi bệnh của nhóm bệnh nhân trên.

7. Một người có một chùm chìa khoá 7 chiếc, trong đó chỉ có hai chiếc mở được cửa. Người đó thử ngẫu nhiên từng chiếc cho đến khi mở được cửa. Gọi X là số lần thử, tìm PPXS của X và tính EX, VX.

8. Trong một chiếc hộp có 4 tấm thẻ được đánh số từ 1 đến 4. Chọn ngẫu nhiên ra 2 tấm thẻ và gọi X là tổng của hai số trên hai thẻ. Tìm PPXS của X.

9. Một người chọn ngẫu từng sản phẩm từ một lô hàng (có số lượng sản phẩm rất lớn) cho đến khi được một phế phẩm thì dừng. Gọi X là số lần cần chọn. Tìm PPXS của X, biết tỷ lệ phế phẩm của lô hàng là 10%. Giả sử có 100 người cùng chọn như vậy, hỏi có khoảng bao nhiêu người phải chọn ít nhất 3 lần?

10. Một xạ thủ có 5 viên đạn. Anh ta được yêu bắn cho đến khi có 2 viên đạn trúng thì dừng. Gọi X là số đạn phải bắn, biết XS bắn trúng mỗi lần là 0,4. Tìm PPXS của X, sau đó tìm kỳ vọng và phương sai của X.

11. Có 10 sản phẩm, trong đó có 2 phế phẩm. Lấy lần lượt ra 3 sản phẩm (không hoàn lại).
    a) Gọi X là số chính phẩm trong 3 sản phẩm đó. Tìm PPXS của X, sau đó tìm kỳ vọng và phương sai của X.

b) Gọi Y là số phế phẩm trong 3 sản phẩm lấy ra, tìm EY và VY.

12. Một người bỏ ra x\$ để tham gia trò chơi tung con xúc xắc 3 lần. Nếu cả 3 lần đều xuất hiện mặt lục thì thu về 36\$, nếu 2 lần xuất hiện mặt lục thì thu về 2,8\$, còn nếu 1 lần xuất hiện mặt lục thì thu về 0,4\$.
a) Tìm x sao cho trò chơi là vô thưởng vô phạt.
b) x bằng bao nhiêu thì trung bình mỗi lần chơi sẽ mất 1\$?

13. Hai cầu thủ thay nhau ném bóng vào rổ cho đến khi nào trúng rổ thì dừng ném, biết rằng XS ném trúng mỗi lần của từng người tương ứng là 0,4 và 0,6. Tìm PPXS của:
a) số lần ném của cầu thủ thứ nhất;
b) số lần ném của cả hai cầu thủ.

14. XS bắn trúng của một khẩu súng là p. Tiến hành bắn liên tiếp trong điều kiện như nhau cho đến khi trúng thì dừng bắn. Tìm số đạn trung bình phải bắn.