Biến ngẫu nhiên

Hoàng Văn Hà
University of Science, VNU - HCM
hvha@hcmus.edu.vn

Nội dung

Định nghĩa

Biến ngẫu nhiên (random variable) có thể được mô tả như một "quy tắc" biểu diễn các kết quả của phép thử ngẫu nhiên dưới dạng số.

Định nghĩa 1.1

Biến ngẫu nhiên X là một ánh xạ từ không gian các biến cố sơ cấp Ω vào $\mathbb{R}$,

$$\begin{array}{cccc} X & : & \Omega & \longrightarrow & \mathbb{R} \\ & & \omega & \longmapsto & X = X(\omega) \end{array}$$

Người ta thường dùng các chữ in $X, Y, Z, \ldots$ để ký hiệu các biến ngẫu nhiên và các chữ thường $x, y, z, \ldots$ để chỉ các giá trị của biến ngẫu nhiên.

Định nghĩa

Ví dụ 1

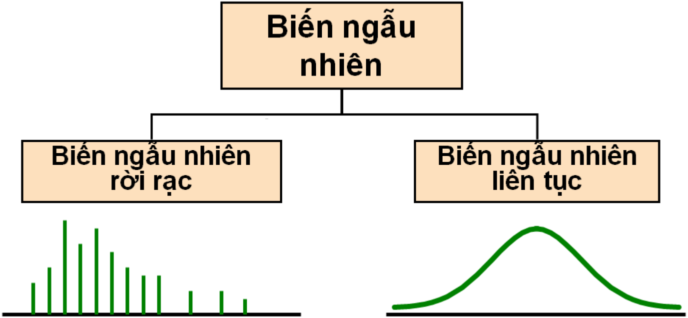
Thực hiện phép thử gieo đồng thời 3 đồng xu cân đối, trong trường hợp này chúng ta có các biến cố sơ cấp sau

$$\begin{array}{llll} \omega_1 = (SSS), & \omega_2 = (SSN), & \omega_3 = (SNN), & \omega_4 = (SNS), \\ \omega_5 = (NNN), & \omega_6 = (NNS), & \omega_7 = (NSS), & \omega_8 = (NSN). \end{array}$$

Nếu gọi biến ngẫu nhiên X là số đồng xu ngửa xuất hiện thì X nhận các giá trị sau

$$\begin{array}{llll} X(\omega_1) = 0, & X(\omega_2) = 1, & X(\omega_3) = 2, & X(\omega_4) = 1, \\ X(\omega_5) = 3, & X(\omega_6) = 2, & X(\omega_7) = 1, & X(\omega_8) = 2. \end{array}$$

Phân loại biến ngẫu nhiên



Biến ngẫu nhiên rời rạc (Discrete random variable)

Định nghĩa 1.2

Một biến ngẫu nhiên được gọi là rời rạc nếu tập hợp các giá trị mà nó có thể nhận là một tập hữu hạn hoặc đếm được.

Ví dụ 2

Các biến ngẫu nhiên sau là các biến ngẫu nhiên rời rạc:

- *Số sản phẩm kém chất lượng trong một lô hàng*
- *Số con trong một gia đình*
- *Số cuộc điện thoại đến một tổng đài ở bưu điện trong một ngày*
- *Số bit lỗi được truyền đi trong một kênh truyền tín hiệu số*

Biến ngẫu nhiên liên tục (Continuous random variable)

Định nghĩa 1.3

Một biến ngẫu nhiên được gọi là liên tục nếu tập hợp các giá trị mà nó nhận được là một khoảng dạng (a, b) *(hoặc* $(a, b], [a, b), [a, b]$*) hoặc toàn bộ* $\mathbb{R}$.

Ví dụ 3

Các biến ngẫu nhiên sau là biến ngẫu nhiên liên tục:

- *Nhiệt độ không khí ở mỗi thời điểm nào đó*
- *Thời gian hoạt động bình thường của một bóng đèn điện tử*
- *Độ pH của một chất hóa học nào đó*
- *Độ dài của một vật mẫu*

Các giá trị ở ví dụ trên có thể lấy bất kỳ giá trị nào tùy thuộc vào khả năng đo lường.

Định nghĩa 1.4

Một hệ thức cho phép biễu diễn mối quan hệ giữa các giá trị có thể có của biến ngẫu nhiên với xác suất nhận các giá trị tương ứng gọi là quy luật phân phối xác suất của biến ngẫu nhiên.

Định nghĩa 1.5

Hàm phân phối xác suất (Cumulative distribution function - c.d.f) của biến ngẫu nhiên X (xác định trên không gian các biến cố sơ cấp Ω) là hàm $F(x)$ được định nghĩa

$$F(x) = \mathbb{P}(X \leq x) \tag{1}$$

với mọi $x \in (-\infty, +\infty)$.

Tính chất của hàm phân phối xác suất

Mệnh đề 1.1

Hàm phân phối xác suất $F(x)$ có các tính chất cơ bản sau

i) *Hàm phân phối là hàm không giảm.*

ii) *Liên tục phải, có giới hạn bên trái tại mọi điểm.*

iii) $F(-\infty) = \lim\limits_{x\to-\infty} F(x) = 0, \quad F(+\infty) = \lim\limits_{x\to+\infty} F(x) = 1.$

iv) $\mathbb{P}(a < X \le b) = F(b) - F(a)$ *với mọi* $a, b \in \mathbb{R}$ *và* $a \le b$.

Hàm khối xác suất (probability mass function)

Định nghĩa 2.1

Xét một biến ngẫu nhiên rời rạc X có thể nhận các giá trị $x_1, x_2, \ldots,$ phân phối xác suất (probability distribution) hay hàm khối xác suất (probability mass function - p.m.f) của X được cho bởi

$$(1) \quad p_X(x) = \mathbb{P}(X = x) \geq 0, \quad \forall x \in \{x_1, x_2, \ldots\},$$

$$(2) \quad \sum_x p_X(x) = 1.$$

- Khi chỉ xét một biến ngẫu nhiên X, để đơn giản kí hiệu, ta sử dụng $p(x)$ thay vì $p_X(x)$.
- Trong một số tài liệu, ta có thể sử dụng kí hiệu $f_X(x)$ (hay $f(x)$) để biểu diễn hàm khối xác suất, thay vì dùng $p_X(x)$ (hay $p(x)$).

Bảng phân phối xác suất

Để mô tả biến ngẫu nhiên X nhận giá trị nào đó với xác suất tương ứng là bao nhiên thì người ta dùng bảng phân phối xác suất. Bảng phân phối xác suất là một bảng có hai dòng:

- Dòng thứ nhất là các giá trị có thể của biến ngẫu nhiên X.
- Dòng thứ hai là xác suất biến ngẫu nhiên X nhận các giá trị tương ứng.

Bảng phân phối của một biến ngẫu nhiên X có dạng như sau:

X	x_1	x_2	$\cdots$	x_n	$\cdots$
$\mathbb{P}$	$p(x_1)$	$p(x_2)$	$\cdots$	$p(x_n)$	$\cdots$

Ví dụ 4

Cho biến ngẫu nhiên rời rạc X có phân phối xác suất cho bởi

$$p(x) = \frac{2x+1}{25}, \quad x = 0, 1, 2, 3, 4.$$

(a) *Lập bảng phân phối xác suất.*

(b) *Tính $\mathbb{P}(X \leq 1)$ và $\mathbb{P}(2 \leq X < 4)$.*

Ví dụ 5

Một xạ thủ có 4 viên đạn, bắt lần lượt từng viên vào một mục tiêu một cách độc lập. Xác suất bắn trúng mục tiêu ở mỗi lần bắn là 0.7. Nếu có một viên trúng hoặc hết đạn thì dừng. Gọi X là số viên đạn đã bắn, lập bảng phân phối xác suất cho X.

Định nghĩa 2.2

Xét biến ngẫu nhiên rời rạc X có thể nhận các giá trị $x_1, x_2, \dots, x_n$. Hàm phân phối xác suất của X, ký hiệu là $F(x)$, được xác định như sau

$$F(x) = \mathbb{P}(X \leq x) = \sum_{x_i \leq x} p(x_i) \tag{2}$$

Cụ thể:

$$F(x) = \mathbb{P}(X \leq x) = \begin{cases} 0, & x < x_1 \\ p(x_1), & x_1 \leq x < x_2 \\ p(x_1) + p(x_2), & x_2 \leq x < x_3 \\ \quad\vdots & \\ p(x_1) + \cdots + p(x_{n-1}), & x_{n-1} \leq x < x_n \\ 1, & x \geq x_n \end{cases} \tag{3}$$

Ví dụ 6

Một lô hàng có 300 sản phẩm trong đó có 50 sản phẩm kém chất lượng. Chọn ngẫu nhiên lần lượt 2 sản phẩm không hoàn lại. Gọi X là số sản phẩm kém chất lượng trong 2 sản phẩm được chọn.

(a) *Lập bảng phân phối xác suất cho X.*

(b) *Tìm hàm phân phối xác suất của X.*

Ví dụ 7

Xét biến ngẫu nhiên rời rạc X có thể nhận các giá trị thuộc $S = \{1, 2, 3, \ldots\}$. Giả sử phân phối xác suất của X cho bởi

$$p_X(k) = \frac{1}{2^k},\ k = 1, 2, 3, \ldots$$

(a) *Chứng tỏ rằng $p_X(\cdot)$ thỏa các điều kiện của một hàm khối xác suất và tìm hàm phân phối xác suất của X.*

(b) *Tính $\mathbb{P}(2 < X \leq 5)$.*

(c) *Tính $\mathbb{P}(X > 4)$.*

Định nghĩa 2.3

Cho biến ngẫu nhiên liên tục X, nếu hàm số $f(x)$ xác định trên $\mathbb{R}$ và thỏa các tính chất sau

i) $f(x) \geq 0 \; \forall x \in \mathbb{R}$,

ii) $\int\limits_{-\infty}^{\infty} f(x)dx = 1$,

thì $f(x)$ được gọi là hàm mật độ xác suất (Probability density function - p.d.f) của biến ngẫu nhiên X. Ta có:

$$\mathbb{P}(X \in I) = \int\limits_{I} f(x)dx, \quad \forall I \subset \mathbb{R}.$$

Chú ý:

1) Mọi hàm $f(x)$ không âm, và thỏa điều kiện $\int\limits_{-\infty}^{\infty} f(x)dx = 1$ đều là hàm mật độ của 1 biến ngẫu nhiên X nào đó.
2) Từ định nghĩa về hàm mật độ ta có hàm phân phối xác suất của biến ngẫu nhiên X có hàm mật độ $f(x)$ là

$$F(x) = \mathbb{P}(X \leq x) = \int\limits_{-\infty}^{x} f(u)du$$

3)

$$F'(x) = \frac{dF(x)}{dx} = f(x)$$

Chú ý:

4) Nếu X là biến ngẫu nhiên liên tục thì xác suất X bằng một điểm cụ thể bằng 0, tức là

$$\mathbb{P}(X = c) = 0.$$

Do vậy, ta có

$$\begin{aligned}\mathbb{P}(a \leq X \leq b) &= \mathbb{P}(a < X \leq b)\\ &= \mathbb{P}(a \leq X < b)\\ &= \mathbb{P}(a < X < b).\end{aligned}$$

Ví dụ 8

Cho hàm số

$$f(x) = \begin{cases} \frac{3}{8}x^2, & \textit{khi}\, 0 \leq x \leq 2 \\ 0, & \textit{nơi khác} \end{cases}$$

(a) *Chứng tỏ $f(x)$ là hàm mật độ của một biến ngẫu nhiên X,*

(b) *Tìm hàm phân phối $F(x)$,*

(c) *Tính xác suất $\mathbb{P}(1 \leq X \leq 3/2)$.*

Ví dụ 9

Biết tuổi thọ của một thiết bị điện tử trong máy photocopy (đv: giờ) là biến ngẫu nhiên liên tục có hàm mật độ xác suất là

$$f(x) = \begin{cases} \dfrac{e^{-x/1000}}{1000} & khi\ x \geq 0, \\ 0 & khi\ x < 0. \end{cases}$$

Xác định xác suất để

(a) *Một thiết bị có tuổi thọ trên 3000 giờ,*

(b) *Một thiết bị có tuổi thọ trong khoảng từ 1000 đến 2000 giờ,*

(c) *Xác định số giờ mà có khoảng* 10% *thiết bị hỏng trước thời gian đó,*

(d) *Tìm hàm phân phối xác suất* $F(x)$.

Ví dụ 10

Thời gian để hoàn thành một phản ứng hóa học (đv: mili-giây) là biến ngẫu nhiên liên tục X có hàm phân phối xác suất

$$F(x) = \begin{cases} 1 - e^{-0.01x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

(a) *Xác định hàm mật độ xác suất của X,*

(b) *Xác suất để phản ứng hoàn thành trong vòng 200 mili-giây là bao nhiêu?*

Ví dụ 11

Một người bắn tên vào một tấm bia hình tròn tâm O, bán kính $R = 1$. Giả sử người này luôn bắn trúng bia. Gọi X là khoảng cách từ điểm tên trúng bia đến tâm O. Hãy xác định hàm phân phối xác suất của X.

Hàm của biến ngẫu nhiên rời rạc

Định nghĩa 3.1

Xét X là một biến ngẫu nhiên rời rạc có hàm khối xác suất $p_X(x)$. Nếu $u(x)$ là một hàm số của x, và Y là một biến ngẫu nhiên xác định bởi $Y = u(X)$, thì Y là một biến ngẫu nhiên rời rạc có hàm khối xác suất là

$$p_Y(y) = \sum_{x:y=u(x)} p_X(x).$$

Nếu X và Y là hai biến ngẫu nhiên độc lập, và u và v là các hàm số, thì các biến ngẫu nhiên $u(X)$ và $v(Y)$ độc lập.

Ví dụ 12

Cho biến ngẫu nhiên rời rạc X có bảng phân phối xác suất

X	-1	0	1	2
$\mathbb{P}$	0.1	0.3	0.4	0.2

(a) *Lập bảng phân phối xác suất cho biến ngẫu nhiên* $Y = 2X + 3$,

(b) *Lập bảng phân phối xác suất cho biến ngẫu nhiên* $Z = X^2$.

Ví dụ 13

Tìm hàm mật độ xác suất của biến ngẫu nhiên $Y = X^2$ với X là biến ngẫu nhiên liên tục có hàm mật độ xác suất

$$f(x) = \begin{cases} \dfrac{1}{2} & \text{khi } x \in (-1, 1) \\ 0 & \text{nơi khác} \end{cases}$$

Bởi vì $Y = X^2$ nên miền giá trị của biến ngẫu nhiên Y là $0 \leq y < 1$. Hàm phân phối xác suất $G(y)$ của Y: $G(y) = \mathbb{P}(Y \leq y)$, hay

$$G(y) = \mathbb{P}(X^2 \leq y) = \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{2} = \sqrt{y}$$

Suy ra $g(y) = \dfrac{dG(y)}{dy} = \dfrac{1}{2\sqrt{y}}$ là hàm mật độ xác suất của Y.

Ví dụ 14

Xét X là biến ngẫu nhiên liên tục có hàm mật độ xác suất

$$f(x) = \begin{cases} e^{-x}, & khi\ x > 0 \\ 0, & nơi\ khác \end{cases}$$

Tìm hàm mật độ xác suất của biến ngẫu nhiên $Y = 2\ln(X)$.

Kỳ vọng của biến ngẫu nhiên

Định nghĩa 4.1 (Kỳ vọng của biến ngẫu nhiên rời rạc)

Giả sử biến ngẫu nhiên rời rạc X có bảng phân phối xác suất

X	x_1	x_2	$\cdots$	x_n
$\mathbb{P}$	$p_X(x_1)$	$p_X(x_2)$	$\cdots$	$p_X(x_n)$

Kỳ vọng (Expectation/Expected value) *của X, ký hiệu $\mathbb{E}(X)$ được định nghĩa như sau*

$$\mathbb{E}(X) = \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i) = \sum_{i=1}^{n} x_i p_X(x_i). \tag{4}$$

Kỳ vọng $\mathbb{E}(X)$ thường được ký hiệu tắt là μ ('mu').

Nếu biến ngẫu nhiên X nhận vô hạn đếm được các giá trị $x_1, x_2, x_3, \ldots$ với các xác suất tương ứng $p_X(x_1), p_X(x_2), p_X(x_3), \ldots$ và chuỗi $\sum_{i=1}^{+\infty} x_i p_X(x_i)$ hội tụ tuyệt đối (tức là $\sum_{i=1}^{+\infty} |x_i| p_X(x_i) < \infty$) thì kỳ vọng của X được xác định bởi

$$\mathbb{E}(X) = \sum_{i=1}^{+\infty} x_i p_X(x_i).$$

Định nghĩa 4.2 (Kỳ vọng của biến ngẫu nhiên liên tục)

Xét biến ngẫu nhiên liên tục X có hàm mật độ xác suất $f(x)$, kỳ vọng của X được cho bởi

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f(x) dx. \tag{5}$$

Ý nghĩa của kỳ vọng

- Là giá trị trung bình theo xác suất của tất cả các giá trị có thể có của biến ngẫu nhiên.
- Kỳ vọng phản ánh giá trị trung tâm của phân phối xác suất.
- Xét biến ngẫu nhiên rời rạc X có phân phối xác suất $p(x_i)$, $i = 1, \ldots, n$. Nếu ta xem $x_1, \ldots, x_n$ là một hệ chất điểm mà tại đó đặt các khối lượng tương ứng $p(x_1), \ldots, p(x_n)$ thì kỳ vọng chính là trọng tâm của hệ chất điểm đó.

Kỳ vọng của biến ngẫu nhiên rời rạc

Ví dụ 15

Một hộp chứa 10 viên bi, trong đó có 3 viên bi nặng 10g, 5 viên nặng 50g, 2 viên nặng 20g. Chọn ngẫu nhiên ra 1 viên bi và gọi X là khối lượng của viên bi đó. Tính $\mathbb{E}(X)$.

Ví dụ 16

Có một trò chơi với luật chơi như sau: tung đồng thời 3 con xúc sắc. Nếu xuất hiện cả 3 mặt sáu điểm thì thắng được 10.000 đ, nếu xuất hiện 2 mặt sáu thì được 5000 đ, 1 mặt sáu thì được 1000 đ. Nếu không xuất hiện mặt sáu điểm nào thì không nhận được gì cả. Một người chơi sẽ đóng d đồng cho một lần chơi. Hỏi d phải bằng bao nhiêu để trò chơi công bằng?

Ví dụ 17 (Xét nghiệm nhóm)

Để xét nghiệm một loại bệnh nào đó (chẳng hạn Covid-19), người ta thực hiện xét nghiệm máu của một nhóm 100 người. Tuy nhiên, thay vì xét nghiệm từng người một, người ta quyết định xét nghiệm gộp cho từng nhóm, mỗi nhóm 10 người. Đối với mỗi nhóm, người ta trộn mẫu máu của 10 người lại và thực hiện xét nghiệm chung một lần. Giả sử rằng mẫu gộp sẽ cho kết quả dương tính nếu như có ít nhất một người trong nhóm đó có bệnh. Nếu như cả 10 người trong nhóm đều không mắc bệnh, ta chỉ cần thực hiện 1 xét nghiệm là đủ. Tuy nhiên, nếu như mẫu gộp cho kết quả dương tính, ta cần thực hiện thêm 10 xét nghiệm cho từng người để xác định cụ thể ai mắc bệnh. Giả sử rằng xác suất một người mắc bệnh là 0.1, *và độc lập với những người khác.*

a) *Xác suất để ta chỉ cần thực hiện chỉ 1 xét nghiệm cho một nhóm 10 người bằng bao nhiêu?*
b) *Xác suất để phải thực hiện 11 xét nghiệm cho một nhóm 10 người bằng bao nhiêu?*
c) *Tính số xét nghiệm trung bình cần thực hiện cho một nhóm 10 người.*
d) *Để tối ưu hóa số lượng xét nghiệm cần thực hiện, thay vì chia 100 người thành các nhóm 10 người, hãy thử tính số xét nghiệm trung bình cần thực hiện nếu như ta chia thành các nhóm 2, 4, 5, 20, 25, 50? Kích cỡ nhóm nào cho số lượng xét nghiệm trung bình nhỏ nhất?*

Ví dụ 18

Xét biến ngẫu nhiên X có hàm mật độ xác suất

$$f(x) = \begin{cases} C(1-x^2) & 0 \le x \le 1, \\ 0 & \textit{nơi khác.} \end{cases}$$

(a) *Tìm C.*

(b) *Tính $\mathbb{E}(X)$.*

Ví dụ 19

Giả sử thời gian hoạt động (Đv: giờ) của một thiết bị điện tử đến khi nó bị hỏng là một biến ngẫu nhiên liên tục có hàm mật độ xác suất cho bởi

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \ge 0, \\ 0 & x < 0. \end{cases}$$

Tìm thời gian hoạt động trung bình của thiết bị điện tử này.

Mệnh đề 4.1 (Tính chất của kỳ vọng)

Cho X, Y là hai biến ngẫu nhiên bất kỳ và $c \in \mathbb{R}$ thì kỳ vọng của biến ngẫu nhiên có các tính chất sau

i) $\mathbb{E}(c) = c$.

ii) $\mathbb{E}(cX) = c\mathbb{E}(X)$.

iii) $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

iv) *Nếu hai biến ngẫu nhiên X và Y độc lập thì* $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

v) *Nếu $X \geq 0$ thì* $\mathbb{E}(X) \geq 0$.

vi) *Nếu $X \leq Y$ thì* $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

Mệnh đề 4.2 (Kỳ vọng của hàm của biến ngẫu nhiên)

i) *Nếu X là biến ngẫu nhiên rời rạc có hàm xác suất là $f(x)$ thì*

$$\mathbb{E}\left[g(X)\right] = \sum_x g(x)p_X(x) \tag{6}$$

ii) *Nếu X là biến ngẫu nhiên liên tục có hàm mật độ xác suất $f(x)$ thì*

$$\mathbb{E}\left[g(X)\right] = \int_{-\infty}^{\infty} g(x)f(x)dx \tag{7}$$

Phương sai của biến ngẫu nhiên

Định nghĩa 4.3

Nếu biến ngẫu nhiên X có kỳ vọng $\mathbb{E}(X)$ thì phương sai (Variance) của X, ký hiệu $\mathbb{V}ar(X)$, được định nghĩa như sau

$$\mathbb{V}ar(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right]. \tag{8}$$

- Ký hiệu khác của phương sai: $D(X)$ hoặc $V(X)$.
- Công thức rút gọn để tính phương sai:

$$\mathbb{V}ar(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

Phương sai của biến ngẫu nhiên

- Nếu X là biến ngẫu nhiên rời rạc có hàm khối xác suất $p_X(x)$, ký hiệu $\mu = \mathbb{E}(X)$, công thức tính phương sai là

$$\mathbb{V}ar(X) = \sum_x (x-\mu)^2 p_X(x) = \sum_x x^2 p_X(x) - \mu^2$$

- Nếu X là biến ngẫu nhiên liên tục có hàm mật độ xác suất $f(x)$, ký hiệu $\mu = \mathbb{E}(X)$, công thức tính phương sai là

$$\mathbb{V}ar(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2$$

Độ lệch chuẩn

Định nghĩa 4.4

Độ lệch chuẩn (Standard deviation) của biến ngẫu nhiên X, ký hiệu σ ('sigma'), là căn bậc hai của phương sai $\mathbb{V}ar(X)$

$$\sigma = \sqrt{\mathbb{V}ar(X)}.$$

Phương sai thường được ký hiệu tắt là σ^2.

Chú ý

- Phương sai $\mathbb{V}ar(X)$ có cùng đơn vị đo với bình phương của X, tức là, nếu X đo bằng m thì đơn vị của $\mathbb{V}ar(X)$ là m^2.
- Độ lệch chuẩn σ có cùng đơn vị đo với X. Do vậy, độ lệch chuẩn là độ đo tự nhiên của cho sự phân tán của biến ngẫu nhiên.

Phương sai của biến ngẫu nhiên

Ý nghĩa của Phương sai:

- Phương sai biểu thị độ tập trung hay phân tán của các giá trị của một biến ngẫu nhiên X xung quanh giá trị kỳ vọng (trung bình) của nó. Nếu phương sai nhỏ thì các giá trị của X tập trung gần kỳ vọng, nếu phương sai lớn thì các giá trị của X phân tán nhiều.
- Trong công nghiệp, nếu X là kích cỡ của một sản phẩm (ví dụ chiều dài) thì phương sai biểu thị độ chính xác tương ứng với kích cỡ đó. Trong canh tác, phương sai biểu thị mức độ ổn định của năng suất. Trong đo lường, phương sai thể hiện độ "ổn định" của phép đo, . . .

Ví dụ 20

Tính phương sai của các biến ngẫu nhiên trong Ví dụ 15 và Ví dụ 18.

Mệnh đề 4.3 (Tính chất của phương sai)

Cho X, Y là hai biến ngẫu nhiên bất kỳ và $c \in \mathbb{R}$ thì phương sai của biến ngẫu nhiên có các tính chất sau

i) $\mathbb{V}ar(c) = 0$,

ii) $\mathbb{V}ar(cX) = c^2\mathbb{V}ar(X)$,

iii) $\mathbb{V}ar(X + c) = \mathbb{V}ar(X)$,

iv) $\mathbb{V}ar(X + Y) = \mathbb{V}ar(X) + \mathbb{V}ar(Y)$ *nếu* X, Y *độc lập.*

Trung vị

Định nghĩa 4.5

Trung vị (median) của biến ngẫu nhiên X, kí hiệu m_X hoặc $\mathrm{med}(X)$, là giá trị của X thỏa

- $\mathbb{P}(X \leq m_X) \geq 1/2$ *và* $\mathbb{P}(X \geq m_X) \geq 1/2$ *nếu X rời rạc,*
- $\mathbb{P}(X \leq m_X) = F_X(m_X) = 1/2$ *nếu X liên tục.*

Đối với biến ngẫu nhiên liên tục, trung vị là hoành độ của điểm mà tại đó diện tích giới hạn bởi hàm mật độ xác suất được chia làm hai phần có diện tích bằng nhau.

Ví dụ 21

Tìm trung vị của biến ngẫu nhiên trong Ví dụ 18.

Yếu vị

Định nghĩa 4.6

Yếu vị (Mode) của biến ngẫu nhiên X, kí hiệu $\mathrm{mode}(X)$, là giá trị của nó mà tại đó có

- *xác suất cực đại nếu X rời rạc.*
- *mật độ xác suất cực đại nếu X liên tục.*

Một biến ngẫu nhiên có thể có một hay nhiều mode.

Bài tập 1

Cho biến ngẫu nhiên rời rạc X có bảng phân phối xác suất

X	a	0.1	0.3	0.4	2
$\mathbb{P}$	0.3	0.2	0.2	b	0.1

(a) Xác định a, b biết $\mathbb{E}(X) = 0.3$,

(b) Tìm hàm phân phối xác suất của X.

Theo thống kê, một người Mỹ 25 tuổi sẽ sống thêm trên 1 năm với xác suất là 0.992 và người đó chết trong vòng 1 năm tới là 0.008. Một công ty bảo hiểm A đề nghị người đó bảo hiểm sinh mạng cho 1 năm với số tiền chi trả là 10000 USD, phí bảo hiểm là 100 USD. Hỏi trung bình công ty A lãi bao nhiêu khi bán bảo hiểm cho người đó?

Bài tập 3

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. Biết rằng 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 kiếm được bao nhiêu tiền chép tranh mỗi tuần?

Bài tập 4

Tuổi thọ X (năm) của người dân ở một địa phương là biến ngẫu nhiên có hàm phân phối xác suất như sau

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{khi } x > 0 \\ 0, & \text{khi } x \leq 0 \end{cases}$$

với $\lambda = 0.013$. Tính

(a) Tỷ lệ người dân thọ từ 60 đến 70 tuổi,

(a) Hàm mật độ của X,

(c) Tuổi thọ trung bình và $\mathbb{V}ar(X)$.

Bài tập 5

Tuổi thọ X (tháng) của một bộ phận trong một dây chuyền sản xuất là biến ngẫu nhiên liên tục có hàm mật độ

$$f(x) = \begin{cases} \frac{25}{2}(x+10)^{-2}, & \text{khi } x \in (0, 40) \\ 0, & \text{khi } x \notin (0, 40) \end{cases}$$

Tính

(a) Xác suất tuổi thọ của bộ phận này nhỏ hơn 6 tháng,

(a) Tuổi thọ trung bình của bộ phận này,

(c) Tìm hàm phân phối xác suất của X.