2.1 ĐỊNH NGHĨA VÀ PHÂN LOẠI BIẾN NGẪU NHIÊN

2.1.1 Định nghĩa biến ngẫu nhiên

Biến ngẫu nhiên X là đại lượng nhận các giá trị nào đó phụ thuộc vào các yếu tố ngẫu nhiên, nghĩa là với mọi giá trị thực $x \in$ R thì "X nhận giá trị nhỏ hơn bằng x", ký hiệu $\{X \le x\}$, là một biến cố ngẫu nhiên.

Tập hợp tất cả các giá trị của X được gọi là *miền giá trị* của X, ký hiệu R_X

Ví dụ 2.1: Nếu gọi X là tổng số chấm xuất hiện khi gieo hai con xúc xắc thì X là một biến ngẫu nhiên có miền giá trị

$$R_X = \{2, 3, \ldots, 12\}$$

và $\{X = k\} = A_k \,; k = 2, 3, \ldots, 12$

Định nghĩa 2.2: Hai biến ngẫu nhiên X, Y là độc lập nếu X nhận các giá trị nào đó không phụ thuộc Y và ngược lại. Nói cách khác với mọi số thực x, y hai biến cố $\{X \le x\}$, $\{Y \le y\}$ là độc lập.

2.1.2 Hàm phân bố xác suất

Các biến ngẫu nhiên được xét trong các phép thử khác nhau (tương ứng với các không gian xác suất khác nhau) nhưng các phân bố xác suất của chúng có thể như nhau.

Phân bố xác suất được nghiên cứu thông qua hàm phân bố xác suất

Hàm phân bố xác suất (cumulative distribution function, viết tắt CDF) của biến ngẫu nhiên X là hàm số $F_X(x)$ xác định với mọi $x \in \mathrm{R}$ bởi công thức:

$$F_X(x) = P\left\{X \le x\right\}; \ -\infty < x < \infty$$

trong đó $\{X \le x\}$ là ký hiệu biến cố "biến ngẫu nhiên X nhận giá trị nhỏ hơn hay bằng x".

Các tính chất của hàm phân bố

1. $0 \le F_X(x) \le 1$ với mọi $x \in \mathrm{R}$

2. $F_X(x)$ là hàm không giảm, liên tục bên phải

   Nếu X là biến ngẫu nhiên liên tục thì $F_X(x)$ là hàm liên tục.

3. $F_X(-\infty) = \lim\limits_{x \to -\infty} F_X(x) = 0; \; F_X(+\infty) = \lim\limits_{x \to +\infty} F_X(x) = 1$

4. $P\{a < X \le b\} = F_X(b) - F_X(a) \quad P\{X > a\} = 1 - F_X(a)$

   $P\{X < a\} = F_X(a^-) = \lim\limits_{x < a, x \to a} F_X(x)$

Ví dụ 2.3: Một nguồn thông tin sinh ra các ký hiệu ngẫu nhiên từ bốn ký tự $\{a, b, c, d\}$ với xác suất $P(a)=1/2$, $P(b)=1/4$ và $P(c)=P(d)=1/8$.

Mã hóa các ký hiệu này theo các mã nhị phân sau

a	0
b	10
c	110
d	111

Đặt X là biến ngẫu nhiên ký hiệu độ dài của mã, đó là số các bit

1. Tìm miền giá trị của X
2. Giả sử các ký hiệu được sinh độc lập. Tính các xác suất
$$P\{X=1\}, P\{X=2\} \text{ và } P\{X=3\}$$
3. Tìm hàm phân bố $F_X(x)$ và vẽ đồ thị

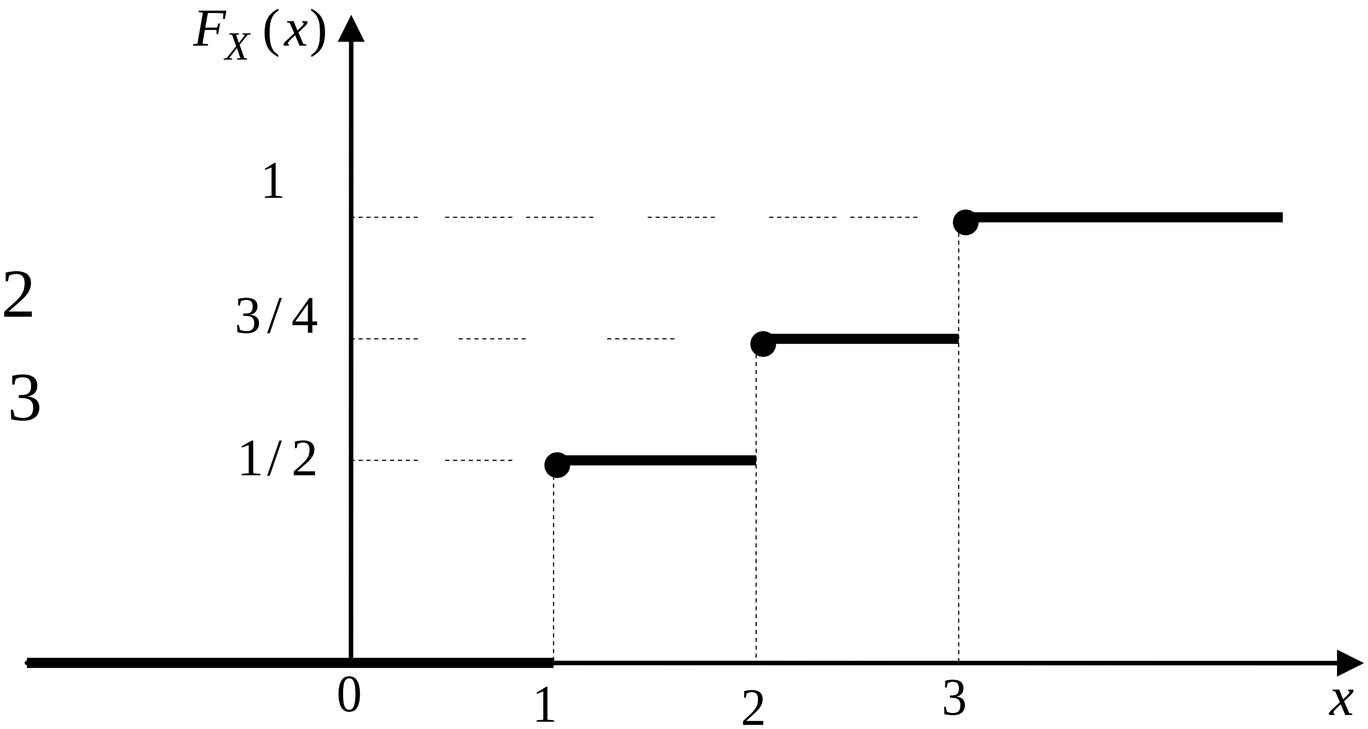
1. Miền giá trị $R_X = \{1, 2, 3\}$

2. $P\{X = 1\} = P(a) = \frac{1}{2}$; $P\{X = 2\} = P(b) = \frac{1}{4}$;

$$P\{X = 3\} = P\{c, d\} = P(c) + P(d) = \frac{1}{4}$$

3.

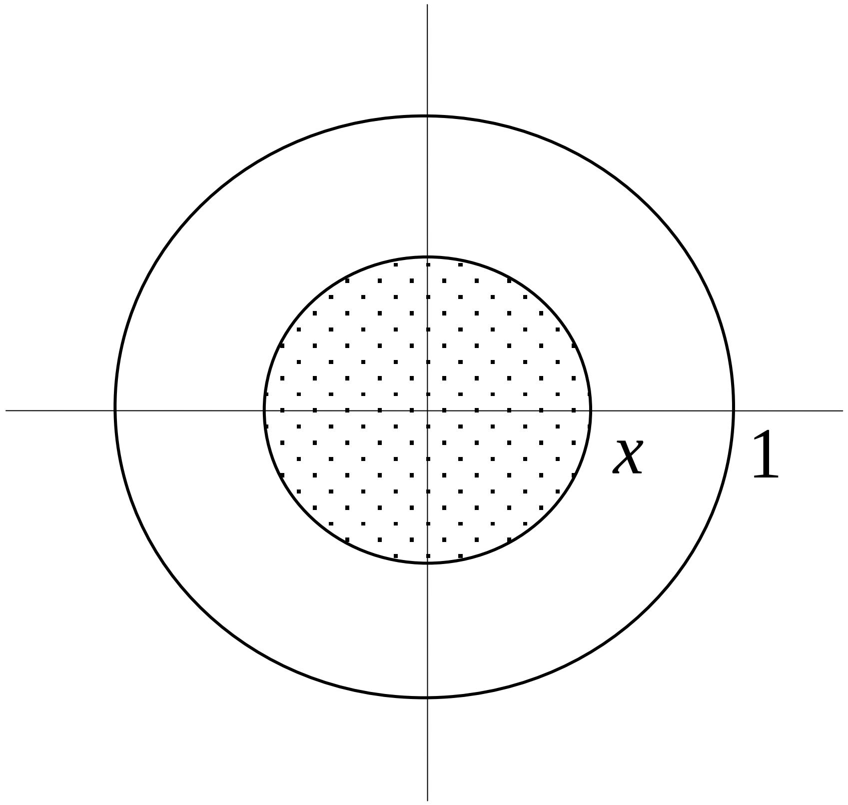
$$F_X(x) = \begin{cases} 0 & x < 1 \\ 1/2 & 1 \le x < 2 \\ 3/4 & 2 \le x < 3 \\ 1 & x \ge 3 \end{cases}$$



Ví dụ 2.4: Xét phép thử ném phi tiêu vào một đĩa tròn có bán kính bằng 1. Ký hiệu X là biến ngẫu nhiên đo khoảng cách từ điểm mũi phi tiêu cắm vào đĩa đến tâm của đĩa. Giả sử mũi phi tiêu luôn cắm vào đĩa và đồng khả năng tại mọi điểm của đĩa.

1. Tìm miền giá trị của X

2. Tìm hàm phân bố $F_X(x)$ và vẽ đồ thị

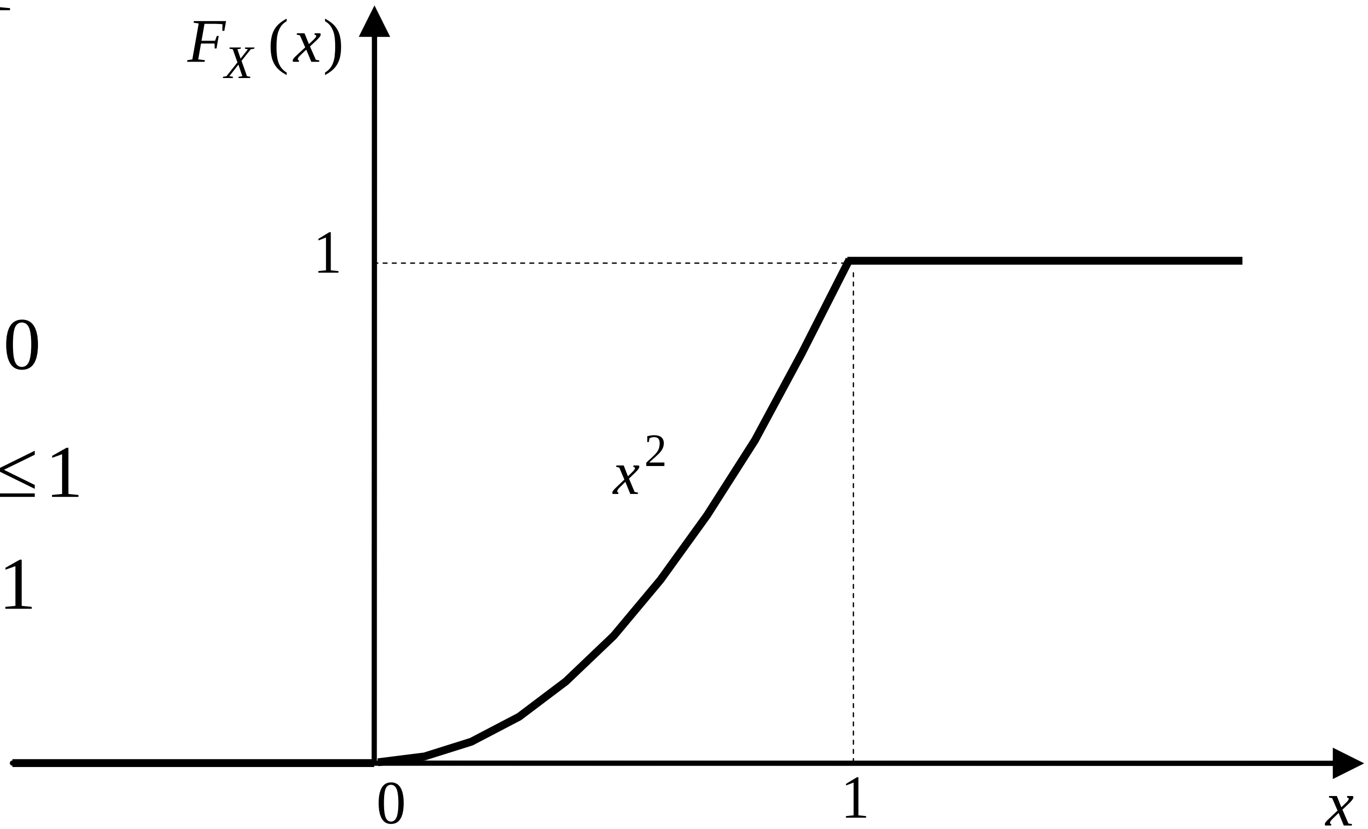


1. Miền giá trị $R_X = \left\{ x \in R \middle| 0 \le x < 1 \right\}$

2. $P\{X \le x\} = \dfrac{\pi . x^2}{\pi . 1^2} = x^2$

Hàm phân bố

$$F_X(x) = \begin{cases} 0 & x \le 0 \\ x^2 & 0 < x \le 1 \\ 1 & x > 1 \end{cases}$$



2.1.3 Phân loại

❖ ***Biến ngẫu nhiên rời rạc***

Biến ngẫu nhiên X là rời rạc nếu miền giá trị gồm một số hữu hạn hoặc vô hạn đếm được các giá trị, nghĩa là có thể liệt kê các giá trị của miền giá trị thành một dãy. Do đó hàm phân bố có đồ thị dạng hình thang.

❖ ***Biến ngẫu nhiên liên tục***

X là biến ngẫu nhiên liên tục nếu miền giá trị của nó có thể lấp đầy một hoặc một số các khoảng hữu hạn hoặc vô hạn và xác suất biến ngẫu nhiên nhận giá trị tại từng điểm đều bằng 0 (nghĩa là $P\{X = a\}=0$ với mọi a). Do đó hàm phân bố là hàm số liên tục.

Ví dụ 2.6:

- Gọi X là số chấm xuất hiện khi gieo một con xúc xắc thì X là biến ngẫu nhiên rời rạc nhận các giá trị 1, 2, 3, 4, 5, 6.
- Gọi T là tuổi thọ của một thiết bị đang hoạt động thì T là biến ngẫu nhiên liên tục nhận giá trị trong một khoảng.
- Gọi Z là số khách hàng vào một điểm phục vụ trong 1 đơn vị thời gian, Z là biến ngẫu nhiên rời rạc nhận các giá trị 0, 1, 2, …
- Số cuộc gọi đến một tổng đài là biến ngẫu nhiên rời rạc nhận các giá trị 0, 1, 2, …
- Sai số Y khi đo lường một đại lượng vật lý nào đó là biến ngẫu nhiên liên tục nhận giá trị trong một khoảng.

2.2 BIẾN NGẪU NHIÊN RỜI RẠC

2.2.1 Hàm khối lượng xác suất và bảng phân bố xác suất của biến ngẫu nhiên rời rạc

Hàm $$p_X(x) = P\{X = x\}$$

được gọi là **hàm khối lượng xác suất** (probability mass function) của biến ngẫu nhiên rời rạc X

Hàm phân bố của X

$$F_X(x) = P\{X \le x\} = \sum_{x_k \le x;\, x_k \in R_X} p_X(x_k)$$

Tính chất của hàm khối lượng xác suất

1. $\sum_{x_k \in R_X} p_X(x_k) = 1$

2. $p_X(x_k) > 0$, với mọi $x_k \in R_X$

3. $p_X(x) = 0$, với mọi $x \notin R_X$

Nếu $R_X = \{x_1, x_2, ...\}$ thì hàm phân bố xác suất có dạng:

$$F_X(x) = \begin{cases} 0 & khi \quad x < x_1 \\ p_X(x_1) + \cdots + p_X(x_{k-1}) & \text{khi} \quad x_{k-1} \le x < x_k , \ \forall k > 1 \end{cases}$$

Bảng phân bố xác suất

Bảng phân bố xác suất có hai hàng, hàng trên ghi các giá trị mà biến ngẫu nhiên nhận được, hàng dưới là giá trị của hàm khối lượng xác suất tương ứng

Bảng phân bố xác suất của biến ngẫu nhiên X có dạng

$\boldsymbol{X}$	x_1	x_2	• • •
$\boldsymbol{P}$	$p_X(x_1)$	$p_X(x_2)$	• • •

Ví dụ 2.8: Chọn ngẫu nhiên 3 bi từ một túi có 6 bi đen, 4 bi trắng. Gọi X là số bi trắng trong 3 bi vừa chọn thì X là một biến ngẫu nhiên rời rạc. Tìm bảng phân bố xác suất và hàm phân bố xác suất.

$$P\{X=0\}=\frac{C_6^3}{C_{10}^3}=\frac{5}{30},\ P\{X=1\}=\frac{C_6^2C_4^1}{C_{10}^3}=\frac{15}{30}$$

$$P\{X=2\}=\frac{C_6^1C_4^2}{C_{10}^3}=\frac{9}{30},\ P\{X=3\}=\frac{C_4^3}{C_{10}^3}=\frac{1}{30}$$

Hàm khối lượng xác suất

$$p_X(0)=\frac{5}{30},\ p_X(1)=\frac{15}{30},\ p_X(2)=\frac{9}{30},\ p_X(3)=\frac{1}{30}$$

$p_X(x)=0$ với mọi x khác 0, 1, 2, 3

Bảng phân bố xác suất

X	0	1	2	3
P	5/30	15/30	9/30	1/30

Hàm phân bố xác suất

$$F_X(x) = \begin{cases} 0 & \text{khi} \quad x < 0 \\ 5/30 & \text{khi} \quad 0 \le x < 1 \\ 20/30 & \text{khi} \quad 1 \le x < 2 \\ 29/30 & \text{khi} \quad 2 \le x < 3 \\ 1 & \text{khi} \quad x \ge 3 \end{cases}$$

2.2.2 Các phân bố rời rạc thường gặp

2.2.2.1 Phân bố Bernoulli

Biến ngẫu nhiên rời rạc X nhận hai giá trị 0, 1 với hàm khối lượng xác suất

$$p_X(k) = P\{X = k\} = p^k q^{1-k};\ k = 0,1$$

trong đó $0 < p < 1,\ q = 1 - p,$

được gọi là có phân bố Bernoulli tham số p .

Xét phép thử Bernoulli với sự thành công của phép thử là sự xuất hiện của biến cố A với xác suất xuất hiện là p. Gọi X là số lần thành công trong một lần thử thì X là biến ngẫu nhiên rời rạc có phân bố Bernoulli tham số p.

2.2.2.2 Phân bố nhị thức $\mathcal{B}(n;p)$

Biến ngẫu nhiên rời rạc X nhận các giá trị $0, 1, \ldots, n$ với xác suất tương ứng

$$p_X(k) = P\{X = k\} = C_n^k p^k q^{n-k};\ k = 0, 1, \ldots, n$$

trong đó n là số tự nhiên và $0 < p < 1$, $q = 1 - p$, được gọi là có phân bố nhị thức tham số n, p, ký hiệu $X \sim \mathcal{B}(n;p)$

Hàm phân bố

$$F_X(x) = \begin{cases} 0 & \text{khi } x < 0 \\ \sum_{k=0}^{m} C_n^k p^k q^{n-k} & \text{khi } m \le x < m+1,\ 0 \le m \le n-1 \\ 1 & \text{khi } x \ge n \end{cases}$$

Thực hiện n phép thử Bernoulli với xác suất thành công của biến cố A trong mỗi lần thử là p

Với mỗi $i = 1, 2, \ldots, n$; nếu ở lần thử thứ i biến cố A xuất hiện ta cho X_i nhận giá trị 1, nếu biến cố A không xuất hiện ta cho X_i nhận giá trị 0

Như vậy X_i là biến ngẫu nhiên rời rạc có phân bố Bernoulli tham số p

Gọi X là số thành công trong n phép thử Bernoulli này thì

$$X = X_1 + X_2 + \cdots + X_n \sim \mathrm{B}(n; p)$$

Ví dụ 2.11: Một nguồn nhị phân phát ra hai ký số (digit) 1 và 0 một cách ngẫu nhiên với xác suất tương ứng 0,6 và 0,4.

1. Tính xác suất có đúng hai ký số 1 và ba ký số 0 trong dãy có năm ký số
2. Tính xác suất có đúng hai ký số 1 và ba ký số 0 trong dãy có năm ký số

Giải: Gọi X là số các ký số 1 trong dãy có năm ký số

$$X \sim \mathrm{B}(5;0,6)$$

1. $P\{X=2\}=C_5^2(0,6)^2(0,4)^3=0,23$

2. $P\{X\le 2\}=\sum_{k=0}^{2}C_5^k(0,6)^k(0,4)^{5-k}=0,317$

$$P\{X\ge 3\}=1-P\{X\le 2\}=1-0,317=0,683$$

2.2.2.3 Phân bố Poisson

Biến ngẫu nhiên X được gọi là có phân bố Poisson tham số $\lambda>0$, ký hiệu $X \sim \mathcal{P}(\lambda)$, nếu X nhận các giá trị $k = 0, 1, 2, \ldots$ với hàm khối lượng xác suất

$$p_X(k) = P\{X = k\} = e^{-\lambda}\frac{\lambda^k}{k!}; \lambda > 0;\ k = 0, 1, 2, \ldots$$

Hàm phân bố

$$F_X(x) = e^{-\lambda}\sum_{k=0}^{n}\frac{\lambda^k}{k!},\quad n \le x < n+1$$

Trong thực tế với một số giả thiết thích hợp thì các biến ngẫu nhiên là các quá trình đếm sau:

1. Số cuộc gọi đến một tổng đài,
2. Số khách hàng đến 1 điểm phục vụ,
3. Số xe cộ qua 1 ngã tư,
4. Số tai nạn (xe cộ); số các sự cố xảy ra ở một địa điểm

…

trong một khoảng thời gian xác định nào đó sẽ có phân bố Poisson với tham số λ, trong đó λ là tốc độ trung bình diễn ra trong khoảng thời gian này.

Ví dụ 2.12: Ở một tổng đài điện thoại các cuộc gọi đến một cách ngẫu nhiên, độc lập. Ký hiệu $X(t)$ là số cuộc gọi đến tổng đài trong khoảng thời gian t phút.

Có thể chứng minh được $X(t)$ có phân bố Poisson tham số λt, trong đó λ là số cuộc gọi trung bình trong 1 phút.

Giả sử trung bình có 2 cuộc gọi trong 1 phút. Tìm xác suất:

1. Có đúng 5 cuộc gọi đến trong 2 phút (biến cố A).
2. Không có một cuộc gọi nào trong 30 giây (biến cố B).
3. Có ít nhất 1 cuộc gọi trong 10 giây (biến cố C).

Giải: Theo giả thiết $\lambda=2$, vậy ta có

1. $X(2) \sim \mathcal{P}(4)$, do đó

$$P(A) = P\{X(2) = 5\} = e^{-4}\frac{4^5}{5!} \approx 0{,}156$$

2. $X(1/2) \sim \mathcal{P}(1)$, do đó

$$P(B) = P\{X(1/2) = 0\} = e^{-1} \approx 0{,}3679$$

3. $X(1/6) \sim \mathcal{P}(1/3)$, do đó

$$P(C) = P\{X(1/6) \geq 1\} = 1 - P\{X(1/6) = 0\} = 1 - e^{-1/3} \approx 0{,}2835$$

2.3 BIẾN NGẪU NHIÊN LIÊN TỤC

2.3.1 Hàm mật độ xác suất của biến ngẫu nhiên liên tục

Giả sử X là một biến ngẫu nhiên liên tục có hàm phân bố xác suất $F_X(x)$

Hàm $f_X(x)$ thỏa mãn

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt, \text{ với mọi } x \in R$$

được gọi là **hàm mật độ xác suất** của biến ngẫu nhiên X (probability density function, viết tắt PDF)

Tính chất của hàm mật độ xác suất

1. $F_X^{'}(x) = f_X(x)$ tại các điểm x mà $f_X(x)$ liên tục

   Vậy hàm phân bố của biến ngẫu nhiên liên tục là một nguyên hàm của hàm mật độ, và hàm mật độ là đạo hàm của hàm phân bố

2. $f_X(x) \geq 0$ với mọi $x \in R$

3. $\int_{-\infty}^{\infty} f_X(x)dx = 1$

4. $$P\{a < X < b\} = P\{a \leq X \leq b\} = P\{a \leq X < b\} = P\{a < X \leq b\} = \int_a^b f_X(x)dx$$

Ví dụ 2.13: Hàm phân bố xác suất của biến ngẫu nhiên liên tục X có dạng

$$F_X(x) = \begin{cases} 0 & khi \quad x < 0 \\ kx^2 & khi \quad 0 \le x < 1 \\ 1 & khi \quad x \ge 1 \end{cases}$$

Xác định hệ số k và tìm hàm mật độ xác suất

Giải: Từ tính chất liên tục của hàm phân bố xác suất của biến ngẫu nhiên liên tục, ta có $k=1$

Hàm mật độ xác suất

$$f_X(x) = \begin{cases} 0 & khi \quad x \le 0 \\ 2x & khi \quad 0 < x < 1 \\ 0 & khi \quad x \ge 1 \end{cases}$$

Ví dụ 2.14: Biến ngẫu nhiên liên tục X với hàm mật độ xác suất có dạng

$$f_X(x) = \begin{cases} 0 & khi \quad x < 1 \\ \dfrac{k}{x^2} & khi \quad x \geq 1 \end{cases}$$

Hãy xác định: Hệ số k; Hàm phân bố $F_X(x)$ và $P\{2 < X < 3\}$

$$1 = \int_{-\infty}^{\infty} f_X(x)dx = \int_{1}^{\infty} \frac{k}{x^2} dx = -\lim_{a\to\infty}\left(\frac{k}{x}\bigg|_1^a \right) = k \quad \Rightarrow k = 1$$

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \begin{cases} 0 & \text{ví i} \quad x < 1 \\ 1 - \dfrac{1}{x} & \text{ví i} \quad x \geq 1 \end{cases}$$

$$P\{2 < X < 3\} = F_X(3) - F_X(2) = \left(1 - \frac{1}{3}\right) - \left(1 - \frac{1}{2}\right) = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

2.3.2 Các phân bố liên tục thường gặp

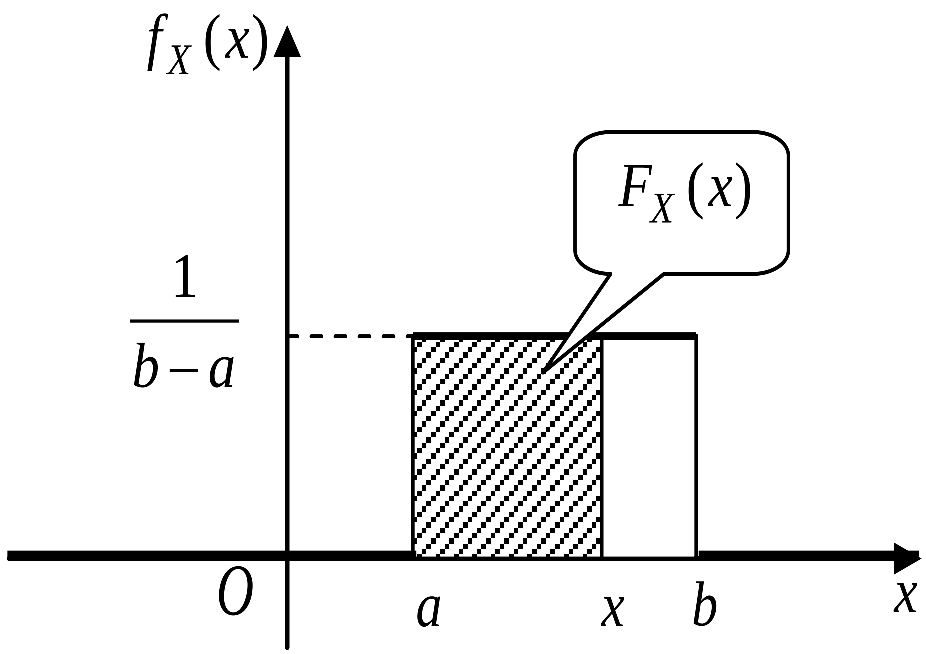
2.3.2.1 Phân bố đều U(a; b)

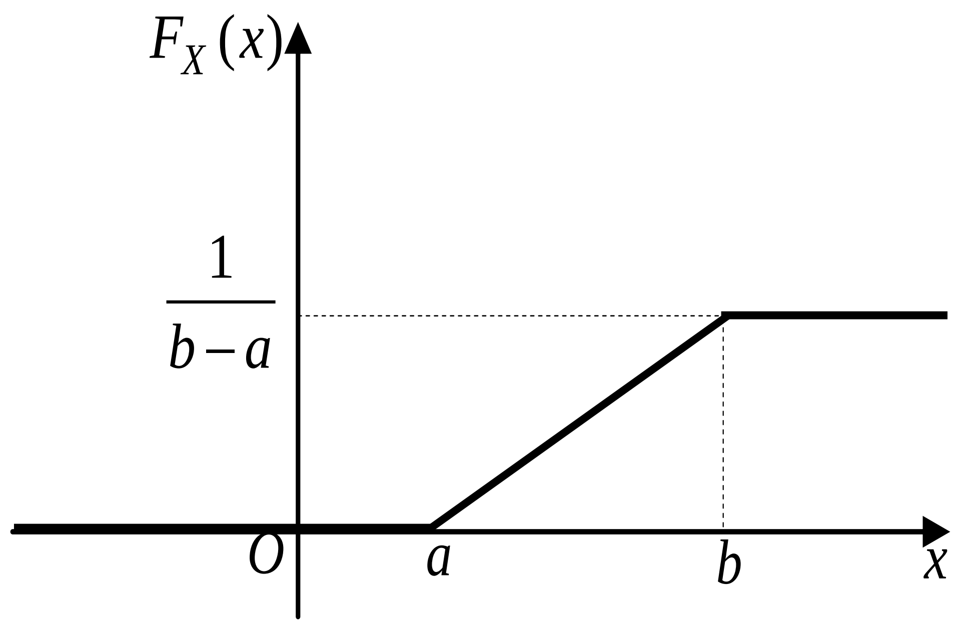
Biến ngẫu nhiên X được gọi là có phân bố đều trong khoảng $(a; b)$ nếu hàm mật độ xác suất của nó xác định bởi

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & khi \quad a < x < b \\ 0 & khi \quad x \notin (a,b) \end{cases}$$

Hàm phân bố xác suất

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \begin{cases} 0 & khi \quad x \le a \\ \dfrac{x-a}{b-a} & khi \quad a < x < b \\ 1 & khi \quad x \ge b \end{cases}$$


$f_X(x)$
$F_X(x)$
$\frac{1}{b-a}$
O
a
x
b
x


$F_X(x)$
$\frac{1}{b-a}$
O
a
b
x

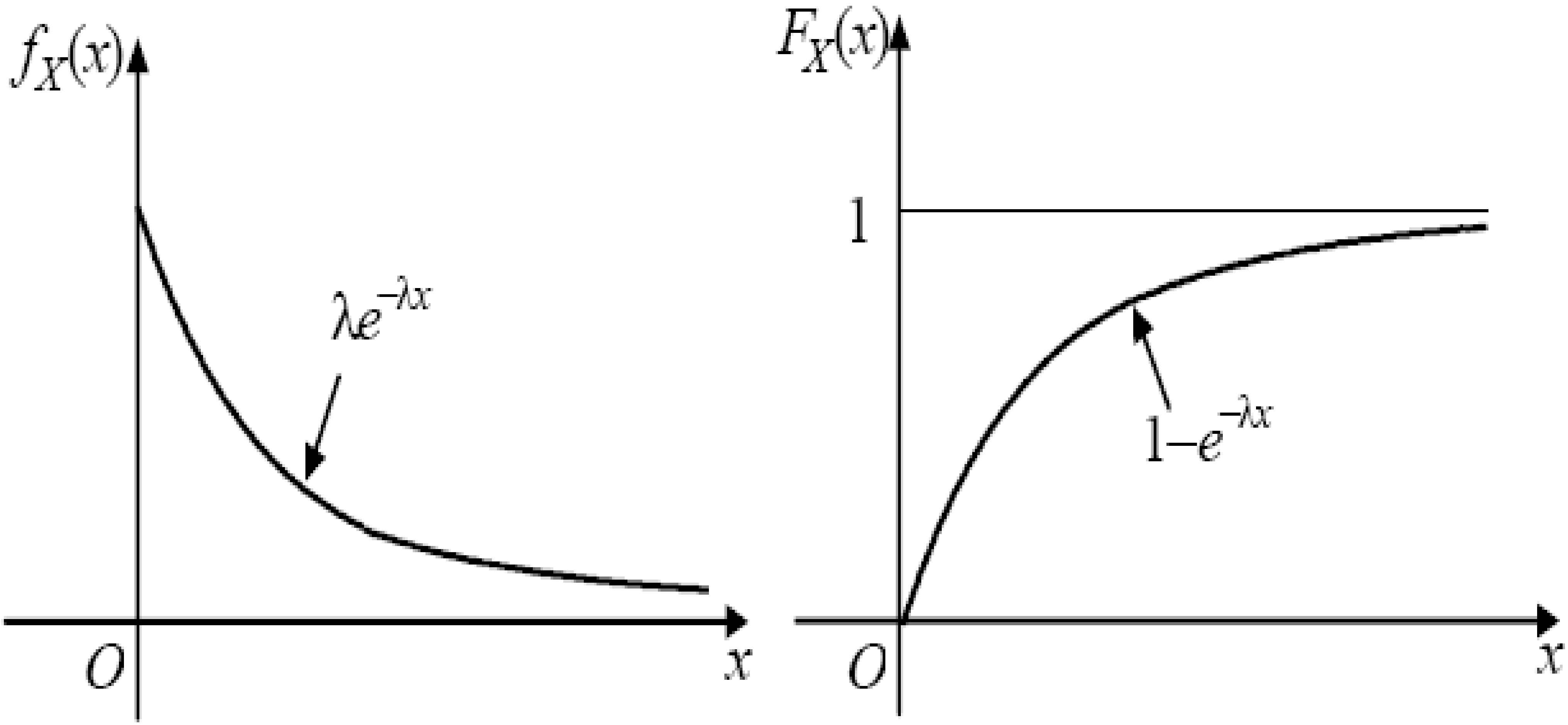
2.3.2.2 Phân bố mũ

Biến ngẫu nhiên liên tục X được gọi là có phân bố mũ tham số $\lambda > 0$ nếu hàm mật độ xác suất xác định như sau

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & khi \quad x > 0 \\ 0 & khi \quad x \le 0 \end{cases}$$

Hàm phân bố xác suất

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \begin{cases} 1 - e^{-\lambda x} & khi \quad x > 0 \\ 0 & khi \quad x \le 0 \end{cases}$$

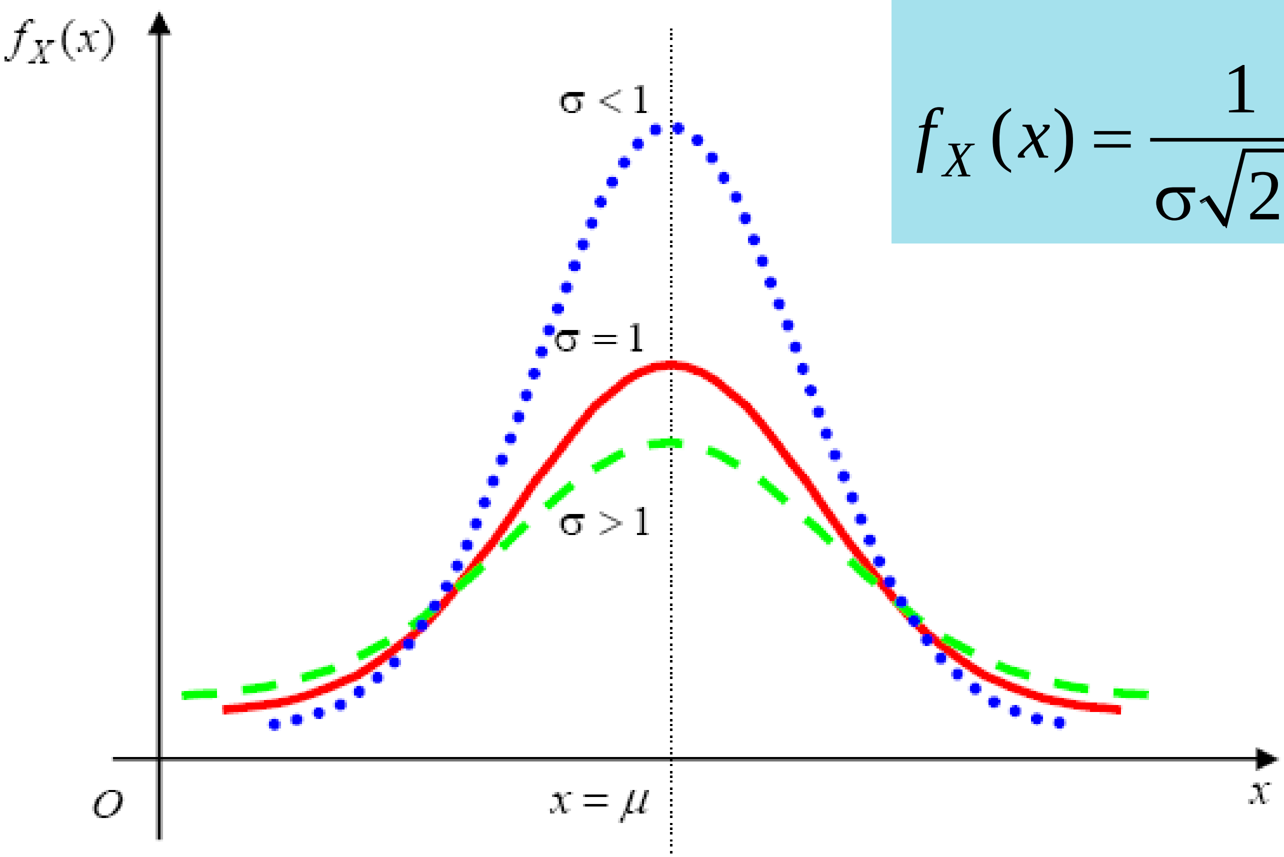


$f_X(x)$
$\lambda e^{-\lambda x}$
O
x
$F_X(x)$
1
$1-e^{-\lambda x}$
O
x

2.3.2.4 Phân bố chuẩn $\mathbf{N}(\mu;\sigma^2)$

Biến ngẫu nhiên liên tục X có phân bố chuẩn $\mathbf{N}(\mu;\sigma^2)$, ký hiệu $X \sim \mathbf{N}(\mu;\sigma^2)$, nếu hàm mật độ xác suất có dạng

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \; ; \; \forall x \in R$$



Phân bố chuẩn tắc $\mathbf{N}(0;1)$

Hàm mật độ xác suất

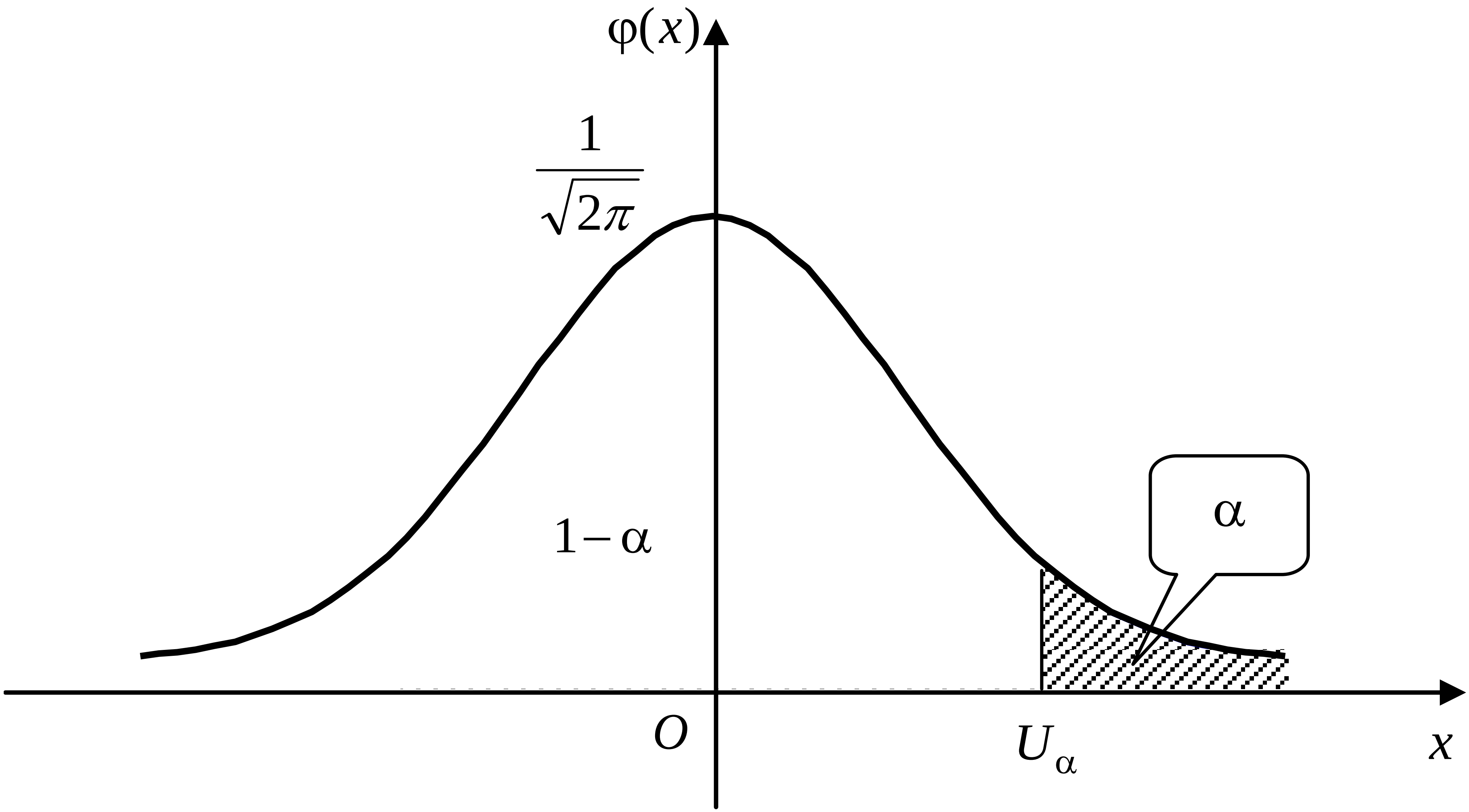
$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \forall x \in R$$

Hàm phân bố xác suất

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt, \quad \forall x \in R$$

Giá trị U_α gọi là **giá trị tới hạn mức α** của phân bố chuẩn tắc nếu

$$\Phi^{-1}(1-\alpha) = U_\alpha$$



$\varphi(x)$
$\frac{1}{\sqrt{2\pi}}$
$1-\alpha$
α
O
U_α
x

Các tính chất của hàm phân bố xác suất $\Phi(x)$

1. $\Phi(x) + \Phi(-x) = 1, \quad \Phi(-x) = 1 - \Phi(x).$

2. Nếu $X \sim \mathbf{N}(0;1)$ thì

$$\forall a > 0,\ P\{|X| < a\} = 2\Phi(a) - 1,\ P\{|X| > a\} = 2\left(1 - \Phi(a)\right)$$

$$P\{X > U_\alpha\} = \alpha\ ;\ \ P\left\{|X| > U_{\frac{\alpha}{2}}\right\} = \alpha\ ;\ \ P\left\{|X| < U_{\frac{\alpha}{2}}\right\} = 1 - \alpha$$

3. Nếu $X \sim \mathbf{N}(\mu;\sigma^2)$ thì $\dfrac{X-\mu}{\sigma} \sim \mathbf{N}(0;1)$

$$F_X(x) = P\{X \le x\} = P\left\{\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right\} = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$P\{a < X < b\} = P\{a < X \le b\} = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Ví dụ 2.17: Giả sử $X \sim \mathbf{N}(\mu;\sigma^2);\mu = 2100,\ \sigma = 200$. Hãy tìm:

1. $P\{X < 2400\}$, $P\{1700 < X < 2200\}$
2. Xác định a để $P\{X > a\} = 0{,}03$

Giải:

$$P\{X < 2400\} = P\{X \le 2400\} = \Phi\left(\frac{2400-2100}{200}\right) = \Phi(1,5) = 0,9332$$

$$P\{1700 < X < 2200\} = \Phi\left(\frac{2200-2100}{200}\right) - \Phi\left(\frac{1700-2100}{200}\right)$$

$$= \Phi(0{,}5) - \Phi(-2) = 0,6688$$

$$P\{X > a\} = 1 - \Phi\left(\frac{a-2100}{200}\right) = 0,03 \Rightarrow \Phi\left(\frac{a-2100}{200}\right) = 0,97$$

?

2.3.2.5 Phân bố "Chi (khi) bình phương"

Biến ngẫu nhiên liên tục X có phân bố "Chi (khi) bình phương" n bậc tự do, ký hiệu $X \sim \chi_n^2$ nếu hàm mật độ xác suất có dạng

$$f_X(x) = \begin{cases} \dfrac{(x/2)^{n/2-1}}{\Gamma(n/2)} \dfrac{1}{2} e^{-(x/2)} & khi \quad x > 0 \\ 0 & khi \quad x \le 0 \end{cases}$$

trong đó, $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt,\ x > 0$ là hàm Gamma.

Nếu là các biến ngẫu nhiên độc lập có cùng phân bố chuẩn tắc **N**(0;1) thì $X_1, X_2, \ldots, X_n$

$$\sum_{i=1}^{n} X_i^2 = X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi_n^2$$

2.3.2.6 Phân bố Student $\mathbf{T}(n)$

Biến ngẫu nhiên liên tục T có phân bố Student n bậc tự do, ký hiệu $T \sim \mathbf{T}(n)$, nếu hàm mật độ xác suất có dạng

$$f_T(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma(n/2)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < x < \infty$$

Có thể chứng minh được rằng nếu $Z \sim \mathbf{N}(0;1)$, $V \sim \chi_n^2$; Z và V độc lập thì

$$T = \frac{Z}{\sqrt{V/n}} \sim \mathbf{T}(n)$$

Giá trị tới hạn mức α của phân bố "khi bình phương" n bậc tự do, ký hiệu $\chi^2_\alpha(n)$, được xác định như sau

$$P\left\{\chi^2 > \chi^2_\alpha(n)\right\} = \alpha$$

Giá trị tới hạn mức α của phân bố Student n bậc tự do, ký hiệu $t_\alpha(n)$, được xác định như sau

$$P\left\{T > t_\alpha(n)\right\} = \alpha$$

Bảng tính các giá trị tới hạn $t_\alpha(n)$ cho trong Phụ lục III và $\chi^2_\alpha(n)$ trong Phụ lục IV

2.4 CÁC THAM SỐ ĐẶC TRƯNG CỦA BIẾN NGẪU NHIÊN

2.4.1 Kỳ vọng toán

Kỳ vọng của biến ngẫu nhiên X ký hiệu $\mathrm{E}X$ và xác định như sau

$$\mathrm{E}\,X = \sum_{x_i \in R_X} x_i\, p_X(x_i) \quad \text{nếu } X \text{ rời rạc}$$

$$\mathrm{E}\,X = \int_{-\infty}^{\infty} x f_X(x)\,dx \quad \text{nếu } X \text{ liên tục}$$

Ví dụ 2.19: Theo thống kê việc một người Mỹ 25 tuổi sẽ sống thêm trên một năm có xác suất là 0,992, còn xác suất để người đó chết trong vòng một năm tới là 0,008. Một chương trình bảo hiểm đề nghị người đó bảo hiểm sinh mạng cho 1 năm với số tiền chi trả 1000 đô la, còn tiền đóng là 10 đô la. Hỏi lợi nhuận của công ty bảo hiểm nhận được là bao nhiêu?

Giải:

Lợi nhuận là biến ngẫu nhiên X với 2 giá trị là +10 đô la (nếu người bảo hiểm không chết) và –990 đô la (nếu người đó chết). Bảng phân bố xác suất tương ứng

X	–990	+10
P	0,008	0,992

$$\mathrm{E}\,X = (-990)\cdot 0{,}008 + 10\cdot 0{,}992 = 2$$

Ví dụ 2.20: Tuổi thọ của một loại côn trùng nào đó là một biến ngẫu nhiên X (đơn vị là tháng) với hàm mật độ xác suất như sau

$$f_X(x) = \begin{cases} kx^2(4-x) & \text{khi } 0 \le x \le 4 \\ 0 & \text{khi } x \notin [0,4] \end{cases}$$

Tìm tuổi thọ trung bình của loài côn trùng trên

Giải:

$$\int_0^4 x^2(4-x)dx = \frac{64}{3} \Rightarrow k = \frac{3}{64}$$

Tuổi thọ trung bình

$$\mathrm{E}\,X = \frac{3}{64}\int_0^4 x^3(4-x)dx = \frac{3}{64}\left(x^4 - \frac{x^5}{5}\right)\Bigg|_0^4 = \frac{12}{5} \text{ (tháng)}$$

Tính chất kỳ vọng

1. $\mathrm{E}(C)=C$ với mọi hằng số C
2. $\mathrm{E}(CX)=C\,\mathrm{E}(X)$ với mọi hằng số C
3. $\mathrm{E}\left(X_1+\cdots+X_n\right)=\mathrm{E}\left(X_1\right)+\cdots+\mathrm{E}\left(X_n\right)$
4. Nếu $X_1, \ldots, X_n$ độc lập thì $\mathrm{E}\left(X_1\cdots X_n\right)=\mathrm{E}\left(X_1\right)\cdots\mathrm{E}\left(X_n\right)$
5. Cho hàm số $g(x)$, kỳ vọng của biến ngẫu nhiên $Y=g(X)$ được tính theo công thức

$$\mathrm{E}\,Y=\begin{cases}\displaystyle\sum_{x_i\in R_X} g(x_i)\,p_X(x_i) & \text{khi X roi rac có } p_X(x_i)=P\left\{X=x_i\right\}\\ \displaystyle\int_{-\infty}^{\infty} g(x)\,f_X(x)\,dx & khi\ X \text{ liên tuc có hàm MĐXS } f_X(x)\end{cases}$$

2.4.2 Phương sai

Phương sai (variance) hay độ lệch (deviation) bình phương trung bình của biến ngẫu nhiên X là đại lượng đo sự phân tán bình phương trung bình của X xung quanh giá trị trung bình $\mathrm{E}X$. Nói cách khác phương sai của X là kỳ vọng của $(X - \mathrm{E}X)^2$

Phương sai của X được ký hiệu $\mathrm{D}X$ hoặc $\mathrm{Var}X$

$$\mathrm{D}X = \mathrm{E}\left(X - \mathrm{E}X\right)^2$$

Độ lệch chuẩn của X : $\sigma_X = \sqrt{\mathrm{D}X}$

Khai triển vế phải công thức trên và áp dụng các tính chất của kỳ vọng ta có thể tính phương sai theo công thức sau

$$\mathrm{D}X = \mathrm{E}X^2 - \left(\mathrm{E}X\right)^2$$

$$\mathrm{E}\,X^2 = \sum_i x_i^2 p_X(x_i) \quad \text{nếu } X \text{ rời rạc}$$

$$\mathrm{E}\,X^2 = \int_{-\infty}^{\infty} x^2 f_X(x)dx \quad \text{nếu } X \text{ liên tục}$$

Ví dụ 2.24: Tính phương sai của biến ngẫu nhiên xét trong ví dụ 2.19.

$$\mathrm{E}\,X^2 = (-990)^2 \cdot 0,008 + 10^2 \cdot 0,992 = 7940$$

$$\Rightarrow \mathrm{D}X = \mathrm{E}X^2 - \left(\mathrm{E}X\right)^2 = 7940 - 4 = 7936$$

$$\Rightarrow \sigma_X = \sqrt{\mathrm{D}X} = \sqrt{7936} \approx 89,08$$

Điều này nói lên rằng mặc dù kinh doanh bảo hiểm có lãi nhưng rủi ro khá lớn

Tính chất của phương sai

1. $\mathrm{D}(a)=0$ với mọi hằng số a

2. $\mathrm{D}(aX+b)=a^2\mathrm{D}(X)$ với mọi hằng số a, b

3. Nếu $X_1, ..., X_n$ độc lập và có các phương sai hữu hạn thì

$$\mathrm{D}\left(a_1X_1+\cdots+a_nX_n\right)=a_1^2\mathrm{D}\left(X_1\right)+\cdots+a_n^2\mathrm{D}\left(X_n\right)$$

Nói riêng: Nếu X, Y độc lập và $\mathrm{D}X$, $\mathrm{D}Y$ hữu hạn thì

$$\mathrm{D}\left(X\pm Y\right)=\mathrm{D}X+\mathrm{D}Y$$

Phân vị

Phân vị mức α của biến ngẫu nhiên X, ký hiệu v_α , là giá trị phân chia miền giá trị của X thỏa mãn

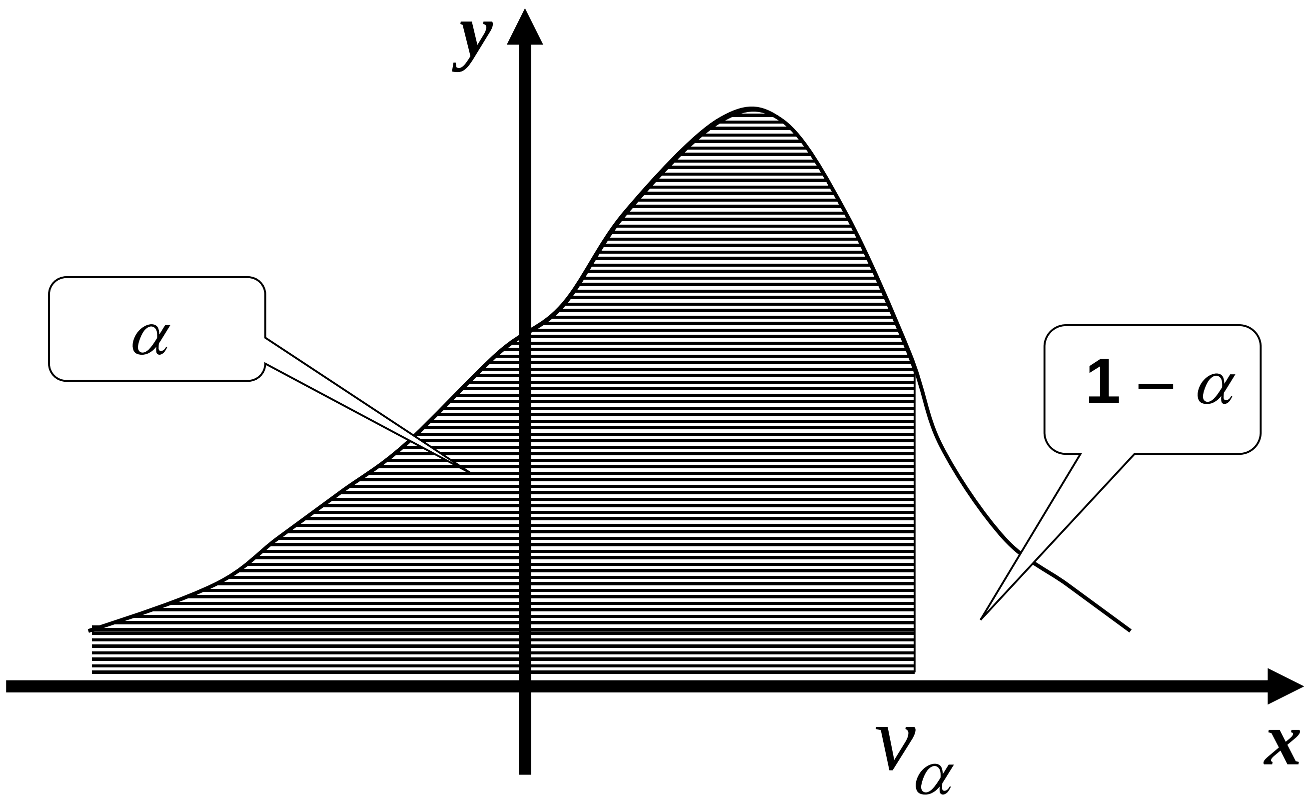
$$P\{X < v_\alpha\} \leq \alpha \leq P\{X \leq v_\alpha\}$$

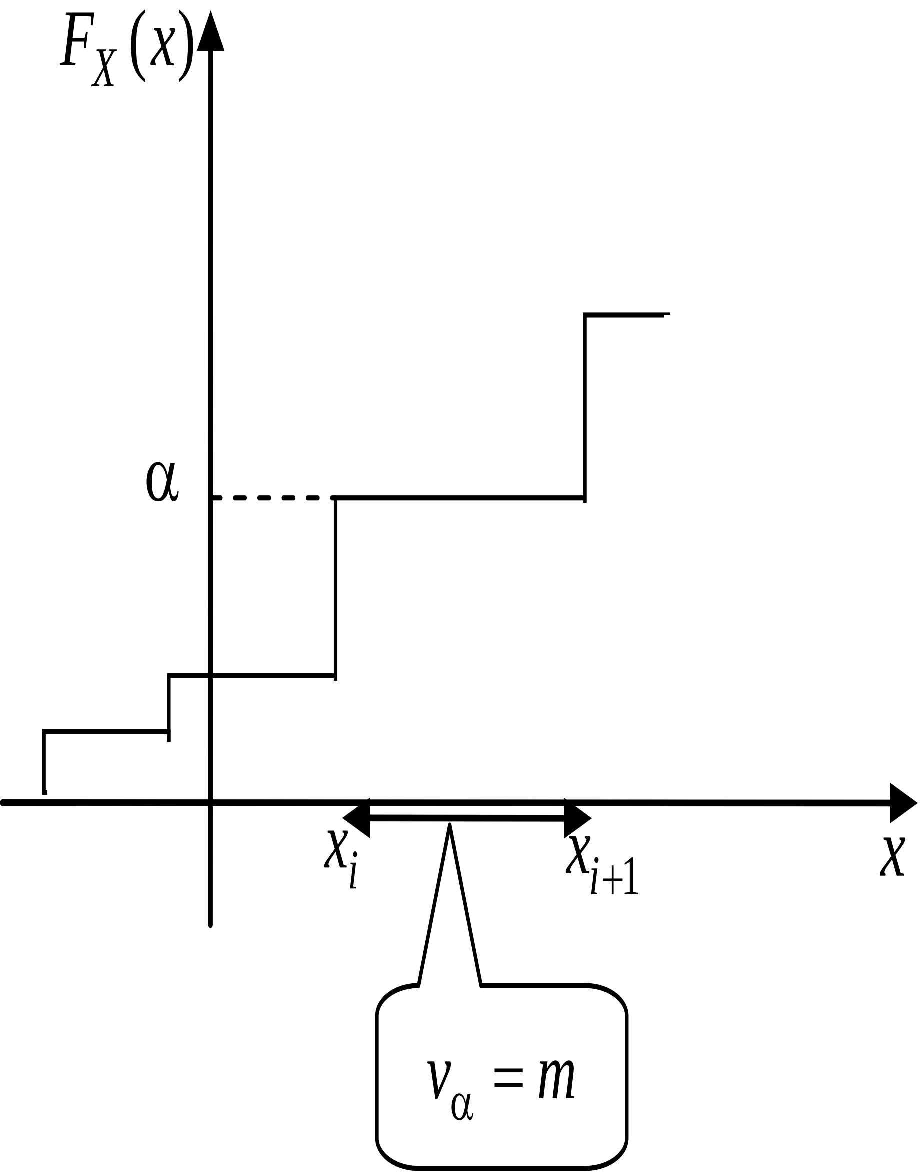
Nghĩa là $$F_X(v_\alpha -) \leq \alpha \leq F_X(v_\alpha)$$

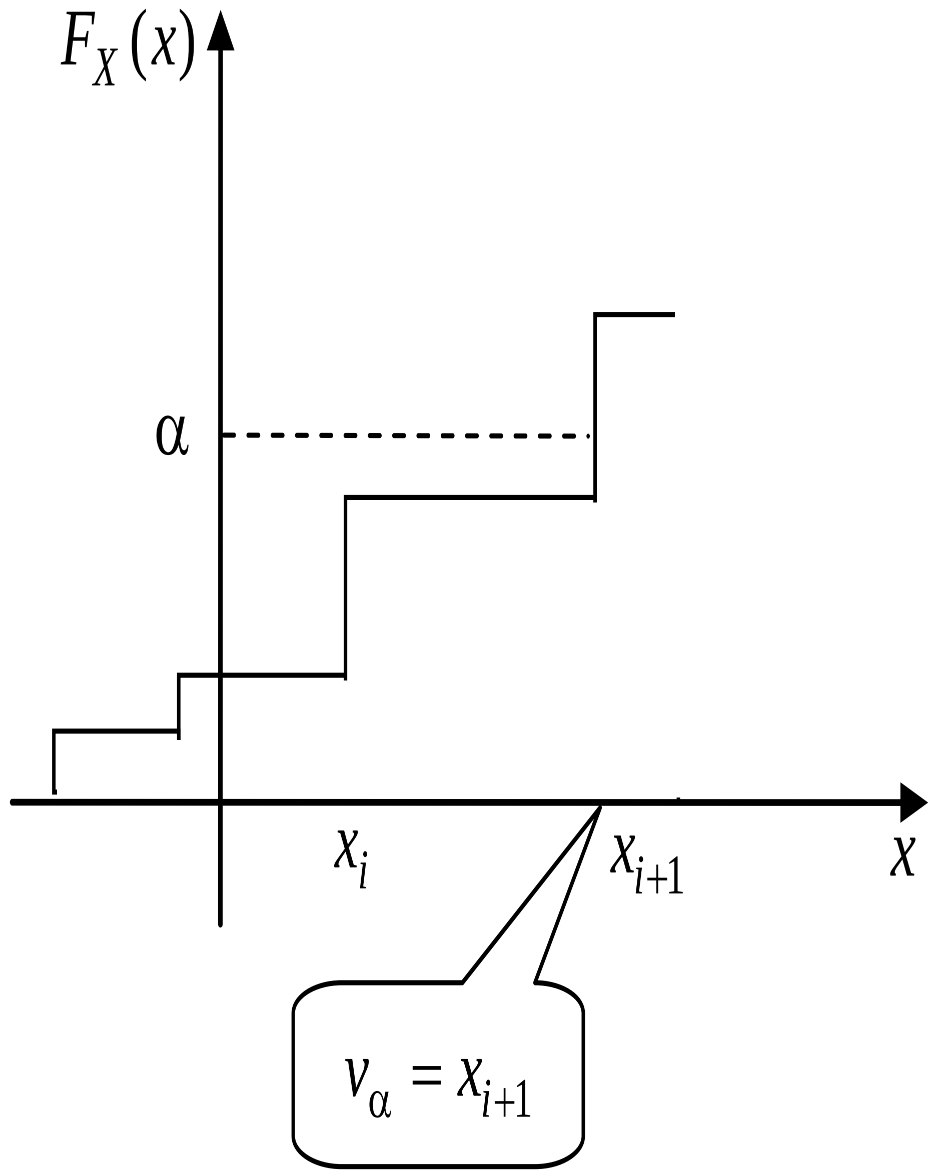
Trung vị

Phân vị mức 1/2 được gọi là median hay trung vị của X, ký hiệu $\mathrm{Med}\, X$

Như vậy trung vị là điểm phân chia phân bố xác suất thành hai phần bằng nhau


y
α
1 – α
v_α
x


$F_X(x)$
α
x_i
x_{i+1}
x
$v_\alpha = m$


$F_X(x)$
α
x_i
x_{i+1}
x
$v_\alpha = x_{i+1}$

Mốt

Mốt (Mode) của biến ngẫu nhiên X, ký hiệu $\text{Mod}\,X$, là giá trị mà biến ngẫu nhiên X nhận với xác suất lớn nhất

Một biến ngẫu nhiên có thể có nhiều Mốt

- Mốt của biến ngẫu nhiên rời rạc X

$$x_{i_0} = \text{Mod}\,X \quad \Leftrightarrow \quad p_X(x_{i_0}) = \max\left\{ p_X(x_1), p_X(x_2), \ldots \right\}$$

- Mốt của biến ngẫu nhiên liên tục X

$$c = \text{Mod}\,X \quad \Leftrightarrow \quad f_X(c) = \max\left\{ f_X(x),\ x \in 3 \right\}$$

Moment, hệ số bất đối xứng, hệ số nhọn

1. Moment cấp k $\quad m_k = \mathrm{E}X^k;\ k = 1, 2, \ldots$

2. Moment quy tâm cấp k $\quad \mu_k = \mathrm{E}\left(X - \mathrm{E}X\right)^k;\ k = 1, 2, \ldots$

3. Hệ số bất đối xứng $\quad \alpha_3 = \dfrac{\mu_3}{\sigma^3};\ \sigma = \sqrt{\mathrm{D}X}$

4. Hệ số nhọn $\quad \alpha_4 = \dfrac{\mu_4}{\sigma^4}$

α_3 đo mức độ bất đối xứng của luật phân bố

α_4 đặc trưng cho độ nhọn của đồ thị hàm mật độ xác suất

BÀI TẬP