

Xác suất thống kê ứng dụng

Nguyễn Ngọc Tứ

Bài giảng 5

Nội dung

- Biến ngẫu nhiên liên tục: hàm mật độ và hàm phân phối
- Các đặc trưng số của biến ngẫu nhiên liên tục
- Phân phối chuẩn

Bnn liên tục và hàm mật độ

Một biến ngẫu nhiên X gọi là liên tục nếu thỏa mãn:

- (1) các giá trị có thể có của X nằm trong một khoảng số thực hoặc là hợp của các khoảng rời nhau;
- (2) $P(X = c) = 0$ với bất kì số c là một giá trị của X .

Bnn liên tục và hàm mật độ

Một biến ngẫu nhiên X gọi là liên tục nếu thỏa mãn:

- (1) các giá trị có thể có của X nằm trong một khoảng số thực hoặc là hợp của các khoảng rời nhau;
- (2) $P(X = c) = 0$ với bất kì số c là một giá trị của X .

Hàm mật độ xác suất của X là hàm số $f(x)$ sao cho với bất kì hai giá trị a, b nào và $a \leq b$ thì

$$P(a \leq X \leq b) = \int_a^b f(x)dx.$$

Hơn thế nữa, $f(x)$ phải thỏa mãn hai điều kiện sau

1. $f(x) \geq 0$ với mọi x .
2. $\int_{-\infty}^{\infty} f(x)dx = 1$.

Hàm mật độ xác suất - Ví dụ

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

$$f(x) = \begin{cases} \frac{1}{360} & 0 \leq x < 360, \\ 0 & \text{nếu khác.} \end{cases}$$

Rõ ràng, $f(x) \geq 0$ và $\int_{-\infty}^{\infty} f(x)dx = \int_0^{360} f(x)dx = 1$. Chẳng hạn, ta tính

$$P(90 \leq X \leq 180) = \int_{90}^{180} \frac{1}{360} dx = 0.25$$

Hàm mật độ xác suất - Ví dụ

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

$$f(x) = \begin{cases} \frac{1}{360} & 0 \leq x < 360, \\ 0 & \text{nếu khác.} \end{cases}$$

Rõ ràng, $f(x) \geq 0$ và $\int_{-\infty}^{\infty} f(x)dx = \int_0^{360} f(x)dx = 1$. Chẳng hạn, ta tính

$$P(90 \leq X \leq 180) = \int_{90}^{180} \frac{1}{360} dx = 0.25$$

Định nghĩa. Một biến ngẫu nhiên liên tục X gọi là một **phân phối đều** trên một khoảng $[A, B]$ nếu hàm mật độ của X là

$$f(x; A, B) = \begin{cases} \frac{1}{B - A} & A \leq x \leq B \\ 0 & \text{nếu khác.} \end{cases}$$

Hàm phân phối tích lũy

Hàm phân phối tích lũy $F(x)$ của biến ngẫu nhiên liên tục X được xác định bởi:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(y)dy, \quad \forall x \in \mathbb{R}.$$

Cho X là một biến ngẫu nhiên liên tục có hàm mật độ $f(x)$ và hàm phân phối $F(x)$. Khi đó, với mọi số thực a , ta có

$$P(X > a) = 1 - F(a),$$

và cho hai số a và b với $a < b$, ta có

$$P(a \leq X \leq b) = F(b) - F(a).$$

Hàm phân phối tích lũy - Ví dụ

Cho X là biến ngẫu nhiên liên tục có hàm mật độ là

$$f(x) = \begin{cases} \frac{1}{8} + \frac{3}{8}x & 0 \leq x \leq 2 \\ 0 & \text{nếu khác} \end{cases}$$

Xét giá trị x bất kì trong đoạn $[0, 2]$, ta có

$$F(x) = \int_{-\infty}^x f(y) dy =$$

Hàm phân phối tích lũy - Ví dụ

Cho X là biến ngẫu nhiên liên tục có hàm mật độ là

$$f(x) = \begin{cases} \frac{1}{8} + \frac{3}{8}x & 0 \leq x \leq 2 \\ 0 & \text{nếu khác} \end{cases}$$

Xét giá trị x bất kì trong đoạn $[0, 2]$, ta có

$$F(x) = \int_{-\infty}^x f(y) dy = \int_0^x \left(\frac{1}{8} + \frac{3}{8}y \right) dy = \frac{x}{8} + \frac{3}{16}x^2$$

Hàm phân phối tích lũy - Ví dụ

Cho X là biến ngẫu nhiên liên tục có hàm mật độ là

$$f(x) = \begin{cases} \frac{1}{8} + \frac{3}{8}x & 0 \leq x \leq 2 \\ 0 & \text{nếu khác} \end{cases}$$

Xét giá trị x bất kì trong đoạn $[0, 2]$, ta có

$$F(x) = \int_{-\infty}^x f(y) dy = \int_0^x \left(\frac{1}{8} + \frac{3}{8}y \right) dy = \frac{x}{8} + \frac{3}{16}x^2$$

Khi đó,

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{8} + \frac{3}{16}x^2 & 0 \leq x \leq 2 \\ 1 & x > 2 \end{cases}$$

Hàm phân phối tích lũy - Ví dụ

Chẳng hạn, ta tính

$$\begin{aligned}P(1 \leq X \leq 1.5) &= F(1.5) - F(1) \\&= \left[\frac{1}{8}1.5 + \frac{3}{16}1.5^2 \right] - \left[\frac{1}{8}1 + \frac{3}{16}1^2 \right] \\&= \frac{19}{64} = 0.297\end{aligned}$$

và

Hàm phân phối tích lũy - Ví dụ

Chẳng hạn, ta tính

$$\begin{aligned}P(1 \leq X \leq 1.5) &= F(1.5) - F(1) \\&= \left[\frac{1}{8}1.5 + \frac{3}{16}1.5^2 \right] - \left[\frac{1}{8}1 + \frac{3}{16}1^2 \right] \\&= \frac{19}{64} = 0.297\end{aligned}$$

và

$$\begin{aligned}P(X > 1) &= 1 - P(X \leq 1) = 1 - F(1) = 1 - \left[\frac{1}{8}1 + \frac{3}{16}1^2 \right] \\&= \frac{11}{16} = 0.688\end{aligned}$$

Trung vị - Kỳ vọng - Phương sai

Trung vị của một phân phối liên tục là giá trị $\tilde{\mu}$ thỏa mãn $0.5 = F(\tilde{\mu})$.
Nghĩa là, trung vị là giá trị chia phân phối thành 2 phần bằng nhau.

Trung vị - Kỳ vọng - Phương sai

Trung vị của một phân phối liên tục là giá trị $\tilde{\mu}$ thỏa mãn $0.5 = F(\tilde{\mu})$.

Nghĩa là, trung vị là giá trị chia phân phối thành 2 phần bằng nhau.

Kỳ vọng hay **giá trị trung bình** của một biến ngẫu nhiên liên tục X với hàm mật độ $f(x)$ là

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

Trung vị - Kỳ vọng - Phương sai

Trung vị của một phân phối liên tục là giá trị $\tilde{\mu}$ thỏa mãn $0.5 = F(\tilde{\mu})$.

Nghĩa là, trung vị là giá trị chia phân phối thành 2 phần bằng nhau.

Kỳ vọng hay **giá trị trung bình** của một biến ngẫu nhiên liên tục X với hàm mật độ $f(x)$ là

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

Cho X là biến ngẫu nhiên liên tục và $h(X)$ là một hàm của X . Khi đó,

$$E[h(X)] = \mu_{h(X)} = \int_{-\infty}^{\infty} h(x) \cdot f(x) dx$$

Phương sai của X

$$\sigma_X^2 = \text{Var}(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 \cdot f(x) dx = E[(X - \mu_X)^2] = E(X^2) - \mu_X^2$$

Độ lệch chuẩn của X là $\sigma_X = \sqrt{\text{Var}(X)}$.

Ví dụ

Cho X là biến ngẫu nhiên liên tục có hàm mật độ

$$f(x) = \begin{cases} \frac{3}{2}(1 - x^2) & 0 \leq x \leq 1, \\ 0 & \text{nếu khác.} \end{cases}$$

Khi đó,

$$E(X) =$$

Ví dụ

Cho X là biến ngẫu nhiên liên tục có hàm mật độ

$$f(x) = \begin{cases} \frac{3}{2}(1 - x^2) & 0 \leq x \leq 1, \\ 0 & \text{nếu khác.} \end{cases}$$

Khi đó,

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^1 x \cdot \frac{3}{2}(1 - x^2) dx = \frac{3}{8}$$

$$E(X^2) =$$

Ví dụ

Cho X là biến ngẫu nhiên liên tục có hàm mật độ

$$f(x) = \begin{cases} \frac{3}{2}(1 - x^2) & 0 \leq x \leq 1, \\ 0 & \text{nếu khác.} \end{cases}$$

Khi đó,

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^1 x \cdot \frac{3}{2}(1 - x^2) dx = \frac{3}{8}$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx = \int_0^1 x^2 \cdot \frac{3}{2}(1 - x^2) dx = \frac{1}{5}$$

và

$$\text{Var}(X) =$$

Ví dụ

Cho X là biến ngẫu nhiên liên tục có hàm mật độ

$$f(x) = \begin{cases} \frac{3}{2}(1 - x^2) & 0 \leq x \leq 1, \\ 0 & \text{nếu khác.} \end{cases}$$

Khi đó,

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^1 x \cdot \frac{3}{2}(1 - x^2) dx = \frac{3}{8}$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx = \int_0^1 x^2 \cdot \frac{3}{2}(1 - x^2) dx = \frac{1}{5}$$

và

$$\text{Var}(X) = \frac{1}{5} - \left(\frac{3}{8}\right)^2 = \frac{19}{320} \quad \text{và} \quad \sigma_X = 0.244$$

Phân phối chuẩn

Định nghĩa. Một biến ngẫu nhiên liên tục X gọi là có phân phối chuẩn với hai tham số μ và σ , trong đó $-\infty < \mu < \infty$ và $0 < \sigma$, nếu hàm mật độ của X là

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}, \quad -\infty < x < \infty$$

Phân phối chuẩn

Định nghĩa. Một biến ngẫu nhiên liên tục X gọi là có phân phối chuẩn với hai tham số μ và σ , trong đó $-\infty < \mu < \infty$ và $0 < \sigma$, nếu hàm mật độ của X là

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}, \quad -\infty < x < \infty$$

Phân phối chuẩn với hai tham số $\mu = 0, \sigma = 1$ gọi là **phân phối chuẩn tắc**. Một biến ngẫu nhiên Z có phân phối chuẩn tắc với hàm mật độ xác suất là

$$f(z; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad -\infty < z < \infty$$

Hàm phân phối của Z là $P(Z \leq z) = \int_{-\infty}^z f(y; 0, 1) dy = \Phi(z)$.

Phân phối chuẩn - Ví dụ

Cho Z là biến ngẫu nhiên có phân phối chuẩn tắc. Tính các xác suất sau:

a) $P(Z \leq 1.25)$ b) $P(Z > 1.25)$

c) $P(Z \leq -1.25)$ d) $P(-0.38 \leq Z \leq 1.25)$

Phân phối chuẩn - Ví dụ

Cho Z là biến ngẫu nhiên có phân phối chuẩn tắc. Tính các xác suất sau:

- a) $P(Z \leq 1.25)$ b) $P(Z > 1.25)$
c) $P(Z \leq -1.25)$ d) $P(-0.38 \leq Z \leq 1.25)$

Ta có:

- a. $P(Z \leq 1.25) = \Phi(1.25)$, tra bảng ta được $\Phi(1.25) = 0.8944$. Do đó,
 $P(Z \leq 1.25) = 0.8944$

Phân phối chuẩn - Ví dụ

Cho Z là biến ngẫu nhiên có phân phối chuẩn tắc. Tính các xác suất sau:

- a) $P(Z \leq 1.25)$ b) $P(Z > 1.25)$
c) $P(Z \leq -1.25)$ d) $P(-0.38 \leq Z \leq 1.25)$

Ta có:

- a.** $P(Z \leq 1.25) = \Phi(1.25)$, tra bảng ta được $\Phi(1.25) = 0.8944$. Do đó,
 $P(Z \leq 1.25) = 0.8944$
b. $P(Z > 1.25) = 1 - P(Z \leq 1.25) = 0.1056$

Phân phối chuẩn - Ví dụ

Cho Z là biến ngẫu nhiên có phân phối chuẩn tắc. Tính các xác suất sau:

- a) $P(Z \leq 1.25)$ b) $P(Z > 1.25)$
c) $P(Z \leq -1.25)$ d) $P(-0.38 \leq Z \leq 1.25)$

Ta có:

- a. $P(Z \leq 1.25) = \Phi(1.25)$, tra bảng ta được $\Phi(1.25) = 0.8944$. Do đó,
 $P(Z \leq 1.25) = 0.8944$
b. $P(Z > 1.25) = 1 - P(Z \leq 1.25) = 0.1056$
c. $P(Z \leq -1.25) = \Phi(-1.25) = 0.1056$

Phân phối chuẩn - Ví dụ

Cho Z là biến ngẫu nhiên có phân phối chuẩn tắc. Tính các xác suất sau:

- a) $P(Z \leq 1.25)$ b) $P(Z > 1.25)$
c) $P(Z \leq -1.25)$ d) $P(-0.38 \leq Z \leq 1.25)$

Ta có:

- a. $P(Z \leq 1.25) = \Phi(1.25)$, tra bảng ta được $\Phi(1.25) = 0.8944$. Do đó,
 $P(Z \leq 1.25) = 0.8944$
b. $P(Z > 1.25) = 1 - P(Z \leq 1.25) = 0.1056$
c. $P(Z \leq -1.25) = \Phi(-1.25) = 0.1056$
d. $P(-0.38 \leq Z \leq 1.25) = P(Z \leq 1.25) - P(Z \leq -0.38) =$
 $\Phi(1.25) - \Phi(-0.38) = 0.8944 - 0.3520 = 0.5424$.

Phân phối chuẩn

Mệnh đề. Cho $X \sim N(\mu, \sigma^2)$. Khi đó,

$$Z = \frac{X - \mu}{\sigma}$$

có phân phối chuẩn tắc. Do đó,

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

$$P(X \leq a) = \Phi\left(\frac{a - \mu}{\sigma}\right) \quad P(X \geq b) = 1 - \Phi\left(\frac{b - \mu}{\sigma}\right)$$

Ví dụ. Cho X là biến ngẫu nhiên có phân phối chuẩn với giá trị trung bình 1.25 và độ lệch chuẩn 0.46. Tính xác suất:

a) $P(1 \leq X \leq 1.75)$ b) $P(X > 2)$

Phân phối chuẩn

a) Ta có

$$1 \leq X \leq 1.75$$

nếu và chỉ nếu

$$\frac{1 - 1.25}{0.46} \leq \frac{X - 1.25}{0.46} \leq \frac{1.75 - 1.25}{0.46}$$

Khi đó,

Phân phối chuẩn

a) Ta có

$$1 \leq X \leq 1.75$$

nếu và chỉ nếu

$$\frac{1 - 1.25}{0.46} \leq \frac{X - 1.25}{0.46} \leq \frac{1.75 - 1.25}{0.46}$$

Khi đó,

$$\begin{aligned} P(1 \leq X \leq 1.75) &= P\left(\frac{1 - 1.25}{0.46} \leq Z \leq \frac{1.75 - 1.25}{0.46}\right) \\ &= P(-0.54 \leq Z \leq 1.09) = \Phi(1.09) - \Phi(-0.54) \\ &= 0.8621 - 0.2946 = 0.5675 \end{aligned}$$

b)

$$P(X > 2)$$

Phân phối chuẩn

a) Ta có

$$1 \leq X \leq 1.75$$

nếu và chỉ nếu

$$\frac{1 - 1.25}{0.46} \leq \frac{X - 1.25}{0.46} \leq \frac{1.75 - 1.25}{0.46}$$

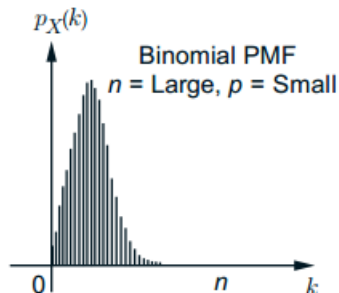
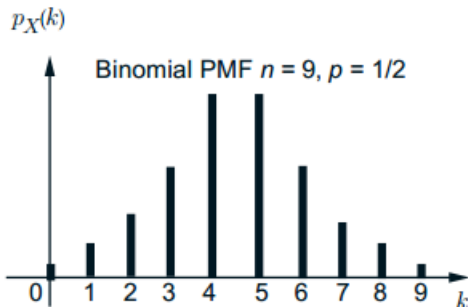
Khi đó,

$$\begin{aligned} P(1 \leq X \leq 1.75) &= P\left(\frac{1 - 1.25}{0.46} \leq Z \leq \frac{1.75 - 1.25}{0.46}\right) \\ &= P(-0.54 \leq Z \leq 1.09) = \Phi(1.09) - \Phi(-0.54) \\ &= 0.8621 - 0.2946 = 0.5675 \end{aligned}$$

b)

$$P(X > 2) = P\left(Z > \frac{2 - 1.25}{0.46}\right) = P(Z > 1.63) = 1 - \Phi(1.63) = 0.0516$$

Xấp xỉ phân phối nhị thức



Xấp xỉ phân phối nhị thức bởi Phân phối chuẩn

Cho $X \sim B(n, p)$. Khi đó nếu đồ thị của phân phối nhị thức không quá lệch (không quá mất đối xứng) thì ta có thể xấp xỉ X như một phân phối chuẩn với $\mu = np$ và $\sigma = \sqrt{np(1-p)}$. Đặc biệt, nếu x là một giá trị của biến ngẫu nhiên X thì

$$P(X \leq x) = B(x; n, p) \simeq \Phi\left(\frac{x - np}{\sqrt{np(1-p)}}\right)$$

Trong các bài tập cụ thể, tính xấp xỉ này được áp dụng nếu $np \geq 10$ và $n(1-p) \geq 10$ thỏa mãn. Bởi vì đây là điều kiện cần để phân phối nhị thức có tính chất đối xứng.

Phân phối chuẩn

Ví dụ. Giả sử một trường đại học có 25% sinh viên được nhận hỗ trợ tài chính. Gọi X là số lượng sinh viên được nhận hỗ trợ trong 50 sinh viên sao cho $p = 0.25$. Khi đó, ta có $\mu = 12.5$ và $\sigma = 3.06$.

Vì $np = 50 \cdot 0.25 = 12.5 \geq 10$ và $n(1 - p) = 37.5 \geq 10$ nên có thể xấp xỉ phân phối nhị thức. Xác suất mà nhiều nhất 10 sinh viên nhận hỗ trợ

$$P(X \leq 10) = B(10; 50, 0.25) \simeq \Phi\left(\frac{10 - 12.5}{3.06}\right) = \Phi(-0.81699) =$$

Xác suất có từ 5 đến 15 sinh viên được nhận hỗ trợ tài chính

$$\begin{aligned} P(5 \leq X \leq 15) &= B(15; 50, 0.25) - B(4; 50, 0.25) \\ &\simeq \Phi\left(\frac{15 - 12.5}{3.06}\right) - \Phi\left(\frac{4 - 12.5}{3.06}\right) = \end{aligned}$$

Phân phối mũ

Định nghĩa. Một biến ngẫu nhiên X gọi là có phân phối mũ với tham số $\lambda (\lambda > 0)$ nếu hàm mật độ xác suất của X là

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & \text{nếu khác.} \end{cases}$$

Hàm phân phối của X là

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

Giá trị trung bình và phương sai của X lần lượt là $\mu = \frac{1}{\lambda}; \sigma^2 = \frac{1}{\lambda^2}$

Phân phối mũ

Ví dụ. Cho X là biến ngẫu nhiên có phân phối mũ với giá trị trung bình là 6. Tính xác suất:

a) $P(X \leq 10)$

b) $P(5 \leq X \leq 10)$

Phân phối mũ

Ví dụ. Cho X là biến ngẫu nhiên có phân phối mũ với giá trị trung bình là 6. Tính xác suất:

a) $P(X \leq 10)$

b) $P(5 \leq X \leq 10)$

Ta có: $6 = EX = \frac{1}{\lambda}$, do đó $\lambda = 0.1667$.

a) $P(X \leq 10) = F(10; 0.1667) = 1 - e^{-(0.1667)(10)} = 1 - 0.189 = 0.811$

Phân phối mũ

Ví dụ. Cho X là biến ngẫu nhiên có phân phối mũ với giá trị trung bình là 6. Tính xác suất:

a) $P(X \leq 10)$

b) $P(5 \leq X \leq 10)$

Ta có: $6 = EX = \frac{1}{\lambda}$, do đó $\lambda = 0.1667$.

a) $P(X \leq 10) = F(10; 0.1667) = 1 - e^{-(0.1667)(10)} = 1 - 0.189 = 0.811$

b) $P(5 \leq X \leq 10) = F(10; 0.1667) - F(5; 0.1667)$
 $= (1 - e^{-1.667}) - (1 - e^{-8.335}) = 0.246$

Phân phối Gamma

Định nghĩa. Cho $\alpha > 0$, khi đó **hàm gamma** $\Gamma(\alpha)$ được xác định bởi

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

Một số tính chất quan trọng của hàm gamma:

1. Cho bất kì $\alpha > 1$, $\Gamma = (\alpha - 1) \cdot \Gamma(\alpha - 1)$.
2. Cho n là một số nguyên dương bất kì, ta có $\Gamma(n) = (n - 1)!$
3. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Đặt

$$f(x; \alpha) = \begin{cases} \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)} & x \geq 0, \\ 0 & \text{nếu khác.} \end{cases} \quad (1)$$

Khi đó, ta thấy $f(x; \alpha) \geq 0$ và $\int_0^{\infty} f(x; \alpha) dx = \Gamma(\alpha) / \Gamma(\alpha) = 1$, vì thế $f(x; \alpha)$ thỏa mãn hai tính chất cơ bản của hàm mật độ xác suất.

Phân phối Gamma

Định nghĩa. Một biến ngẫu nhiên liên tục X gọi là có **phân phối gamma** nếu hàm mật độ của X là

$$f(x; \alpha, \beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x} & x \geq 0, \\ 0 & \text{nếu khác.} \end{cases} \quad (2)$$

trong đó α và β là hai tham số dương.

Hàm gamma chuẩn tắc khi $\beta = 1$ và hàm mật độ của X lúc đó là (1).

Kì vọng và phương sai của X

$$E(X) = \mu = \alpha\beta \quad \text{Var}(X) = \sigma^2 = \alpha\beta^2$$

Hàm phân phối của biến ngẫu nhiên X có phân phối gamma chuẩn tắc là

$$F(x; \alpha) = \int_0^x \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)} dy \quad x > 0$$

còn được gọi là **hàm gamma không đầy đủ**.