Chương 2. Biến ngẫu nhiên

Nguyễn Minh Trí

Trường Đại học Công nghệ Thông tin

Mục tiêu

- Phân biệt được biến ngẫu nhiên rời rạc và liên tục.
- Tìm các hàm phân phối, hàm mật độ xác suất.
- Tính được kỳ vọng, phương sai, độ lệch chuẩn.
- Biết một số phân phối đặc biệt.
- Biết vận dụng phân phối đặc biệt để giải quyết các bài toán thực tế.

2.1 Biến ngẫu nhiên

Trong chương 1, xác suất của biến cố A là một số được gán với các biến cố A, trong đó A là một tập hợp các kết quả của một phép thử

$$A = \{\omega_1, \omega_2, \ldots, \omega_k\} \subset \Omega.$$

Biến cố có thể là tổng của hữu hạn, vô hạn đếm được các biến cố sơ cấp. Trong các trường hợp này, xác suất của biến cố là tổng của các xác suất của các biến cố sơ cấp chứa trong nó, tức là nếu biến cố $A = \{\omega_1, \omega_2, \ldots, \omega_k\}$ thì

$$P(A) = P(\omega_1) + P(\omega_2) + \ldots + P(\omega_k).$$

Trong chương 2, ta liên kết mỗi kết quả cả một phép thử với duy nhất 1 số nào đó. Trong hầu hết các bài toán thống kê, ta làm việc với các số mà chúng có sự liên kết với các kết quả của phép thử.

- Quan sát các sản phẩm: số sản phẩm tốt (lỗi).
- Quan sát các loại pin: tuổi thọ của pin.
- Mua một tờ vé số loại 10000 đồng: Nếu trúng số thì người mua được 100000 đồng. (Trúng số $\to 90000$; không trúng $\to -10000$)

Quy tắc liên kết mỗi kết quả của phép thử với một số $\longrightarrow$ **biến ngẫu nhiên**.

2.1 Biến ngẫu nhiên

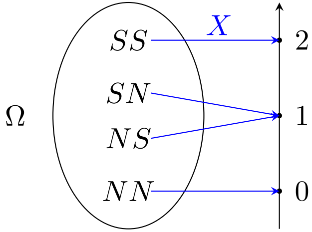
Định nghĩa 2.1 Cho một phép thử có không gian mẫu là Ω. Biến ngẫu nhiên X (random variable) là một quy tắc cho tương ứng mỗi biến cố sơ cấp của phép thử với duy nhất một số.

- Hay nói cách khác, mỗi biến ngẫu nhiên X là một ánh xạ

$$\begin{aligned} X : \Omega &\to \mathbb{R} \\ \omega &\mapsto X(\omega) \end{aligned}$$

- Kí hiệu: $X, Y, Z, \ldots$

Ví dụ 2.2 Xét một phép thử gieo 2 đồng xu. Không gian mẫu $\Omega = \{SS, NS, SN, NN\}$ và biến ngẫu nhiên X xác định bởi $X(\omega)$ là số mặt sấp (S) của biến cố ω.



$$X(\omega) = \begin{cases} 0, & \text{nếu } \omega = NN \\ 1, & \text{nếu } \omega \in \{SN, NS\} \\ 2, & \text{nếu } \omega = SS. \end{cases}$$

Ta thấy X nhận các giá trị $0, 1, 2$.

Ví dụ 2.3 Xét một phép thử bắn một viên đạn vào một tấm bia. Không gian mẫu $\Omega = \{$trúng, không trúng$\}$. Đặt biến ngẫu nhiên X như sau

$$X(\omega) = \begin{cases} 0, & \text{nếu } \omega = \text{ không trúng} \\ 1, & \text{nếu } \omega = \text{ trúng} \end{cases}$$

Khi đó X nhận các giá trị $0, 1$.

Ví dụ 2.4 Tung một đồng xu cho đến khi xuất hiện mặt ngửa (N). Không gian mẫu $\Omega = \{N, SN, SSN, SSSN, \ldots\}$. Đặt biến ngẫu nhiên Y là số lần tung đồng xu đến khi xuất hiện mặt ngửa. Khi đó Y nhận các giá trị $1, 2, 3, \ldots$

Ví dụ 2.5 Xét phép thử là chọn một số ngẫu nhiên trong đoạn $[1; 10]$. Cho quy tắc tương ứng Z xác định bởi $Z(\omega) = \omega - 1$. Khi đó Z là một biến ngẫu nhiên và có giá trị là các số thực thuộc $[0; 9]$.

Cho X là một biến ngẫu nhiên xác định trên không gian mẫu Ω.

- Ta viết $X = a$ thay cho tập hợp $\{\omega \in \Omega \mid X(\omega) = a\}$.
- Ta viết $X \leq a$ thay cho tập hợp $\{\omega \in \Omega \mid X(\omega) \leq a\}$.

Ví dụ 2.6 Xét phép thử là gieo một con xúc xắc. Đặt X là số chấm xuất hiện ở mặt trên.

Ký hiệu	Tập hợp
$X \leq 2$	$\{1, 2\}$
$X = 2$	$\{2\}$
$1 < X \leq 4$	$\{2, 3, 4\}$

Định nghĩa 2.7

1. Biến ngẫu nhiên được gọi là rời rạc (discrete) nếu tập giá trị của nó là một tập hữu hạn hoặc *vô hạn đếm được* các phần tử.
2. Biến ngẫu nhiên được gọi là liên tục (continuous) nếu tập giá trị của nó lấp kín một khoảng trên trục số.
3. Một biến ngẫu nhiên không rời rạc và cũng không liên tục, nó có đặc điểm của cả hai thì được gọi là hỗn hợp (mixed).

(Một tập hợp là vô hạn đếm được nếu ta có thể gán mỗi phần tử của nó với một số tự nhiên.)

Ví dụ 2.8

1. Đặt biến ngẫu nhiên X là số chấm xuất hiện khi gieo một con xúc xắc. Khi đó X nhận các giá trị là 1, 2, 3, 4, 5, 6. Do đó, X là một biến ngẫu nhiên rời rạc. – (hữu hạn)
2. Đặt Y là số lần tung một đồng xu cho đến khi xuất hiện mặt ngửa. Khi đó Y nhận các giá trị $1, 2, 3, \ldots$. Do đó Y là một biến ngẫu nhiên rời rạc. – (vô hạn đếm được)
3. Đặt Z là số đo nhiệt độ trong một ngày tại Thủ Đức. Khi đó Z là nhận các giá trị trong đoạn $[-50; 50]$. Do đó Z là một biến ngẫu nhiên liên tục.
4. Đặt T là thời gian chờ của các phương tiện giao thông tại một cột đèn giao thông. Nếu có đèn xanh thì thời gian chờ bằng 0 (rời rạc). Nếu đèn màu đỏ thì các xe phải dừng chờ trong một khoảng thời gian (liên tục) cho đến khi có đèn xanh. Do đó T là một biến ngẫu nhiên hỗn hợp.

2.2 Phân phối xác suất

2.2 Phân phối xác suất

Định nghĩa 2.9 Cho X là một biến ngẫu nhiên rời rạc có miền giá trị là $\{x_1, x_2, \ldots\}$. Phân phối xác suất (probability distribution) của X là danh sách các giá trị của X và các xác suất của chúng

X	x_1	x_2	$\ldots$
$P(X = x_i)$	p_1	p_2	$\ldots$

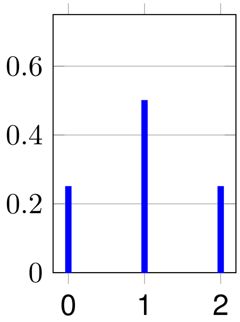
trong đó $p_i = P(X = x_i)$.
Kí hiệu: $p(x) = P(X = x)$.
Chú ý:

- $0 \leq p(x) \leq 1$
- $\sum_{i \geq 1} p(x_i) = 1.$

Ví dụ 2.10 Xét phép thử là tung 2 đồng xu cân đối đồng chất. Không gian mẫu là $\Omega = \{SS, SN, NS, NN\}$. Đặt X là số mặt ngửa (N). Tìm phân phối xác suất của X.

Giải. Ta thấy X nhận các giá trị $0, 1, 2$. Phân phối xác suất của X là

X	0	1	2
$P(X = x_i)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

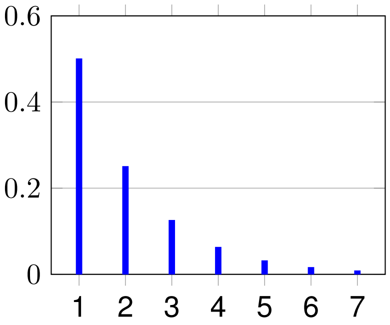


Biểu đồ cột của phân phối xác suất

Ví dụ 2.11 Xét phép thử là tung 1 đồng xu cân đối đồng chất đến khi xuất hiện mặt sấp (S). Không gian mẫu là $\Omega = \{S, NS, NNS, NNNS, \dots\}$. Đặt X là số lần tung đến khi xuất hiện mặt sấp. Tìm phân phối xác suất của X.

Giải. Ta thấy X nhận các giá trị $1, 2, 3 \dots$. Phân phối xác suất của X là

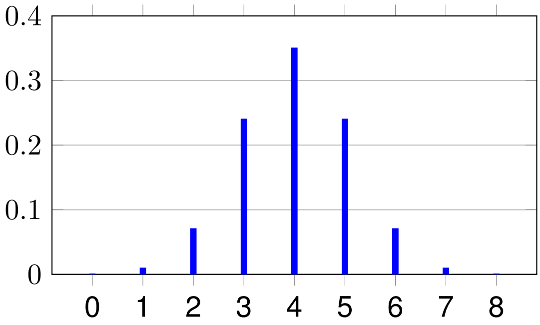
X	1	2	3	$\cdots$	n	$\cdots$
$P(X = x_i)$	$\frac{1}{2}$	$\left(\frac{1}{2}\right)^2$	$\left(\frac{1}{2}\right)^3$	$\cdots$	$\left(\frac{1}{2}\right)^n$	$\cdots$



Ví dụ 2.12 Một hộp có 10 bi đen và 10 bi trắng. Lấy ngẫu nhiên 8 bi. Đặt X là số bi đen. Tìm phân phối xác suất của X.

Giải. Ta thấy X nhận các giá trị $0, 1, 2, \ldots, 8$. Phân phối xác suất của X là

$$P(X = x) = \frac{C_{10}^{x} C_{10}^{8-x}}{C_{20}^{8}}, x = 0, 1, 2, \ldots, 8.$$



Ví dụ 2.13 Cho X là một biến ngẫu nhiên rời rạc có phân phối xác suất như sau

X	0,2	0,4	0,5	0,8	1
$P(X=x)$	0,1	0,2	0,2	0,3	0,2

a. Tính $P(X \leq 0,5)$

b. Tính $P(0,25 < X < 0,75)$

c. Tính $P(X = 0,2 | X \leq 0,5)$

Giải.

a. $P(X \leq 0,5) = P(X = 0,2) + P(X = 0,4) + P(X = 0,5) = 0,1 + 0,2 + 0,2 = 0,5$

b. $P(0,25 < X \leq 0,75) = P(X = 0,4) + P(X = 0,5) = 0,2 + 0,2 = 0,4$

c. $P(X = 0,2|X \leq 0,5) = \dfrac{P(X = 0,2 \cap X \leq 0,5)}{P(X \leq 0,5)} = \dfrac{P(X = 0,2)}{P(X \leq 0,5)} = \dfrac{0,1}{0,5} = 0,2$

Định nghĩa 2.14 Hàm phân phối (distribution function) của biến ngẫu nhiên X được ký hiệu và xác định như sau

$$F(x) = P(X \leq x).$$

- Hàm phân phối còn được gọi hàm phân phối xác suất tích lũy (cumulative distribution function, viết tắt là CDF).
- Nếu X là biến ngẫu nhiên rời rạc thì

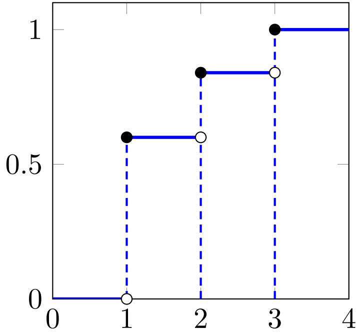
$$F(x) = \sum_{x_i \leq x} P(X = x_i).$$

Ví dụ 2.15 Cho X là một biến ngẫu nhiên có phân phối xác suất như sau

X	1	2	3
$P(X=x_i)$	$0,6$	$0,24$	$0,16$

Hàm phân phối của X

$$F(x) = \begin{cases} 0, & x < 1 \\ 0,6, & 1 \le x < 2 \\ 0,84, & 2 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$



Đồ thị hàm phân phối của X

Định lý 2.16 Cho $F(x)$ là hàm phân phối của biến ngẫu nhiên X. Khi đó

1. $\lim_{x \to -\infty} F(x) = 0$
2. $\lim_{x \to +\infty} F(x) = 1$
3. Hàm số không giảm, tức là nếu $a \leq b$ thì $F(a) \leq F(b)$.
4. Hàm số liên tục phải, tức là $\lim_{x \to a^+} F(x) = F(a)$ với mọi a.

Chú ý. Nếu một hàm số $F : \mathbb{R} \to [0, 1]$ thỏa mãn 4 tính chất của Định lý 2.16 thì người ta chứng minh được rằng F là hàm phân phối của một biến ngẫu nhiên nào đó. Hơn nữa, X là biến ngẫu nhiên liên tục khi và chỉ khi hàm phân phối của nó là một hàm số liên tục.

Ví dụ 2.17 Cho D là một hình trong mặt phẳng Oxy. Nếu A, B là các tập con của D có diện tích bằng nhau và chọn ngẫu nhiên 1 điểm ω trong D thì ta nói

$$P(\omega \in A) = P(\omega \in B).$$

Ta định nghĩa

$$P(\omega \in A) = \frac{\text{diện tích hình } A}{\text{diện tích hình } D}.$$

Cho D là một hình tròn có bán kính R. Chọn ngẫu nhiên một điểm ω trong D. Đặt X là khoảng cách từ tâm của D đến ω. Nếu $0 \leq x \leq R$ thì biến cố

$$\{\omega \mid X(\omega) \leq x\}$$

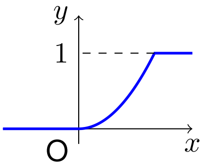
là hình tròn có tâm của D và có bán kính x.

Do đó

$$F(x) = P(X \le x) = \frac{\pi x^2}{\pi R^2} = \frac{x^2}{R^2}$$

Hơn nữa, nếu $x < 0$ thì $P(X \le x) = 0$ hoặc nếu $x > R$ thì $P(X \le x) = 1$.
Hàm phân phối của X xác định như sau

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{x^2}{R^2}, & 0 \le x \le R, \\ 1, & x > R. \end{cases}$$



Định lý 2.18 Cho $F(x)$ là hàm phân phối của biến ngẫu nhiên X. Khi đó

1. $P(X < a) = \lim\limits_{x \to a^-} F(x)$.
2. $P(X = a) = F(a) - \lim\limits_{x \to a^-} F(x)$.
3. Nếu $a < b$ thì $P(a < X \leq b) = F(b) - F(a)$.
4. $P(X > a) = 1 - F(a)$.

Nhận xét: Nếu X là một biến ngẫu nhiên liên tục thì

- $P(X = a) = 0$ với mọi a thuộc miền giá trị của X.
- $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$.

Ví dụ 2.19 Cho X là một biến ngẫu nhiên có hàm phân phối

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

Tính $P(X = 0,5)$ và $P(0,2 < X < 4)$.

Giải.

$$P(X = 0,5) = F(0,5) - \lim_{x \to 0,5^-} F(x) = 0,5^2 - 0,5^2 = 0.$$

$$\begin{aligned} P(0,2 < X < 4) &= P(0,2 < X \leq 4) - P(X = 4) \\ &= (F(4) - F(0,2)) - \left(F(4) - \lim_{x \to 4^-} F(x)\right) \\ &= (1 - 0,2^2) - (1 - 1) = 0,96 \end{aligned}$$

Bài toán. Cho X (số cm) là chiều cao của một người được chọn ngẫu nhiên. Khi đó X là một biến ngẫu nhiên liên tục và nhận giá trị trong đoạn $[10; 260]$. Xác suất $X = 150,123$ cm bằng bao nhiêu?

Khi đó

$$P(X = 150,123) = 0.$$

Nhận xét: Cho X là một biến ngẫu nhiên liên tục có miền giá trị là $[\alpha, \beta]$ và $F(x)$ là hàm phân phối của X. Khi đó

$$P(X = c) = F(c) - \lim_{x \to c^-} F(x) = F(c) - F(c) = 0$$

với mọi $c \in [\alpha, \beta]$.

Câu hỏi. Với $\alpha \leq a < b \leq \beta$, khi đó $P(a \leq X \leq b) = ?$

- Chia đoạn $[a, b]$ thành m phần bằng nhau Δx bởi các điểm

$$a = x_0, x_1, x_2, \ldots, x_m = b.$$

- Khi đó

$$\begin{aligned} P(a \le X \le b) &= \sum_{i=1}^{m} P(X \in [x_{i-1}; x_i]) \\ &= \sum_{i=1}^{m} P(X \in (x_{i-1}; x_i]) = \sum_{i=1}^{m} \left(F(x_i) - F(x_{i-1})\right). \end{aligned}$$

- Nếu $\Delta x \to 0$ và F là một hàm số khả vi trên $[a, b]$ thì

$$F(x_i) - F(x_{i-1}) \approx F'(x_{i-1})\Delta x$$

và do đó

$$P(a \le X \le b) = \int_a^b F'(x)\mathrm{d}x.$$

Đặt $f(x) = F'(x)$ và gọi là hàm mật độ xác suất (probability density function) của biến ngẫu nhiên liên tục X.

Định nghĩa 2.20 Hàm mật độ xác suất (probability density function) của một biến ngẫu nhiên liên tục X là một hàm số $f(x)$ thỏa mãn

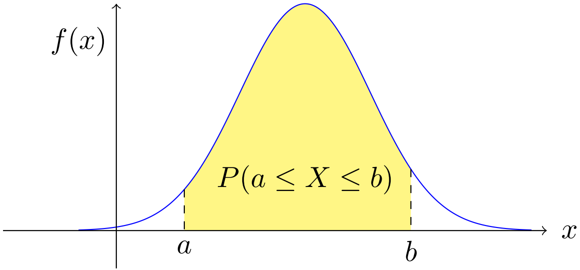
1. $f(x) \geq 0,$
2. $\displaystyle\int_{-\infty}^{+\infty} f(x)dx = 1.$

Tính chất. Cho X là một biến ngẫu nhiên liên tục có hàm mật độ xác suất $f(x)$. Khi đó

1. $\displaystyle P(a \leq X \leq b) = \int_a^b f(x)\mathrm{d}x$
2. $\displaystyle P(X \geq a) = \int_a^{+\infty} f(x)dx.$
3. $\displaystyle F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\mathrm{d}t.$
4. $F'(x) = f(x).$

Nhận xét.

- Đồ thị của hàm mật độ xác suất $f(x)$ không nằm dưới trục hoành.
- Diện tích giới hạn bởi đồ thị của $f(x)$ và trục hoành bằng 1.
- Xác suất của của biến ngẫu nhiên X thuộc đoạn $[a, b]$ bằng diện tích hình giới hạn bởi của đồ thị $f(x)$, trục hoành và hai đường thẳng $x = a; x = b$.



Ví dụ 2.21 Thời gian sử dụng của một thiết bị điện tử (tính bằng năm) là một biến ngẫu nhiên liên tục X có hàm mật độ xác suất

$$f(x) = \begin{cases} \frac{k}{x^3}, & x \geq 1 \\ 0, & x < 1. \end{cases}$$

a. Tìm k.

b. Tìm hàm phân phối của X.

c. Tính xác suất thiết bị đó sử dụng được hơn 5 năm.

Giải. a. Vì $f(x) \geq 0$ với mọi $x \in \mathbb{R}$ nên $\dfrac{k}{x^3} \geq 0$ với mọi $x \geq 1$.Suy ra $k \geq 0$.
Hơn nữa,

$$1 = \int_{-\infty}^{+\infty} f(x)\mathrm{d}x = \int_{1}^{+\infty} \frac{k}{x^3}\mathrm{d}x = \frac{k}{2}.$$ Do đó $k = 2$.

Hàm mật độ của X là

$$f(x) = \begin{cases} \frac{2}{x^3}, & x \geq 1 \\ 0, & x < 1. \end{cases}$$

b. Hàm phân phối xác suất $F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$.

- Với $x < 1$, ta có $f(x) = 0$. Do đó $F(x) = 0$.
- Với $x \geq 1$, ta có

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t = \int_{1}^{x} \frac{2}{t^3}\mathrm{d}t = 1 - \frac{1}{x^2}$$

Như vậy

$$F(x) = \begin{cases} 1 - \dfrac{1}{x^2}, & x \geq 1 \\ 0, & x < 1. \end{cases}$$

$$f(x) = \begin{cases} \frac{2}{x^3}, & x \geq 1 \\ 0, & x < 1. \end{cases}$$

c. Xác suất thiết bị đó sử dụng được nhiều hơn 5 năm

$$P(X > 5) = \int_5^{+\infty} f(x)\mathrm{d}x = \int_5^{+\infty} \frac{2}{x^3}\mathrm{d}x = \frac{1}{25} = 0,04.$$

Cách khác

$$P(X > 5) = 1 - P(X \leq 5) = 1 - F(5) = 1 - \left(1 - \frac{1}{5^2}\right) = 0,04.$$

Bài tập

Bài 2.1 Một túi có 10 bi đỏ và 6 bi đen. Chọn ngẫu nhiên ra 3 bi.

a. Gọi X là số bi đỏ. Tìm bảng phân phối xác suất của X.

b. Giả sử chọn được mỗi bi đỏ được 5$ và mỗi bi đen được 8$. Gọi Y là số tiền tổng cộng tương ứng với 3 bi được chọn ra. Tìm bảng phân phối xác suất của Y.

Bài 2.2 Trong một lô hàng có 2 loại sản phẩm: Loại thứ nhất có 10 hộp, mỗi hộp có 7 sản phẩm hiệu A và 3 sản phẩm hiệu B; loại thứ hai có 15 hộp, mỗi hộp có 8 sản phẩm hiệu A và 4 sản phẩm hiệu B. Lấy ngẫu nhiên một hộp, rồi từ hộp đó lấy ra 2 sản phẩm. Gọi X là số sản phẩm hiệu A. Tìm bảng phân phối và hàm phân phối xác suất của X.

Bài 2.3 Trong 30 sinh viên của một lớp nọ có 8 sinh viên thấp hơn 170 cm. Chọn ngẫu nhiên 4 sinh viên của lớp. Đặt X là số sinh viên có thấp hơn 170cm.

- **a.** Lập bảng phân phối xác suất của X.
- **b.** Tìm hàm phân phối của X.
- **c.** Tính xác suất gặp ít nhất 2 sinh viên thấp hơn 170cm.
- **d.** Tính xác suất gặp từ 1 đến 3 sinh viên thấp hơn 170cm.

Bài 2.4 Đặt X là thời gian (tháng) sử dụng của một loại thiết bị điện tử. Giả sử biến ngẫu nhiên X có hàm phân phối xác suất

$$F(x) = 1 - e^{-3x} \text{ với } x > 0.$$

- **a.** Chọn ngẫu nhiên một thiết bị. Tính xác suất thiết bị đó sử dụng được ít nhất 12 tháng.
- **b.** Tìm hàm mật độ xác suất của X.

Bài 2.5 Thời gian phục vụ mỗi khách hàng tại một cửa hàng là một biến ngẫu nhiên liên tục X (phút) có hàm mật độ xác suất như sau

$$f(x) = \begin{cases} k.e^{-4x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

a. Tìm k.

b. Tìm hàm phân phối của X.

c. Tính xác suất thời gian phục vụ một khách hàng trong khoảng thời gian từ 1 đến 2 phút.

Bài 2.6 Cho một biến ngẫu nhiên X là thời gian sử dụng của một loại pin (tính bằng năm). Giả sử hàm mật độ xác suất của X được cho bởi

$$f(x) = \begin{cases} \frac{1}{2}e^{-x/2} & x > 0 \\ 0, & x \leq 0. \end{cases}$$

a. Tính xác suất loại pin này sử dụng được ít nhất 2 năm hoặc nhiều hơn 4 năm.

b. Giả sử có một viên pin loại này đã được sử dụng 2 năm. Tính xác suất pin này được sử dụng ít nhất 3 năm.

Chương 2. Biến ngẫu nhiên

Nguyễn Minh Trí

Trường Đại học Công nghệ Thông tin

2.3 Kỳ vọng, phương sai

2.3 Kỳ vọng, phương sai

Định nghĩa 2.22 **Kỳ vọng** (expectation) của biến ngẫu nhiên X, ký hiệu $E(X)$ hay μ, được xác định như sau:

- Nếu X là biến ngẫu nhiên rời rạc và nhận các giá trị $x_1, x_2, \ldots$ thì

$$E(X) = \sum_{i \geq 1} x_i P(X = x_i)$$

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

$$E(X) = \int_{-\infty}^{+\infty} x f(x) \mathrm{d}x$$

Kỳ vọng còn được gọi là giá trị trung bình (mean), giá mong đợi (expected value).

Ví dụ 2.23 Trong hộp có 99 quả bóng đen và 1 quả bóng đỏ. Người chơi bỏ 10 000 đồng cho mỗi lần lấy một quả bóng trong hộp. Nếu lấy được bóng đỏ, người chơi được 100 000 đồng. Nếu lấy được bóng đen thì người chơi xem như mất 10 000 đồng. Trung bình một lần lấy bóng, người chơi được/mất bao nhiêu tiền?

Giải. Đặt X là số tiền người chơi đạt được sau khi lấy bóng.

- Nếu lấy được bóng đỏ thì $X = 90000$ (tiền thắng giải trừ tiền mua)
- Nếu lấy được bóng đen thì $X = -10000$ (mất 10 000 đồng cho 1 lần lấy bóng)
- Xác suất lấy được bóng đỏ: $\frac{1}{100}$, xác suất không lấy được bóng đỏ: $\frac{99}{100}$

Phân phối xác suất của X là

X	90000	-10000
$P(X = x)$	$0,01$	$0,99$

$$E(X) = 90000.0,01 + (-10000).0,99 = -90.$$

Như vậy, trung bình mỗi lần lấy bóng, người chơi mất 90 đồng.

Ví dụ 2.24 Trong một thành phố nọ, tỉ lệ những người có bằng đại học là 20%. Gặp ngẫu nhiên 3 người. Đặt X là số người có bằng đại học trong 3 người này. Tính kỳ vọng của X.

Giải. Các giá trị mà X có thể nhận là $0, 1, 2, 3$. Xác suất khi $X = x$ là

$$P(X = x) = C_3^x.0,2^x.(1-0,2)^{3-x}.$$

Kỳ vọng của X

$$\begin{aligned} E(X) &= 0.P(X=0) + 1.P(X=1) + 2.P(X=2) + 3.P(X=3) \\ &= 0.0,512 + 1.0,128 + 2.0,032 + 3.0,008 = 0,6. \end{aligned}$$

Ví dụ 2.25 Giả sử biến ngẫu nhiên X có hàm mật độ xác suất như sau

$$f(x) = \begin{cases} 3x^2, & 0 \le x \le 1 \\ 0, & x \notin [0;1] \end{cases}$$

Tính $E(X)$.

Giải. Ta có

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} x f(x) \mathrm{d}x \\ &= \int_0^1 x 3x^2 \mathrm{d}x \\ &= \frac{3}{4}. \end{aligned}$$

Định lý 2.26

1. Nếu X là một biến ngẫu nhiên rời rạc và $g(x)$ là một hàm số thì

$$E(g(X)) = \sum_{x} g(x)P(X = x)$$

2. Nếu X là một biến ngẫu nhiên liên tục có hàm mật độ xác suất $f(x)$ và $g(x)$ là một hàm số thì

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

3. $E(aX + b) = aE(X) + b$ với $a, b \in \mathbb{R}$.

Ví dụ 2.27 Một trung tâm bảo hành laptop thấy rằng số laptop hiệu T cần sửa trong năm tối đa là 3 máy. Cho X là số laptop T cần sửa trong một năm. Dựa trên các số liệu đã có, bảng phân phối xác suất của X như sau

X	0	1	2	3
$P(X = x_i)$	0,2	0,3	0,4	0,1

a. Trong một năm, trung bình có bao nhiêu laptop T được sửa tại trung tâm?

b. Một cửa hàng Z có hợp đồng với trung tâm bảo hành đối với việc sửa laptop T như sau: Phí dịch vụ hàng năm là 200$ và 50$ cho mỗi laptop được sửa. Tìm trung bình số tiền mà cửa hàng Z trả cho trung tâm bảo hành mỗi năm.

X	0	1	2	3
$P(X = x_i)$	0,2	0,3	0,4	0,1

Giải. a. Trung bình số laptop T được sửa tại trung tâm bảo hành trong một năm

$$E(X) = 0.0,2 + 1.0,3 + 2.0,4 + 3.0,1 = 1,4.$$

b. (Phí dịch vụ hàng năm là 200$ và 50$ cho mỗi laptop được sửa.)
Số tiền phải trả hàng năm là

$$g(X) = 200 + 50X.$$

Số tiền trung bình hàng năm cửa hàng Z trả cho trung tâm bảo hành

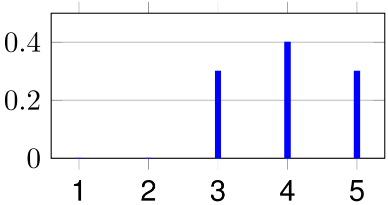
$$E(g(X)) = E(200 + 50X) = 200 + 50.E(X) = 200 + 50.1,4 = 270(\$).$$

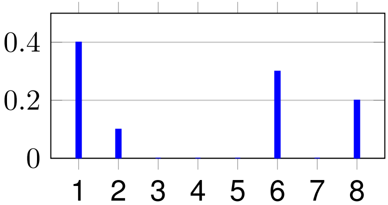
Quan sát 2 hình dưới đây biểu diễn phân phối xác suất của 2 biến ngẫu nhiên.

X	3	4	5
$P(X = x_i)$	0,3	0,4	0,3

Y	1	2	6	8
$P(Y = y_i)$	0,4	0,1	0,3	0,2

$$E(X) = 3.0,3 + 4.0,4 + 5.0,4 = 4; \quad E(Y) = 1.0,4 + 2.0,1 + 6.0,3 + 8.0,2 = 4$$





Cả hai phân phối xác xuất có cùng kỳ vọng $E(X)$ nhưng có độ phân tán khác nhau.

$\rightarrow$ Cần số đo để đánh giá mức độ phân tán.

Định nghĩa 2.28

- **Phương sai** (variance) của X, kí hiệu $\mathrm{Var}(X)$ hoặc σ^2,
  1. Nếu X là biến ngẫu nhiên rời rạc có miền giá trị $\{x_1, x_2, \ldots\}$ thì

$$\mathrm{Var}(X) = \sigma^2 = \sum_{i=1}^{\infty} (x_i - E(X))^2 P(X = x_i)$$

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

$$\mathrm{Var}(X) = \sigma^2 = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x) \mathrm{d}x$$

- **Độ lệch chuẩn** (standard deviation) của X

$$\mathrm{Std}(X) = \sqrt{\mathrm{Var}(X)} = \sigma.$$

Ví dụ 2.29 (Ví dụ 2.24) Trong một thành phố nọ, tỉ lệ những người có bằng đại học là 20%. Gặp ngẫu nhiên 3 người. Đặt X là số người có bằng đại học trong 3 người này. Tính phương sai của X.

Giải.

- Các giá trị mà X có thể nhận là $0, 1, 2, 3$. Xác suất khi $X = x$ là

$$P(X = x) = C_3^x.0,2^x.(1 - 0,2)^{3-x}.$$

- Kỳ vọng của X là $E(X) = 0,6$.
- Phương sai của X

$$\begin{aligned}\mathrm{Var}(X) &= (0 - 0,6)^2.0,512 + (1 - 0,6)^2.0,128 + (2 - 0,6)^2.0,032 + (3 - 0,6)^2.0,008 \\ &= 0,48.\end{aligned}$$

Ví dụ 2.30 (Ví dụ 2.25) Giả sử biến ngẫu nhiên X có hàm mật độ xác suất như sau

$$f(x) = \begin{cases} 3x^2, & 0 \leq x \leq 1 \\ 0, & x \notin [0;1] \end{cases}$$

Tính $\mathrm{Var}(X)$.

Giải. Ta đã có $E(X) = \dfrac{3}{4}$. Phương sai của X là

$$\begin{aligned} \mathrm{Var}(X) &= \int_{-\infty}^{+\infty} (x - E(X))^2 f(x) \mathrm{d}x \\ &= \int_0^1 \left(x - \frac{3}{4}\right)^2 3x^2 \mathrm{d}x \\ &= 0,0375. \end{aligned}$$

Quan sát 2 phân phối xác suất của 2 biến ngẫu nhiên.

X	3	4	5
$P(X = x_i)$	0,3	0,4	0,3

Y	1	2	6	8
$P(Y = y_i)$	0,4	0,1	0,3	0,2

$$E(X) = 3.0,3 + 4.0,4 + 5.0,4 = 4; \quad E(Y) = 1.0,4 + 2.0,1 + 6.0,3 + 8.0,2 = 4$$

Ta có

$$\mathrm{Var}(X) = (3-4)^2.0,3 + (4-4)^2.0,4 + (5-4)^2.0,3 = 0,6$$
$$\mathrm{Var}(Y) = (1-4)^2.0,4 + (2-4)^2.0,1 + (6-4)^2.0,3 + (8-4)^2.0,2 = 8,4$$

Ta thấy $\mathrm{Var}(Y) > \mathrm{Var}(X)$ nên các giá trị của Y phân tán xa giá trị kỳ vọng hơn các giá trị của X.

Định lý 2.31 Cho X là một biến ngẫu nhiên. Khi đó

$$\mathrm{Var}(X) = E(X^2) - E(X)^2.$$

Ví dụ 2.32 (Ví dụ 2.25) Giả sử biến ngẫu nhiên X có hàm mật độ xác suất như sau

$$f(x) = \begin{cases} 3x^2, & 0 \le x \le 1 \\ 0, & x \notin [0;1] \end{cases}$$

Tính $\mathrm{Var}(X)$.

Giải. Ta có

$$\begin{aligned} \mathrm{Var}(X) &= E(X^2) - E(X)^2 \\ &= \int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x - E(X)^2 \\ &= \int_0^1 x^2 3x^2 \mathrm{d}x - \left(\frac{3}{4}\right)^2 = \frac{3}{5} - \frac{9}{16} = \frac{3}{80} = 0,0375. \end{aligned}$$

Định lý 2.33

1. Nếu X là một biến ngẫu nhiên và $g(x)$ là một hàm số thì

$$\operatorname{Var}(g(X)) = E(g^2(X)) - [E(g(X))]^2$$

2. $\operatorname{Var}(aX + b) = a^2 \operatorname{Var}(X)$ với $a, b \in \mathbb{R}$

Ví dụ 2.34 Cho biến ngẫu nhiên X có hàm mật độ xác suất như sau

$$f(x) = \begin{cases} 3x^2, & 0 \le x \le 1 \\ 0, & x \notin [0;1] \end{cases}$$

Cho biến ngẫu nhiên $Y = 3X - 2$. Tính $E(Y)$ và $\mathrm{Var}(Y)$.

Giải.

$$E(Y) = E(3X - 2) = 3E(X) - 2 = 3.\frac{3}{4} - 2 = 0.25$$

$$\mathrm{Var}(Y) = \mathrm{Var}(3X - 2) = 3^2\mathrm{Var}(X) = 9.0,0375 = 0,3375.$$

Định lý 2.35 (Định lý Chebyshev[1]) Cho X là một biến ngẫu nhiên có kỳ vọng μ và phương sai σ^2. Khi đó với mọi $\varepsilon > 0$, ta có

$$P(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

Định lý Chebyshev cho thấy

$$P(\mu-\varepsilon < X < \mu+\varepsilon) \geq 1-\frac{\sigma^2}{\varepsilon^2}.$$

[1] Pafnuty Lvovich Chebyshev (1821 - 1894) là nhà toán học người Nga.

Ví dụ 2.36 Số khách hàng đến một trung tâm bán xe trong một ngày là một biến ngẫu nhiên có trung bình là 18 và độ lệch chuẩn là 2,5. Xác suất có nhiều hơn 8 nhưng ít hơn 28 người đến trung tâm trong một ngày ít nhất là bao nhiêu?

Giải. Đặt X là số khách đến trung tâm trong một ngày. Tính $P(8 < X < 28)$. Theo Định lý Chebyshev,

$$P(\mu - \varepsilon < X < \mu + \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2}.$$

Đặt

$$\begin{cases} \mu - \varepsilon = 8 \\ \mu + \varepsilon = 28 \end{cases} \Rightarrow \varepsilon = 10$$

Do đó

$$P(8 < X < 28) \geq 1 - \frac{2,5^2}{10^2} = \frac{15}{16} = 0,9375.$$

Xác suất số khách đến trung tâm nhiều hơn 8 và ít hơn 28 người ít nhất là 93,75%.

Ví dụ 2.37 Chứng minh rằng trong 40 000 tung đồng xu có ít nhất 99% tỉ lệ mặt ngửa lớn hơn 0,475 và nhỏ hơn 0,525.

Giải. Đặt X là số lần xuất hiện mặt ngửa trong 40 000 lần tung đồng xu.

- Xác suất xuất hiện mặt ngửa trong mỗi lần tung đồng xu là 0,5.
- $P(X=k) = C^k_{40000}(0,5)^k(0,5)^{40000-k} = C^k_{40000}(0,5)^{40000}$
- $E(X) = \mu = 20000$ và $\mathrm{Var}(X) = \sigma^2 = 10000.$
- Vì $0,475 < \dfrac{X}{40000} < 0,525$ nên $19000 < X < 21000$
- Đặt

$$\begin{cases} \mu - \varepsilon = 19000 \\ \mu + \varepsilon = 21000 \end{cases} \Rightarrow \varepsilon = 1000$$

- Theo Định lý Chebyshev

$$P(19000 < X < 21000) \geq 1 - \frac{\sigma^2}{\varepsilon^2} = 1 - \frac{10000}{1000^2} = 0,99.$$

2.4 Hàm sinh moment

2.4 Hàm sinh moment

Định nghĩa 2.38 Cho X là một biến ngẫu nhiên và i là một số nguyên dương. Moment thứ i của biến ngẫu nhiên X là $E(X^i)$.

Nhận xét:

- Moment thứ nhất (kỳ vọng, trung bình): $E(X)$
- Moment thứ hai: $E(X^2) = \mathrm{Var}(X) - E(X)^2$
- Moment thứ ba (skewness): $E(X^3)$ cho biết về tính bất đối xứng của phân phối.
- Moment thứ tư (kurtosis): $E(X^4)$ cho biết đuôi của phân phối của phân phối "nặng" như thế nào.

Định nghĩa 2.39 **Hàm sinh moment** (moment generating function) của một biến ngẫu nhiên X là một hàm số $M : \mathbb{R} \to [0, +\infty)$ được xác định như sau:

$$M(t) = E(e^{tX}).$$

Ta nói hàm sinh moment của X tồn tại nếu có một hằng số dương a sao cho $M(t) < \infty$ với mọi $t \in [-a, a]$.

- X là biến ngẫu nhiên rời rạc nhận các giá trị $x_1, x_2, \ldots$

$$M(t) = \sum_{i \geq 1} e^{tx_i} P(X = x_i).$$

- X là biến ngẫu nhiên liên tục có hàm mật độ $f(x)$

$$M(t) = \int_{-\infty}^{+\infty} e^{tx} f(x) \mathrm{d}x.$$

Ví dụ 2.40 Cho X là một biến ngẫu nhiên rời rạc có phân phối xác suất như sau

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

trong đó $n > 0, 0 < p < 1$ là các hằng số. Tìm hàm sinh moment của X.
Giải. Vì X là biến ngẫu nhiên rời rạc nên

$$\begin{aligned} M(t) = E(e^{tX}) &= \sum_{k=0}^{n} e^{tk} P(X=k) \\ &= \sum_{k=0}^{n} e^{tk} C_n^k p^k (1-p)^{n-k} \\ &= \sum_{k=0}^{n} C_n^k \left(e^t p\right)^k (1-p)^{n-k} = (pe^t + 1 - p)^n. \end{aligned}$$

Ví dụ 2.41 Cho Y là biến ngẫu nhiên liên tục có hàm mật độ xác suất như sau

$$f(y) = \lambda e^{-\lambda y}$$

trong đó $y > 0$ và λ là hằng số. Tìm hàm sinh moment $M(t)$ của Y.

Giải.

$$\begin{aligned} M(t) = E(e^{tY}) &= \int_0^{+\infty} \lambda e^{ty} e^{-\lambda y} \mathrm{d}y \\ &= -\frac{\lambda}{\lambda - t} e^{-(\lambda - t)y} \Big|_0^{+\infty}, \text{ với } t < \lambda \\ &= \frac{\lambda}{\lambda - t}, \text{ với } t < \lambda \end{aligned}$$

Định lý 2.42 Cho X là một biến ngẫu nhiên. Khi đó

1. $M(t) = \sum_{k=0}^{+\infty} E(X^k) \frac{t^k}{k!}$,
2. $E(X^k) = M^{(k)}(0)$.

Ví dụ 2.43 Cho biến ngẫu nhiên X có hàm sinh moment

$$M(t) = (pe^t + 1 - p)^n.$$

Tính kỳ vọng và phương sai của X.

Giải.

- Kỳ vọng: Ta có $M'(t) = n(pe^t + 1 - p)^{n-1}(pe^t)$. Do đó

$$E(X) = M'(0) = np.$$

- Phương sai: Ta có

$$M''(t) = n(n-1)(pe^t + 1 - p)^{n-2}(pe^t)^2 + n(pe^t + 1 - p)^{n-1}(pe^t).$$

  Do đó

$$\mathrm{Var}(X) = E(X^2) - E(X)^2 = M''(0) - (M'(0))^2 = np(1-p).$$

Định lý 2.44 Hàm sinh moment xác định duy nhất một phân phối xác suất.

Ví dụ 2.45 Cho biến ngẫu nhiên X có hàm sinh moment

$$M(t) = \left(\frac{3}{4} + \frac{1}{4}e^t\right)^{20}$$

với mọi $t \in \mathbb{R}$. Tìm phân phối xác suất của X.

Giải. Ta có

$$M(t) = \sum_{k=0}^{20} C_{20}^k \left(\frac{3}{4}\right)^{20-k} \left(\frac{1}{4}\right)^k (e^t)^k.$$

Do đó

$$P(X = k) = C_{20}^k \left(\frac{3}{4}\right)^{20-k} \left(\frac{1}{4}\right)^k$$

với $k = 0, 1, \ldots, 20.$

Định lý 2.46 Cho $X_1, X_2, \ldots, X_n$ là các biến ngẫu nhiên độc lập và Y là biến ngẫu nhiên xác định bởi

$$Y = X_1 + X_2 + \cdots + X_n.$$

Khi đó

$$M_Y(t) = M_{X_1}(t) M_{X_2}(t) \ldots M_{X_n}(t).$$

Bài tập

Bài 2.7 Đặt X là số mặt ngửa khi gieo 4 đồng xu giống nhau. Tính kỳ vọng và phương sai của X. (Đáp số: $E(X) = 2, \mathrm{Var}(X) = 1$)

Bài 2.8 Cho một biến ngẫu nhiên X là một biến ngẫu nhiên liên tục có hàm mật độ như sau

$$f(x) = \begin{cases} |x| & -1 \le x \le 1 \\ 0, & \text{các trường hợp khác} \end{cases}$$

và biến ngẫu nhiên $Y = X^3 - X^2 - X + 1$. Tính kỳ vọng của các biến ngẫu nhiên X và Y. (Đáp số: $E(X) = 0, E(Y) = \dfrac{1}{2}$)

Bài tập

Bài 2.9 Đặt X là số mặt ngửa khi gieo 2 đồng xu giống nhau. Tìm hàm sinh moment của X. (Đáp số: $M(t) = \frac{1}{4} + \frac{1}{2}e^t + \frac{1}{4}e^{2t}$)

Bài 2.10 Cho X là một biến ngẫu nhiên liên tục có hàm mật độ xác suất như sau

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0, & \text{các trường hợp khác} \end{cases}$$

Tính moment thứ n của X. (Đáp số: $E(X^n) = \frac{2}{n+2}$)

2.5 Một số phân phối rời rạc

2.5.1 Phân phối nhị thức

Định nghĩa 2.47 Một dãy các phép thử được gọi là các phép thử Bernoulli (Bernoulli trails) nếu nó thỏa các điều sau:

1. Mỗi phép thử chỉ có hai kết quả (thành công hoặc thất bại).
2. Xác suất thành công trong mỗi phép thử là bằng nhau.
3. Các phép thử độc lập với nhau.

Ví dụ 2.48

- Bạn trả lời 5 câu hỏi dạng Đúng - Sai. Khi đó bạn thực hiện 5 phép thử Bernuolli với xác suất thành công là 0,5.
- Trả lời 10 câu hỏi trắc nghiệm 4 lựa chọn, có duy nhất một lựa chọn đúng là 10 phép thử Bernoulli với xác suất thành công là 0,25.

2.5.1 Phân phối nhị thức

Định nghĩa 2.49 Phân phối nhị thức (binomial distribution) là phân phối xác suất của một biến ngẫu nhiên mà nó là số lần thành công trong một dãy phép thử Bernoulli.

- Đặt X là số lần thành công trong n phép thử Bernoulli với xác suất thành công là p.
- X có phân phối nhị thức, ký hiệu $X \sim B(n;p)$.
- Phân phối xác suất của X là

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

- $E(X) = np$ và $\mathrm{Var}(X) = np(1-p)$

Ví dụ 2.50 Theo một cuộc khảo sát tại một quốc gia nọ có 60% các hộ gia đình có sử dụng thiết bị kết nối không dây. Chọn ngẫu nhiên 5 hộ gia đình. Tính xác suất

a. có đúng 4 hộ có sử dụng thiết bị không dây.

b. có ít nhất 4 hộ có sử dụng thiết bị không dây.

Giải. Đặt X là số hộ gia đình có sử dụng thiết bị không dây trong 5 hộ gia đình. Theo đề bài $X \sim B(5; 0,6)$ với $n = 5$ và $p = 0,6$.

a. Xác suất có đúng 4 hộ có sử dụng thiết bị không dây

$$P(X = 4) = C_5^4.0,6^4.(1 - 0,6)^1 = 0,259$$

b. Xác suất có ít nhất 4 hộ có sử dụng thiết bị không dây

$$\begin{aligned} P(X \geq 4) &= P(X = 4) + P(X = 5) \\ &= C_5^4.0,6^4.0,4^1 + C_5^5.0,6^5.0,4^0 = 0,259 + 0,078 = 0,337 \end{aligned}$$

2.5.2 Phân phối siêu bội

Định nghĩa 2.51 Xét một tập có N phần tử trong đó có k phần tử có tính chất $\mathcal{P}$. Chọn n phần tử khác nhau, không phân biệt thứ tự từ N phần tử. Gọi X là số phần tử có tính chất $\mathcal{P}$ trong n phần tử đã chọn. Khi đó X là một biến ngẫu nhiên và ta nói X có **phân phối siêu bội** (hypergeometric distribution) với ba tham số N, k, n.

- Ký hiệu $X \sim H(N, k, n)$
- Phân phối xác suất

$$P(X = x) = \frac{C_k^x C_{N-k}^{n-x}}{C_N^n}$$

- $E(X) = n\dfrac{k}{N}$
- $\mathrm{Var}(X) = \dfrac{N-n}{N-1} n \dfrac{k}{N} . \left(1 - \dfrac{k}{N}\right)$

Ví dụ 2.52 Trong 20 laptop trưng bày tại một cửa hàng có 5 laptop hiệu A. Chọn ngẫu nhiên 6 laptop trong cửa hàng đó.

a. Tính xác suất chọn được không quá 1 laptop hiệu A.

b. Trung bình số laptop hiệu A được chọn là bao nhiêu?

Giải. Gọi X là số laptop hiệu A trong số 6 laptop được chọn. Theo đề bài $X \sim H(20;5;6)$ với $N = 20; k = 5$ và $n = 6$.
a. Xác suất chọn được không quá 1 laptop hiệu A

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$
$$= \frac{C_5^0 C_{15}^6}{C_{20}^6} + \frac{C_5^1 C_{15}^5}{C_{20}^6} = 0,129 + 0,387 = 0,516$$

Ví dụ 2.53 Trong 20 laptop trưng bày tại một cửa hàng có 5 laptop hiệu A. Chọn ngẫu nhiên 6 laptop trong cửa hàng đó.

a. Tính xác suất chọn được không quá 1 laptop hiệu A.

b. Trung bình số laptop hiệu A được chọn là bao nhiêu?

Giải. a. Gọi X là số laptop hiệu A trong số 6 laptop được chọn. Theo đề bài $X \sim H(20; 5; 6)$ với $N = 20; M = 5$ và $k = 6$.
b. Trung bình số laptop hiệu A được chọn

$$E(X) = k\frac{M}{N} = 6\frac{5}{20} = 1,5$$

2.5.3 Phân phối Poisson

Cho X là số lần xuất hiện của một biến cố nào đó trong một khoảng thời gian (hoặc không gian). Khi đó ta nói X có phân phối Poisson.

Ví dụ 2.54

- Gọi X là số lỗi chính tả trên một trang in (không gian là trang in).
- Gọi X là số ô tô đi qua ngã tư Thủ Đức trong một phút (thời gian là một phút).
- Gọi X là số email tôi nhận được trong 1 ngày (thời gian là một ngày).

Định nghĩa 2.55 Một biến ngẫu nhiên rời rạc X được gọi là có **phân phối Poisson** (Poisson distribution) nếu nó có phân phối xác suất là

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

trong đó λ là trung bình số lần xuất hiện biến cố mà ta quan tâm.

- Ký hiệu $X \sim P(\lambda)$
- $E(X) = \lambda$
- $\mathrm{Var}(X) = \lambda$
- Giá trị của hàm phân phối xác suất tích lũy $F(X) = P(X \leq x)$ của biến ngẫu nhiên có phân phối Poisson có thể tìm trong bảng A3.

Ví dụ 2.56 Người ta thấy rằng trong một quyển sách nọ có trung bình 3 lỗi ở mỗi trang. Chọn ngẫu nhiên một trang, tính xác suất

a. trang đó không có lỗi.

b. trang đó có nhiều nhất 4 lỗi.

Giải. Đặt X là số lỗi của một trang. Trung bình số lỗi của 1 trang là $\lambda = 3$ và $X \sim P(3)$.

a. Xác suất trang đó không có lỗi

$$P(X=0) = \frac{3^0.e^{-3}}{0!} = 0,05$$

b. Xác suất trang đó nhiều nhất 4 lỗi

$$\begin{aligned} P(X \leq 4) &= P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4) \\ &= \frac{3^0.e^{-3}}{0!} + \frac{3^1.e^{-3}}{1!} + \frac{3^2.e^{-3}}{2!} + \frac{3^3.e^{-3}}{3!} + \frac{3^4.e^{-3}}{4!} = 0,815 \end{aligned}$$

2.6 Một số phân phối liên tục

2.6.1 Phân phối đều

- Chọn ngẫu nhiên một số X trong đoạn $[a, b]$. Tất cả các số thuộc đoạn $[a, b]$ đều có khả năng được chọn như nhau. Khi đó ta nói X có *phân phối đều* trên $[a, b]$.
- Bạn đến một trạm xe buýt để chờ một chuyến buýt chạy 10 phút một lần, nhưng không biết xe buýt trước rời trạm lúc nào. Gọi X là thời gian đợi xe buýt. Khi đó X có phân phối đều trên $[0; 10]$.

2.6.1 Phân phối đều

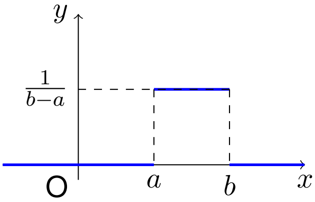
Định nghĩa 2.57 Một biến ngẫu nhiên liên tục được gọi là có phân phối đều (uniform distribution) nếu hàm mật độ của nó có dạng

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{các trường hợp khác} \end{cases}$$

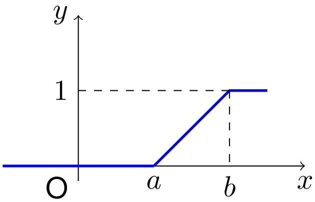
- Ký hiệu $X \sim U(a;b)$
- Hàm phân phối của X

$$F(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

Đồ thị của hàm mật độ



Đồ thị của hàm phân phối



Cho biến ngẫu nhiên X có phân phối đều $X \sim U(a; b)$. Khi đó

1. Kỳ vọng: $E(X) = \dfrac{a+b}{2}$
2. Phương sai: $\mathrm{Var}(X) = \dfrac{(b-a)^2}{12}$
3. Hàm sinh moment (xác định trên (a, b))

$$M(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}$$

Giới thiệu phần mềm R

- Năm 1996, hai nhà thống kê học Ross Ihaka và Robert Gentleman[2] phát hoạ một ngôn ngữ mới cho phân tích thống kê mà họ đặt tên là R.
- R là một phần mềm sử dụng cho phân tích thống kê và vẽ biểu đồ.
- Phần mềm miễn phí, mã nguồn mở.
- Tải về một phiên bản của R và các gói thư viện từ địa chỉ của trang web http://cran.r-project.org.
- Tài liệu tham khảo: http://cran.r-project.org. (Contributed)

[2] Ihaka R. and Gentleman R.: R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3): 299-314, 1996.

Sử dụng R

Cho $X \sim U(a, b)$.

1. Tính $P(X \leq x)$ (giá trị hàm phân phối $F(x)$)

   `punif(`x`,`a`,`b`)`

2. Tính $P(c \leq X \leq d)$

   `punif(`d`,`a`,`b`)-punif(`c`,`a`,`b`)`

3. Vẽ đồ thị hàm mật độ của phân phối đều

   `curve (dunif(x,`a`,`b`), from = `w`, to = `z`)`

4. Tô màu phần diện tích (tiếp theo phần 3)

   `x<- seq(`u, v, c`)` # Giá trị $u \leq x \leq v$ và c là khoảng cách giữa các cột

   `lines (x,dunif(x,`a`,`b`), type='h', col="màu")`

5. Tìm t thỏa mãn $P(X \leq t) = d$.

   `qunif(`d`,`a`,`b`)`

Ví dụ 2.58 Một công ty giao hàng chia các gói hàng của họ thành các gói loại từ 100 gam đến 120 gam (được phân phối đồng đều, nghĩa là tất cả trọng lượng trong mỗi loại đó đều bằng nhau).

- **a.** Tìm hàm mật độ của biến ngẫu nhiên biểu diễn trọng lượng từ 100 gam đến 120 gam.
- **b.** Tìm xác suất các gói có trọng lượng tối đa 115 gam.
- **c.** Tìm xác suất các gói có trọng lượng từ 112 gam đến 116 gam.
- **d.** Tìm trung bình và phương sai.

Giải. **a.** Đặt X là trọng lượng gói hàng từ 100 gam đến 120 gam. Khi đó hàm mật độ của X

$$f(x) = \begin{cases} \dfrac{1}{20}, & 100 \leq x \leq 120 \\ 0, & \text{các trường hợp khác} \end{cases}$$

b. $P(X \leq 115) = \int_{100}^{115} \frac{1}{20} \mathrm{d}x = \frac{x}{20}\Big|_{100}^{115} = \frac{15}{20} = 0,75.$

c. $P(112 \leq X \leq 116) = \int_{112}^{116} \frac{1}{20} \mathrm{d}x = \frac{x}{20}\Big|_{112}^{116} = \frac{4}{20} = 0,2.$

d. $E(X) = \frac{120+100}{2} = 110$ và $\sigma^2 = \frac{1}{12}(120-100)^2 = 33,33.$

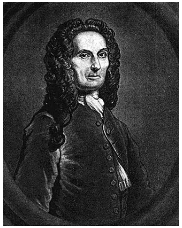
Dùng R:

b. `punif(115, 100, 120)`

c. `punif(116, 100, 120)-punif(112, 100, 120)`

2.6.2 Phân phối chuẩn

Phân phối chuẩn (Normal distribution) được giới thiệu bởi một người Anh gốc Pháp tên là Abraham de Moivre (1733). Sau đó, Gauss, một nhà toán học người Đức, đã dùng phân phối chuẩn để nghiên cứu các dữ liệu về thiên văn học (1809) và do vậy nó cũng được gọi là phân phối Gauss.

Johann Carl Friedrich Gauss (1777-1855) và Abraham de Moivre (1667-1754).
Nguồn ảnh: Wikipeadia.org

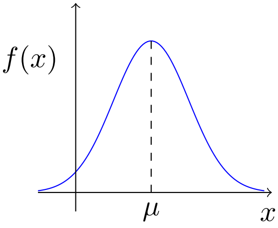
Định nghĩa 2.59 Một biến ngẫu nhiên liên tục X được gọi là có **phân phối chuẩn** (normal distribution) nếu hàm mật độ xác suất của nó có dạng

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

trong đó $\pi = 3,14159\ldots$ và $e = 2,71828\ldots$

- Ký hiệu $X \sim N(\mu; \sigma^2)$.

Hình dạng và vị trí của đường cong phân phối chuẩn (đồ thị của hàm mật độ xác suất) phụ thuộc vào hai tham số μ và σ.



Một số tính chất của phân phối chuẩn

1. $E(X) = \mu$
2. $\mathrm{Var}(X) = \sigma^2$
3. Độ lệch chuẩn $\mathrm{Std}(X) = \sigma$
4. Kỳ vọng nằm ở trung điểm của phân phối.
5. Đồ thị của hàm mật độ xác suất có hình chuông, đối xứng qua đường thẳng đứng đi qua kỳ vọng.
6. Đường cong liên tục và không chạm vào trục x.
7. $P(\mu - \sigma \leq X \leq \mu + \sigma) = 0,68$
8. $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0,95$
9. $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0,997.$

Định nghĩa 2.60 Phân phối chuẩn với $\mu = 0$ và $\sigma = 1$ được gọi là **phân phối chuẩn chuẩn tắc** (standard normal distribution).

- Các phân phối chuẩn $X \sim N(\mu, \sigma^2)$ có thể được chuyển thành phân phối chuẩn chuẩn tắc bằng cách đổi biến

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

- Biến ngẫu nhiên Z có hàm mật độ xác suất

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

và hàm phân phối xác suất

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} \mathrm{d}t$$

Bài toán. Giả sử $X \sim N(\mu; \sigma^2)$. Tính $P(X \leq a)$.

1. Đổi biến $Z = \dfrac{X - \mu}{\sigma} \Rightarrow Z \sim N(0; 1)$.
2. $P(X \leq a) = P(Z \leq \dfrac{a-\mu}{\sigma}) = \Phi(\dfrac{a-\mu}{\sigma})$
3. Tìm $\Phi(z) = P(Z \leq z)$ từ Bảng A4
4. Xác định giá trị của z như sau: cột đầu tiên bên trái là hai chữ số đầu tiên của z, hàng ngang đầu tiên là chữ số thứ ba của z.
5. Giao của cột và hàng này là giá trị $\Phi(z) = P(Z \leq z)$.

Sử dụng R

Cho $X \sim N(\mu;\sigma^2)$.

1. Tính $P(X \leq x)$

   `pnorm(`x`,`μ`,`σ`)`

2. Tính $P(a \leq X \leq b)$

   `pnorm(`b`,`μ`,`σ`)-pnorm(`a`,`μ`,`σ`)`

3. Vẽ đồ thị hàm mật độ của phân phối chuẩn

   `curve (dnorm(x,`μ`,`σ`), from = `a`, to = `b`)`

4. Tô màu phần diện tích (tiếp theo phần 3)

   `x<- seq(`u, v, c`)` # Giá trị $u \leq x \leq v$ và c là khoảng cách giữa các cột

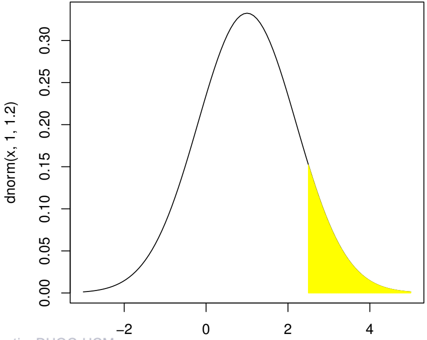
   `lines (x,dnorm(x,`μ`,`σ`), type='h', col="màu")`

5. Tìm t thỏa mãn $P(X \leq t) = a$.

   `qnorm(`a`,`μ`,`σ`)`

Vẽ đồ thị của hàm mật độ của phân phối chuẩn $N(1; 1, 2^2)$ và tô màu phần diện tích tương ứng với $x \geq 2.5$

```
curve (dnorm(x, 1, 1.2), from=-3, to=5)
x<-seq(2.5,5,0.01)
lines (x, dnorm(x, 1, 1.2), type='h', col="yellow")
```



Ví dụ 2.61 Cho biến ngẫu nhiên X có phân phối chuẩn chuẩn tắc $X \sim N(0;1)$.

a. Tính $P(X \leq 1,35)$.

b. Tính $P(-1,37 \leq X \leq 1,68)$.

Giải. a. Xem bảng A4, ta thấy

Table A4, continued. Standard Normal distribution

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545

$$P(X \leq 1,35) = \Phi(1,35) = 0,9115.$$

b. Ta thấy

$$P(-1,37 \leq X \leq 1,68) = \Phi(1,68) - \Phi(-1,37)$$
$$= 0,9535 - 0,0853$$
$$= 0,8682$$

Dùng R

Ví dụ 2.62 Cho biến ngẫu nhiên X có phân phối chuẩn chuẩn tắc $X \sim N(0;1)$.

a. Tính $P(X \leq 1,35)$.

b. Tính $P(-1,37 \leq X \leq 1,68)$.

a. R code: `pnorm(1.35, 0, 1)`
Kết quả: `0.911492`

b. R code: `pnorm(1.68, 0, 1)-pnorm(-1.37, 0, 1)`
Kết quả: `0.8681779`

Ví dụ 2.62 Giả sử thu nhập của các gia đình trong một tháng của một quốc gia có phân phối chuẩn với trung bình là 900$ và độ lệch chuẩn là 200$. Gặp ngẫu nhiên một gia đình.

a. Tính xác suất gia đình này có thu nhập từ 600$ đến 1200$.

b. Chính phủ sẽ trợ cấp cho 0,3% gia đình có mức thu nhập thấp nhất. Hỏi các gia đình này có thu nhập nhiều nhất là bao nhiêu $?

Giải. a. Đặt X là thu nhập của một gia đình. Theo giả thiết $X \sim N(900; 200^2)$ với $\mu = 900$ và $\sigma = 200$. Đặt $Z = \dfrac{X - 900}{200}$, khi đó

$$P(600 \leq X \leq 1200) = P\left(\frac{600-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{1200-\mu}{\sigma}\right)$$
$$= P(-1,5 \leq Z \leq 1,5) = \Phi(1,5) - \Phi(-1,5)$$
$$= 0,9332 - 0,0668 = 0,8664.$$

Xác suất gia đình này có thu nhập từ 600$ đến 1200$ là 86,64%.

b. Gọi t là thu nhập (tính bằng \$)) của các gia đình thỏa mãn

$$P(X \leq t) = 0,3\% = 0,003.$$

Như vậy

$$P(X \leq t) = P(Z \leq \frac{t-\mu}{\sigma}) = \Phi(\frac{t-\mu}{\sigma}) = 0,003$$

-2,75

z	-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	-0.00
-(3.9+)	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
-3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.6	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0002	.0002
-3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
-3.4	.0002	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003
-3.3	.0003	.0004	.0004	.0004	.0004	.0004	.0004	.0005	.0005	.0005
-3.2	.0005	.0005	.0005	.0006	.0006	.0006	.0006	.0006	.0007	.0007
-3.1	.0007	.0007	.0008	.0008	.0008	.0008	.0009	.0009	.0009	.0010
-3.0	.0010	.0010	.0011	.0011	.0011	.0012	.0012	.0013	.0013	.0013
-2.9	.0014	.0014	.0015	.0015	.0016	.0016	.0017	.0018	.0018	.0019
-2.8	.0019	.0020	.0021	.0021	.0022	.0023	.0023	.0024	.0025	.0026
-2.7	.0026	.0027	.0028	.0029	$\boxed{.0030}$	.0031	.0032	.0033	.0034	.0035
-2.6	.0036	.0037	.0038	.0039	.0040	.0041	.0043	.0044	.0045	.0047
-2.5	.0048	.0049	.0051	.0052	.0054	.0055	.0057	.0059	.0060	.0062

b. Gọi t là thu nhập (tính bằng $)) của các gia đình thỏa mãn

$$P(X \leq t) = 0,3\% = 0,003.$$

Như vậy

$$P(X \leq t) = P(Z \leq \frac{t-\mu}{\sigma}) = \Phi(\frac{t-\mu}{\sigma}) = 0,003$$

Ta thấy $z = -2,75$. Suy ra

$$\frac{t-900}{200} = -2,75$$

hay $t = 900 - 2,75.200 = 350.$

Như vậy, các gia đình này có thu nhập không quá 350$.

Dùng R:

R code: `qnorm(0.003, 900, 200)`

Kết quả: `350.4437`

2.6.3 Phân phối mũ

- Thời gian một khách hàng chờ để một nhân viên lễ tân phục vụ là bao nhiêu? (Giả sử nhân viên lễ tân đang phục vụ một khách hàng).
- Thời gian một khách hàng chờ để thực hiện một giao dịch tại một máy ATM? (Giả sử có người đang sử dụng máy ATM đó).

Đối với các câu hỏi trên, ta cần một khoảng thời gian để biến cố tiếp theo xảy ra. Phân phối mũ là phân phối "thời gian chờ đợi" với thời gian trung bình giữa các biến cố là $1/\lambda$. Thời gian chờ đợi một biến cố xảy ra là một biến ngẫu nhiên có *phân phối mũ* với tham số λ.

Ví dụ 2.63 Một người thấy rằng cứ trung bình khoảng 10 phút sẽ có 1 chiếc xe chạy ngang qua nhà mình. Sau khi thấy có một chiếc xe chạy ngang qua nhà. Đặt X là thời gian chờ chiếc xe tiếp theo chạy ngang qua. Ta có thể nói X có phân phối mũ với tham số $\lambda = 0,1$.

2.6.3 Phân phối mũ

Định nghĩa 2.64 Một biến ngẫu nhiên liên tục X được gọi là có phân phối mũ (exponential distribution) với tham số λ nếu hàm mật độ của nó có dạng

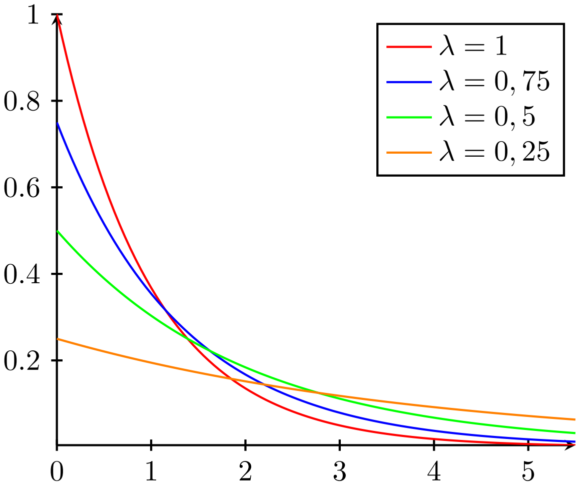
$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases},$$

trong đó $\lambda > 0$.

- Ký hiệu $X \sim \mathrm{Exp}(\lambda)$
- Hàm phân phối của X

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

Đồ thị của hàm mật độ



Cho X là biến ngẫu nhiên có phân phối mũ $X \sim \mathrm{Exp}(\lambda)$. Khi đó

1. Kỳ vọng: $E(X) = \dfrac{1}{\lambda}$
2. Phương sai: $\sigma^2 = \dfrac{1}{\lambda^2}$.
3. Hàm sinh moment: $M(t) = \dfrac{\lambda}{\lambda - t}$ với $t < \lambda$.

Sử dụng R

Cho $X \sim \mathrm{Exp}(\lambda)$.

1. Tính $P(X \leq x)$

   $$\texttt{pexp}(x, \lambda)$$

2. Tính $P(a \leq X \leq b)$

   $$\texttt{pexp}(b, \lambda)\texttt{-pexp}(a, \lambda)$$

3. Vẽ đồ thị hàm mật độ của phân phối mũ

   $$\texttt{curve (dexp(x,}\lambda\texttt{), from = } a\texttt{, to = } b\texttt{)}$$

4. Tô màu phần diện tích (tiếp theo phần 3)
   `x<- seq(`u, v, c`)` # Giá trị $u \leq x \leq v$ và c là khoảng cách giữa các cột
   `lines (x,dexp(x,`λ`), type='h', col="màu")`

5. Tìm t thỏa mãn $P(X \leq t) = a$.

   $$\texttt{qexp}(a, \lambda)$$

Ví dụ 2.65 Cho X là thời gian (năm) sử dụng của một loại laptop. Giả sử X có phân phối mũ và $E(X) = 15$. Chọn ngẫu nhiên một laptop.

a. Tính xác suất laptop đó sử dụng được ít hơn 6 năm.

b. Tính xác suất laptop đó sử dụng được nhiều hơn 18 năm.

c. Tính phương sai và độ lệch chuẩn của X.

Giải. Biến ngẫu nhiên X có phân phối mũ với

$$E(X) = \frac{1}{\lambda} = 15.$$

Hàm mật độ xác suất của X là

$$f(x) = \begin{cases} \frac{1}{15} e^{-\frac{1}{15}x}, & x \geq 0 \\ 0, & x < 0 \end{cases}.$$

a. $P(X < 6) = \int_{-\infty}^{6} f(x)\mathrm{d}x = \int_{0}^{6} \frac{1}{15} e^{-\frac{1}{15}x}\mathrm{d}x = 0,3297.$

b. $P(X > 18) = \int_{18}^{+\infty} f(x)\mathrm{d}x = \int_{18}^{+\infty} \frac{1}{15} e^{-\frac{1}{15}x}\mathrm{d}x = 0,3012.$

c. $\sigma^2 = \frac{1}{\lambda^2} = 225$ và $\sigma = 15.$

Dùng R:

a. `pexp(6, 1/15)`

b. `1-pexp(18, 1/15)`

Ví dụ 2.66 Cho X là thời gian (phút) phục vụ một khách hàng của một nhân viên lễ tân tại một khách sạn. Giả sử X có phân phối mũ với thời gian phục vụ trung bình một khách là 4 phút. Giả sử một khách đã được phục vụ 4 phút. Tính xác suất người đó được phục vụ thêm ít nhất 3 phút nữa.

Giải. Thời gian phục vụ trung bình $E(X) = \dfrac{1}{\lambda} = 4$.

Hàm mật độ xác suất của X là

$$f(x) = \begin{cases} \frac{1}{4}e^{-\frac{1}{4}x}, & x \geq 0 \\ 0, & x < 0 \end{cases}.$$

Khi đó

$$P(X \geq 7 | X > 4) = \frac{P(X \geq 7 \cap X > 4)}{P(X > 4)} = \frac{P(X \geq 7)}{P(X > 4)} = \frac{\int_7^{+\infty} \frac{1}{4}e^{-\frac{1}{4}x}\mathrm{d}x}{\int_4^{+\infty} \frac{1}{4}e^{-\frac{1}{4}x}\mathrm{d}x} = 0,4724.$$

Nhận xét: Ta thấy
$$P(X > 7|X > 4) = P(X > 3).$$
Tổng quát: Giả sử $X \sim \mathrm{Exp}(\lambda)$. Khi đó

$$P(X > t_1 + t_2|X > t_1) = P(X > t_2),$$

với mọi $t_1, t_2 \geq 0$.

Bài tập

Bài 2.11 Vị trí X của điểm có phân phối đều trên khoảng $(-3, 4)$. Xác suất mà vị trí của điểm sẽ không âm là gì? Giá trị kỳ vọng và phương sai của vị trí của điểm bằng bao nhiêu? (Đáp số: $\frac{4}{7}; \frac{1}{2}; \frac{49}{12}$)

Bài 2.12 Cho X là một biến ngẫu nhiên liên tục có phân phối chuẩn. Giả sử kỳ vọng và phương sai của X đều bằng k. Tìm khoảng $[a, b]$ của phân phối đều. (Đáp số: $[k - \sqrt{3k}, k + \sqrt{3k}]$)

Bài 2.13 Thời gian sửa chữa một laptop tại một trung tâm bảo hành có phân phối mũ với trung bình là 75 phút. Tính xác suất thời gian sửa 1 laptop nhiều hơn 2 giờ, biết rằng nó đã được sửa hơn 1,5 giờ. (Đáp số: 0,6703)

Bài 2.14 Giả sử thời gian phục vụ giữa 2 khách hàng liên tiếp bất kỳ tại một bưu điện là độc lập, có phân phối mũ với trung bình là 6 phút. Tính xác suất khách hàng thứ 10 phải chờ nhiều hơn 45 phút. (Đáp số: 0,7762)

Bài 2.15 Tuổi thọ của một loại bóng đèn có phân phối chuẩn với trung bình bằng 800 giờ và độ lệch chuẩn là 40 giờ. Tìm xác suất một bóng đèn có tuổi thọ trong khoảng thời gian từ 778 đến 834 giờ. (Đáp số: 0,5111)

Bài 2.16 Điểm trung bình cho một bài kiểm tra là 74 và độ lệch chuẩn là 7. Nếu 12% học sinh của một lớp được điểm A và điểm có phân phối chuẩn, thì điểm A thấp nhất có thể là bao nhiêu? Giả sử điểm được quy ước là số tự nhiên. (Đáp số: 83)

2.6.4 Phân phối gamma

Định nghĩa 2.67 Biến ngẫu nhiên X được gọi là có phân phối gamma (gamma distribution) với các tham số $\alpha > 0, \beta > 0$ nếu hàm mật độ của có dạng

$$f(x) = \begin{cases} \dfrac{x^{\alpha-1}e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, & x \geq 0 \\ 0, & \text{các trường hợp khác} \end{cases}$$

trong đó $\Gamma(\alpha)$ là giá trị của hàm gamma (gamma function) được xác định bởi

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

- Ký hiệu $X \sim \mathrm{gamma}(\alpha; \beta)$.

- $\Gamma(a) = (a-1)\Gamma(a-1)$ với mọi $a > 1$.
- Nếu n là một số nguyên dương thì $\Gamma(n) = (n-1)!$
- Nếu $\alpha = n$ là một số nguyên dương thì

$$f(x) = \begin{cases} \dfrac{x^{n-1}e^{-x/\beta}}{\beta^n (n-1)!}, & x \geq 0 \\ 0, & \text{các trường hợp khác} \end{cases}$$

hay

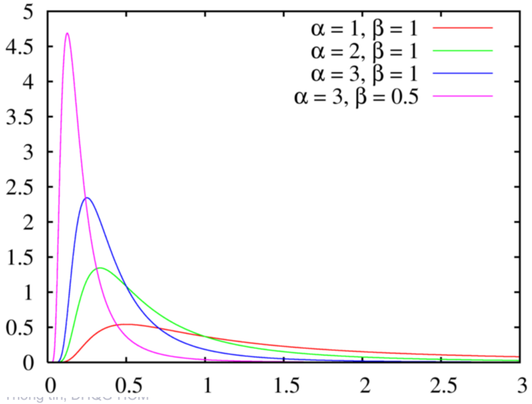
$$f(x) = \begin{cases} \dfrac{\beta^{-1}e^{-x/\beta}(x/\beta)^{n-1}}{(n-1)!}, & x \geq 0 \\ 0, & \text{các trường hợp khác} \end{cases}$$

- Cho α là một số tự nhiên. Phân phối gamma là phân phối thời gian biến cố thứ α xảy ra nếu thời gian giữa các biến cố xảy ra liên tiếp là độc lập, có phân phối mũ với thời gian trung bình giữa các biến cố là β.

Ví dụ 2.68 Giả sử thời gian phục vụ giữa 2 khách hàng liên tiếp bất kỳ tại một bưu điện là độc lập, có phân phối mũ với trung bình là 6 phút. Khi đó thời gian đến khi khách hàng thứ n được phục vụ có phân phối gamma với $\alpha = n$ và $\beta = 6$.

Ví dụ 2.69 Người dùng truy cập một trang web nhất định với trung bình là 12 lượt truy cập mỗi phút. Mỗi khách truy cập thứ sáu đều nhận được một số khuyến mãi dưới dạng biểu ngữ nhấp nháy. Khi đó thời gian giữa các đợt khuyến mãi liên tiếp có phân phối gamma với tham số $\alpha = 6$ và $\beta = 1/12$.

- Tham số α được gọi là tham số hình dạng (shape parameter) và tham số β được gọi là tham số tỷ lệ (scale parameter). Khi tăng α, đồ thị của hàm mật độ sẽ "nhọn" hơn, trong khi tăng β làm tăng "độ phân tán" của phân phối.
- Đồ thị của hàm mật độ



Nhận xét: Cho biến ngẫu nhiên $X \sim \text{gamma}(\alpha, \beta)$.

- Kỳ vọng $E(X) = \alpha\beta$
- Phương sai $\text{Var}(X) = \alpha\beta^2$
- Hàm sinh moment $M(t) = \dfrac{1}{(1-\beta t)^\alpha}$
- Nếu $\alpha = 1$ thì phân phối $\text{gamma}(1, \beta) = \text{Exp}(\frac{1}{\beta})$ là phân phối mũ. Thật vậy, khi đó hàm mật độ xác suất có dạng

$$f(x) = \begin{cases} \dfrac{1}{\beta} e^{-x/\beta}, & x \geq 0; , \beta > 0 \\ 0, & \text{các trường hợp khác} \end{cases}$$

Sử dụng R

Cho $X \sim \text{gamma}(\alpha, \beta)$.

1. Tính $P(X \leq x)$

   `pgamma(`x`,`α`,`β`)`

2. Tính $P(a \leq X \leq b)$

   `pgamma(`b`,`α`,`β`)-pgamma(`a`,`α`,`β`)`

3. Vẽ đồ thị hàm mật độ của phân phối gamma

   `curve (dgamma(x,`α`,`β`), from = `a`, to = `b`)`

4. Tô màu phần diện tích (tiếp theo phần 3)

   `x<- seq(`u, v, c`)` # Giá trị $u \leq x \leq v$ và c là khoảng cách giữa các cột

   `lines (x,dgamma(x,`α`,`β`), type='h', col="màu")`

5. Tìm t thỏa mãn $P(X \leq t) = a$.

   `qgamma(`a`,`α`,`β`)`

Ví dụ 2.70 Mức tiêu thụ nhiên liệu hàng ngày tính bằng triệu lít tại một sân bay có thể được coi là một biến ngẫu nhiên có phân phối gamma với $\alpha = 3, \beta = 1$.

a. Xác suất mà vào một ngày nào đó mức tiêu thụ nhiên liệu sẽ dưới 1 triệu lít là bao nhiêu?

b. Giả sử kho nhiên liệu của sân bay chỉ có thể chứa 2 triệu lít nhiên liệu. Xác suất nhiên liệu sẽ không đủ vào ngày một ngày nào đó là bao nhiêu?

Giải. Đặt X là mức tiêu thụ nhiên liệu (triệu lít) tại sân bay trong một ngày. Theo đề bài $X \sim \text{gamma}(\alpha = 3, \beta = 1)$ và

$$f(x) = \frac{1}{\Gamma(3).1^3} x^{3-1} e^{-x} = \frac{1}{2} x^2 e^{-x}$$

với $x > 0$

a. Theo đề bài, ta tính

$$P(X < 1) = \int_0^1 \frac{1}{2}x^2 e^{-x}\mathrm{d}x = 1 - \frac{5}{2e} = 0,0803.$$

b. Vì sân bay chỉ chứa được 2 triệu lít nhiên liệu nên khi sân bay không cung cấp đủ nhiên liệu tức là $X > 2$.Ta cần tính

$$P(X > 2) = \int_2^{\infty} \frac{1}{2}x^2 e^{-x}\mathrm{d}x = 0,677.$$

Dùng R:

a. `pgamma(1, 3, 1)`

b. `1-pgamma(2, 3, 1)`

Ví dụ 2.71 Thời gian giữa hai cuộc gọi liên tiếp đến tổng đài là một biến ngẫu nhiên độc lập, có phân phối mũ với số lượng 6 cuộc gọi mỗi giờ.

a. Nếu tổng đài bắt đầu trả lời các cuộc gọi từ 9:00 sáng, thời gian trung bình đến khi có cuộc gọi thứ 12 là bao nhiêu?

b. Giả sử thời gian giữa lần gọi thứ 2 và lần thứ 3 vượt quá 4 phút. Xác suất thời gian giữa lần gọi thứ 5 và lần thứ 6 vượt quá 7 phút là bao nhiêu?

Giải. Đặt X_n thời gian đến khi có cuộc gọi thứ n. Vì thời gian giữa các cuộc gọi độc lập, có phân phối mũ với 6 cuộc gọi mỗi giờ nên $X_n \sim \text{gamma}(\alpha = n, \beta = \frac{1}{6})$.

a. Thời gian trung bình đến khi có cuộc gọi thứ 12 ($\alpha = 12$)

$$E(X) = \alpha\beta = 12.\frac{1}{6} = 2.$$

Vì thời gian bắt đầu nhận các cuộc gọi là 9:00 giờ nên thời gian trung bình để có cuộc gọi thứ 12 là 11 giờ.

b. Giả sử thời gian giữa lần gọi thứ 2 và lần thứ 3 vượt quá 4 phút. Xác suất thời gian giữa lần gọi thứ 5 và lần thứ 6 vượt quá 7 phút là bao nhiêu?

Giải. b. Thời gian giữa cuộc gọi thứ 2 và thứ 3 là $t_3 = X_3 - X_2$.
Thời gian giữa cuộc gọi thứ 5 và thứ 6 là $t_6 = X_6 - X_5$. Ta cần tính

$$P\left(t_6 > \frac{7}{60} \mid t_3 > \frac{4}{60}\right).$$

Vì khoảng thời gian giữa các cuộc gọi là độc lập và có phân phối mũ nên

$$P\left(t_6 > \frac{7}{60} \mid t_3 > \frac{4}{60}\right) = P(t_6 > \frac{7}{60})$$
$$= 1 - (1 - e^{-\frac{7}{60}\frac{1}{\beta}}) = 0,4966.$$

2.6.5 Phân phối χ^2

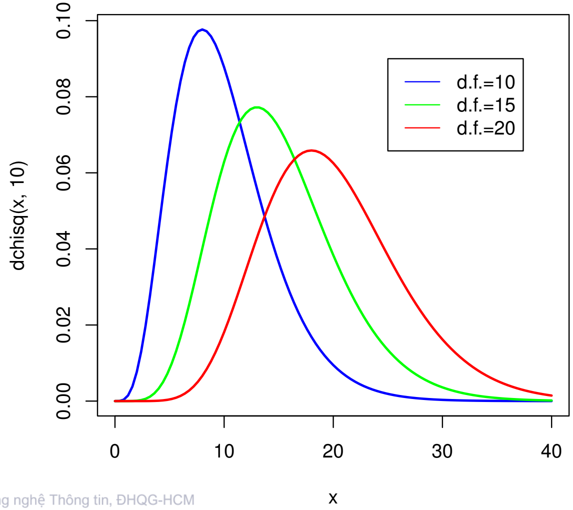
Một trường hợp đặc biệt của phân phối gamma khi $\alpha = \dfrac{n}{2}$ và $\beta = 2$ được gọi là phân phối chi-bình phương (χ^2-distribution).

Định nghĩa 2.72 Cho n là một số nguyên dương. Một biến ngẫu nhiên X được gọi là có **phân phối** χ^2 (chi-bình phương) với tham số n, ký hiệu $X \sim \chi^2(n)$, nếu hàm mật độ xác suất của nó có dạng

$$f(x) = \begin{cases} \dfrac{1}{2^{n/2}\Gamma(n/2)} x^{(n/2)-1} e^{-x/2}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

Tham số n được gọi là **bậc tự do** (degree of freedom) của X.

Đồ thị hàm mật độ của phân phối chi-bình phương



Cho biến ngẫu nhiên $X \sim \chi^2(n)$. Khi đó

- Kỳ vọng $E(X) = n$.
- Phương sai $\mathrm{Var}(X) = 2n$.
- Hàm sinh moment $M(t) = \dfrac{1}{(1-2t)^{n/2}}$ với $t < \dfrac{1}{2}$.
- Nếu biến ngẫu nhiên $X \sim \mathrm{gamma}(\alpha, \beta)$ thì biến ngẫu nhiên $Y = \dfrac{2X}{\beta} \sim \chi^2(2\alpha)$.

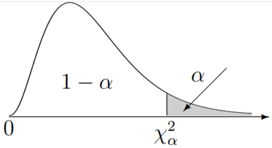
- Nếu $X \sim \chi^2(n)$ thì ta có thể tính giá trị $\chi^2_\alpha(n)$ như sau (Xem bảng A6)

$$P(X > \chi^2_\alpha(n)) = \alpha.$$

Ví dụ 2.72 Cho biến ngẫu nhiên $X \sim \chi^2(15)$. Tìm a sao cho $P(X > a) = 0,95$.
Giải. Đặt $a = \chi^2_\alpha(15)$. Vì $P(X > \chi^2_\alpha(n)) = \alpha$ nên $\alpha = 0,95$.
Xem bảng phụ lục A6, cột tự do (d.f.) bên trái, hàng ngang thứ 2, tại vị trí 0.95. Chọn hàng có bậc tự do 15 và cột giá trị $\alpha = 0.95$. Số nằm tại vị trí giao của hàng và cột này là $\chi^2_{0.95}(15)$.

Table A6. Chi-Square Distribution

χ^2_α; critical values, such that $\boldsymbol{P}\left\{\chi^2 > \chi^2_\alpha\right\} = \alpha$



ν (d.f.)	α, the right-tail probability													
	.999	.995	.99	.975	$\boxed{.95}$	.90	.80	.20	.10	.05	.025	.01	.005	.001
1	0.00	0.00	0.00	0.00	0.00	0.02	0.06	1.64	2.71	3.84	5.02	6.63	7.88	10.8
2	0.00	0.01	0.02	0.05	0.10	0.21	0.45	3.22	4.61	5.99	7.38	9.21	10.6	13.8
3	0.02	0.07	0.11	0.22	0.35	0.58	1.01	4.64	6.25	7.81	9.35	11.3	12.8	16.3
4	0.09	0.21	0.30	0.48	0.71	1.06	1.65	5.99	7.78	9.49	11.1	13.3	14.9	18.5
5	0.21	0.41	0.55	0.83	1.15	1.61	2.34	7.29	9.24	11.1	12.8	15.1	16.7	20.5
6	0.38	0.68	0.87	1.24	1.64	2.20	3.07	8.56	10.6	12.6	14.4	16.8	18.5	22.5
7	0.60	0.99	1.24	1.69	2.17	2.83	3.82	9.80	12.0	14.1	16.0	18.5	20.3	24.3
8	0.86	1.34	1.65	2.18	2.73	3.49	4.59	11.0	13.4	15.5	17.5	20.1	22.0	26.1
9	1.15	1.73	2.09	2.70	3.33	4.17	5.38	12.2	14.7	16.9	19.0	21.7	23.6	27.9
10	1.48	2.16	2.56	3.25	3.94	4.87	6.18	13.4	16.0	18.3	20.5	23.2	25.2	29.6
11	1.83	2.60	3.05	3.82	4.57	5.58	6.99	14.6	17.3	19.7	21.9	24.7	26.8	31.3
12	2.21	3.07	3.57	4.40	5.23	6.30	7.81	15.8	18.5	21.0	23.3	26.2	28.3	32.9
13	2.62	3.57	4.11	5.01	5.89	7.04	8.63	17.0	19.8	22.4	24.7	27.7	29.8	34.5
14	3.04	4.07	4.66	5.63	6.57	7.79	9.47	18.2	21.1	23.7	26.1	29.1	31.3	36.1
$\boxed{15}$	3.48	4.60	5.23	6.26	$\boxed{7.26}$	8.55	10.3	19.3	22.3	25.0	27.5	30.6	32.8	37.7

[3]

[3] Nguồn: Michael Baron, *Probability and statistics for computer scientists*, 3rd edition

Ví dụ 2.73 Cho biến ngẫu nhiên $X \sim \chi^2(15)$. Tìm a sao cho $P(X > a) = 0,95$.
Giải. Đặt $a = \chi^2_\alpha(15)$. Vì $P(X > \chi^2_\alpha(n)) = \alpha$ nên $\alpha = 0,95$.
Xem bảng A6, cột tự do (d.f.) bên trái, hàng ngang thứ 2, tại vị trí 0.95. Chọn hàng có bậc tự do 15 và cột giá trị $\alpha = 0.95$. Số nằm tại vị trí giao của hàng và cột này là $\chi^2_{0.95}(15)$.
Như vậy

$$P(X > 7,26) = 0,95.$$

Suy ra $a = 7,26$.

Sử dụng R

Cho $X \sim \chi^2(n)$.

1. Tính $P(X \leq x)$

   pchisq(x,n)

2. Tính $P(a \leq X \leq b)$

   pchisq(b,n)-pchisq(a,n)

3. Vẽ đồ thị của hàm mật độ của phân phối chi-bình phương

   curve (dchisq(x,n), from = a, to = b)

4. Tô màu phần diện tích (tiếp theo phần 3)

   x<- seq(u, v, c) # Giá trị $u \leq x \leq v$ và c là khoảng cách giữa các cột

   lines (x,dchisq(x,n), type='h', col="màu")

5. Tìm t thỏa mãn $P(X \leq t) = a$.

   qchisq(a,n)

Ví dụ 2.74 Cho biến ngẫu nhiên $X \sim \chi^2(20)$.

a. Tìm a sao cho $P(X > a) = 0,95$.

b. Tính $P(X \leq 12,4)$.

Giải. a. Xem bảng A6, hàng tại vị trí cột tự do 20 và cột giá trị $\alpha = 0.95$. Số nằm tại vị trí giao của hàng và cột này là $\chi^2_{0.95}(20)$.
Như vậy
$$P(X > 10,9) = 0,95.$$
Suy ra $a = 10,9$.
b. Từ bảng A6, ta có
$$P(X \leq 12,4) = 1 - P(X > 12,4) \quad = 1 - P(X > \chi^2_{0,9}(20)) = 1 - 0,9 = 0,1.$$

Dùng R:

a. `qchisq(0.05, 20)` #$(P(X \leq a) = 0.05)$

b. `pchisq(12.4, 20)`

Định lý 2.75 Cho biến ngẫu nhiên $X \sim N(\mu, \sigma^2)$. Khi đó $\left(\dfrac{X-\mu}{\sigma}\right)^2 \sim \chi^2(1)$.

Ví dụ 2.76 Cho biến ngẫu nhiên $X \sim N(7,4)$. Tính $P(15,364 \leq (X-7)^2 \leq 20,095)$.

Giải. Vì $X \sim N(7,4)$ nên $\mu = 7$ và $\sigma = 2$. Khi đó

$$\begin{aligned}
P(15,364 \leq (X-7)^2 \leq 20,095) &= P\left(\frac{15,364}{4} \leq \left(\frac{X-7}{2}\right)^2 \leq \frac{20,095}{4}\right) \\
&= P\left(3,841 \leq Z^2 \leq 5,024\right) \\
&= P\left(Z^2 \leq 5,024\right) - P\left(Z^2 \leq 3,841\right) \\
&= \left(1 - P\left(Z^2 > \chi^2_{0,025}(1)\right)\right) - \left(1 - P\left(Z^2 > \chi^2_{0,05}(1)\right)\right) \\
&= (1-0,025) - (1-0,05) \\
&= 0,975 - 0,95 = 0,025
\end{aligned}$$

Định lý 2.77 Cho $X_1, X_2, \ldots, X_n$ là n biến ngẫu nhiên độc lập có cùng phân phối $N(0,1)$. Khi đó biến ngẫu nhiên $Y = X_1^2 + X_2^2 + \ldots + X_n^2$ có phân phối chi-bình phương với *bậc tự do* n.

Nhận xét: Cho $X_1, X_2, \ldots, X_n$ là các biến ngẫu nhiên độc lập, có cùng phân phối $N(\mu, \sigma^2)$. Khi đó các biến ngẫu nhiên $Z_i = \dfrac{X_i - \mu}{\sigma}$ là độc lập và $\sum_{i=1}^{n} Z_i^2$ có phân phối chi-bình phương với bậc tự do n.

Ví dụ 2.78 Cho $X_1, X_2, \ldots, X_5$ là các biến ngẫu nhiên có phân phối $N(5,1)$. Tìm a sao cho

$$P\left(\sum_{i=1}^{5}(X_i-5)^2 \leq a\right) = 0,9.$$

Giải. Ta thấy $\sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{5}\left(\frac{X_i-5}{1}\right)^2 = \sum_{i=1}^{5}(X_i-5)^2$ có phân phối chi-bình phương với bậc tự do là 5. Khi đó

$$P\left(\sum_{i=1}^{5} Z_i^2 \leq a\right) = 1 - P\left(\sum_{i=1}^{5} Z_i^2 > a\right)$$

Suy ra $P\left(\sum_{i=1}^{5} Z_i^2 > a\right) = 0,1$. Do đó $a = \chi^2_{0,1}(5) = 9,24$.

2.6.6 Phân phối Student

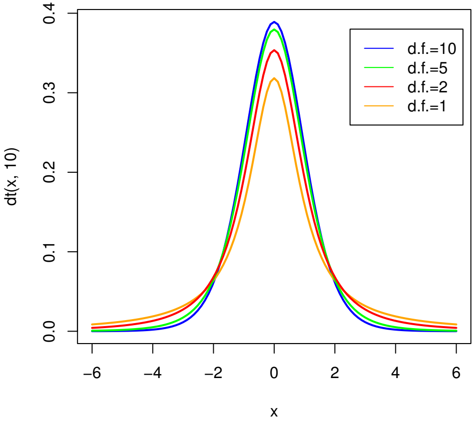
Định nghĩa 2.79 Cho Y, Z là các biến ngẫu nhiên độc lập với $Y \sim \chi^2(n)$ và $Z \sim N(0,1)$. Khi đó biến ngẫu nhiên

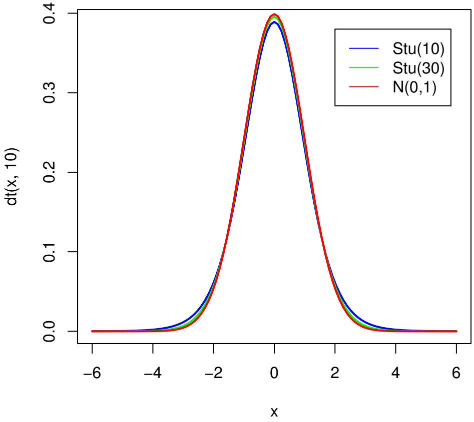
$$T = \frac{Z}{\sqrt{Y/n}}$$

được gọi là có **phân phối Student** hay t-phân phối (Student distribution, t-distribution) với bậc tự do n, ký hiệu $T \sim Stu(n)$.
Khi đó hàm mật độ xác suất của T là

$$f(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})}\left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}.$$

Đồ thị của hàm mật độ của một số phân phối Student

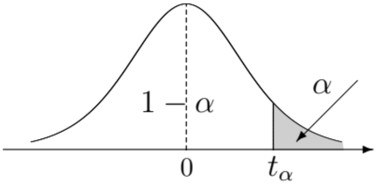




- Khi $n \geq 30$, phân phối Student gần giống phân phối chuẩn.
- Sử dụng bảng A5 để tính xác suất[4]

Table A5. Student's T-distribution

t_α; critical values, such that $\boldsymbol{P}\{t > t_\alpha\} = \alpha$



Dùng bảng A5, tìm t_α sao cho

$$P(T > t_\alpha) = \alpha.$$

[4] Nguồn ảnh: Michael Baron, *Probability and statistics for computer scientists*, 3rd edition

Ví dụ 2.80 Cho T là một biến ngẫu nhiên có phân phối Student với bậc tự do là 9 và $\alpha = 0,01$.

ν (d.f.)	α, the right-tail probability									
	.10	.05	.025	.02	.01	.005	.0025	.001	.0005	.0001
1	3.078	6.314	12.706	15.89	31.82	63.66	127.3	318.3	636.6	3185
2	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60	70.71
3	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92	22.20
4	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610	13.04
5	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.894	6.869	9.676
6	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959	8.023
7	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408	7.064
8	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041	6.442
9	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781	6.009
10	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587	5.694

Dùng bảng A5[5], ta có

$$P(T > 2,821) = 0,01.$$

[5] Nguồn: Michael Baron, *Probability and statistics for computer scientists*, 3rd edition

Sử dụng R

Cho $X \sim Stu(n)$.

1. Tính $P(X \leq x)$

   `pt(`x`,`n`)`

2. Tính $P(a \leq X \leq b)$

   `pt(`b`,`n`)-pt(`a`,`n`)`

3. Vẽ đường cong phân phối Student

   `curve (dt(x,`n`), from = `a`, to = `b`)`

4. Tô màu phần diện tích (tiếp theo phần 3)

   `x<- seq(`u, v, c`)` # Giá trị $u \leq x \leq v$ và c là khoảng cách giữa các cột

   `lines (x,dt(x,`n`), type='h', col="màu")`

5. Tìm t thỏa mãn $P(X \leq t) = a$.

   `qt(`a`,`n`)`

Định lý 2.81 Cho $X_1, X_2, \ldots, X_n$ là các biến ngẫu nhiên độc lập, cùng có phân phối chuẩn $N(\mu, \sigma^2)$. Đặt

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ và } S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2.$$

Khi đó biến ngẫu nhiên $T = \dfrac{\overline{X} - \mu}{S/\sqrt{n}}$ có phân phối Student với bậc tự do $n-1$.

Bài tập

Bài 2.17 Số lượng khách hàng đến một cửa hàng trung bình trong 1 giờ là 30 người. Giả sử thời gian đến giữa hai khách hàng liên tiếp là một biến ngẫu nhiên có phân phối mũ. Tính xác suất người bán hàng phải chờ hơn 5 phút để phục vụ khách hàng thứ 2. (Đáp số: 0,287)

Bài 2.18 Tuổi thọ của chip nhớ máy tính có phân phối gamma với kỳ vọng là 12 năm và độ lệch chuẩn là 4 năm. Xác suất mà một con chip như vậy có tuổi thọ từ 8 đến 10 năm là bao nhiêu? (Đáp số: 0,185)

Bài 2.19 Cho X là một biến ngẫu nhiên có phân phối chi-bình phương với bậc tự do 20.

a. Tính $P(X \leq 12,443)$. (Đáp số: 0,1)

b. Tìm a sao cho $P(X > a) = 0,95$. (Đáp số: $a = 10,851$)

Bài 2.20 Cho X là một biến ngẫu nhiên có phân phối Student với bậc tự do 10. Tính $P(|X| > 2,228)$. (Đáp số: 0,05)

Bài 2.21 Cho X là một biến ngẫu nhiên có phân phối Student với bậc tự do 14. Tìm b sao cho Tính $P(|X| < b) = 0,9$. (Đáp số: 1,761)