Biến ngẫu nhiên và các phân phối xác suất thường dùng

Nguyễn Thị Hồng Nhung
nthnhung@hcmus.edu.vn

Ngày 9 tháng 10 năm 2023

Nội dung

Biến ngẫu nhiên

Ví dụ 1

Thực hiện phép thử tung hai đồng xu cân đối đồng chất. Không gian mẫu của phép thử này như sau

$$\Omega = \{SS, SN, NS, NN\}$$

Gọi X là số mặt ngửa xuất hiện. Khi đó, X sẽ nhận các giá trị sau $X(\{SS\}) = 0, X(\{SN\}) = X(\{NS\}) = 1, X(\{NN\}) = 2.$

Biến ngẫu nhiên

Ví dụ 2

Thực hiện phép thử: gieo một con xúc xắc.
Gọi u_i là sự kiện mặt nhận được có i chấm. Khi đó, không gian mẫu là:

$$\Omega = \{u_1, u_2, u_3, u_4, u_5, u_6\}.$$

Người chơi sẽ nhận được 1\$ *nếu gieo được mặt có 1 chấm, nhận được* 2\$ *nếu gieo được mặt có 2 hoặc 3 chấm và nhận được* 4\$ *nếu gieo được mặt có 4 hoặc 5 hoặc 6 chấm.*
Gọi $X(\$)$ là số tiền người chơi nhận được sau một lần gieo xúc xắc. Khi đó, X sẽ nhận các giá trị sau $\{1, 2, 4\}$.

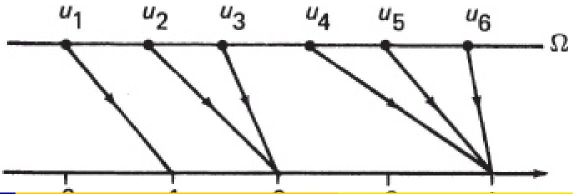
Hình: *1*

Biến ngẫu nhiên

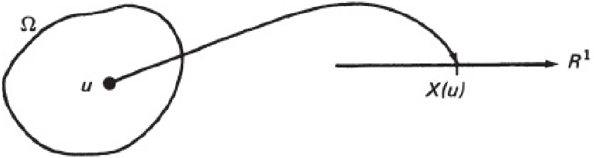
Định nghĩa 1

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

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



Biến ngẫu nhiên



Chain/hinh32.jpg

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

Nếu một biến số ngẫu nhiên có tập giá trị hữu hạn hoặc vô hạn đếm được của các giá trị khác nhau, biến số ngẫu nhiên được gọi là rời rạc (Xem ví dụ 2) . Ngược lại nếu tập giá trị của biến số ngẫu nhiên là vô hạn thì biến số ngẫu nhiên đó được gọi là biến số ngẫu nhiên liên tục.

Phân phối xác suất của biến ngẫu nhiên

Kí hiệu

Cho $\mathbf{X} \subset \mathcal{X}$. Ta kí hiệu

$$(\mathbf{X} \subset \mathcal{X}) = \{\omega \in \Omega : X(\omega) \in \mathcal{X}\}.$$

Chẳng hạn, ta viết

$$(X = x) = \{\omega \in \Omega : X(\omega) = x\}.$$

$$(X \leq x) = \{\omega \in \Omega : X(\omega) \leq x\}.$$

Quy luật phân phối xác suất

Định nghĩa 2

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

Định nghĩa 3 (Hàm phân phối)

Hàm phân phối xác suất của biến ngẫu nhiên X(xác định trên không gian các biến cố sơ cấp Ω) là hàm $F(x)$ được định nghĩa

$$\begin{aligned} F_X : \mathbb{R} &\rightarrow [0, 1] \\ x &\mapsto F_X(x) = \mathbb{P}(X \leq x) \end{aligned}$$

Quy luật phân phối xác suất

Định lý 1

Hàm phân phối của biến ngẫu nhiên X, $F_X(x)$, có các tính chất

- $\lim\limits_{x \to -\infty} F_X(x) = 0, \ \lim\limits_{x \to +\infty} F_X(x) = 1,$
- $F_X(x)$ *là một hàm không giảm,*
- $F_X(x)$ *là một hàm liên tục phải tại một giá trị x bất kỳ,*
- *Nếu $a \leq b$ thì $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$.*

Định nghĩa

Định nghĩa 4 (Biến ngẫu nhiên rời rạc)

Biến ngẫu nhiên được gọi là rời rạc nếu tập hợp các giá trị mà nó có thể nhận là tập hữu hạn hoặc vô hạn đếm được của các giá trị khác nhau.

Định nghĩa 5 (Hàm xác suất)

Xác suất để biến ngẫu nhiên rời rạc X nhận giá trị $k \in \mathbb{R}$, ký hiệu $p_X(k)$ được xác định như sau

$$p_X(k) = \mathbb{P}(X = k) = \mathbb{P}(\{\omega \in \Omega : X(\omega) = k\})$$

$p_X(k)$ được gọi là hàm xác suất của biến ngẫu nhiên rời rạc X.

Ví dụ

Ví dụ 3

- *Biến ngẫu nhiên trong phép thử tung hai đồng xu ở ví dụ trên là một biến ngẫu nhiên rời rạc.*
- *Số cuộc điện thoại đến một tổng đài ở bưu điện trong một ngày.*
- *Số sản phẩm bị lỗi của một lô hàng.*
- ...

Hàm xác suất

Tính chất 1 (Hàm xác suất)

Xét một biến ngẫu nhiên rời rạc X có thể nhận các giá trị $x_1, x_2, x_3, \ldots, x_n$. Một hàm giá trị xác suất(gọi tắt là hàm xác suất) là hàm thỏa

(1) $p_X(x_i) = \mathbb{P}(X = x_i)$

(2) $p_X(x_i) \geq 0$

(3) $\sum_{i=1}^{n} p_X(x_i) = 1$

Bảng phân phối xác suất

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

- Dòng thứ nhất là các giá trị mà biến ngẫu nhiên X nhận được.
- Dòng thứ hai là xác suất biến ngẫu nhiên X nhận các giá trị tương ứng.

X	x_1	x_2	...	x_n	...
$p(x_i) = \mathbb{P}(X = x_i)$	$p(x_1)$	$p(x_2)$	...	$p(x_n)$	...

Bảng phân phối xác suất

Ví dụ 4

Gọi X là số mặt sấp xuất hiện khi thực hiện phép thử: tung một cặp đồng xu cân đối đồng chất. Ta có, không gian mẫu

$$\Omega = \{SS, SN, NS, NN\}.$$

Khi đó, $X \in \{0, 1, 2\}$. $\mathbb{P}(X = 0) = \mathbb{P}(\{NN\}) = \frac{1}{4}$, $\mathbb{P}(X = 1) = \mathbb{P}(\{SN, NS\}) = \frac{2}{4}$, $\mathbb{P}(X = 2) = \mathbb{P}(\{SS\}) = \frac{1}{4}$. *Bảng phân phối của biến ngẫu nhiên X*

X	$x_1 = 0$	$x_2 = 1$	$x_3 = 2$
$p(x_i) = \mathbb{P}(X = x_i)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

Hàm phân phối xác suất của B.N.N rời rạc

Định nghĩa 6 (Hàm phân phối xác suất của b.n.n rời rạc)

Hàm phân phối xác suất của một biến ngẫu nhiên rời rạc X, ký hiệu $F(x)$ được xác định như sau

$$F_X(x) = \mathbb{P}(X \leq x) = \sum_{j \leq x} p_X(j). \tag{1}$$

Cụ thể

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

Bảng phân phối xác suất-Hàm phân phối xác suất

Ví dụ 5

Cho biến ngẫu nhiên rời rạc X có hàm xác suất

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

(a) *Lập bảng phân phối xác suất.*

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

Bảng phân phối xác suất-Hàm phân phối xác suất

Ví dụ 6

Một lô hàng có 850 sản phẩm trong đó có 50 sản phẩm kém chất lượng. Chọn ngẫu nhiên lần lượt 2 sản phẩm không hoàn lại. Gọi X là số sản phẩm không đạt chất lượng trong 2 sản phẩm được chọn.

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

(b) *Viết hàm phân phối xác suất.*

Biến ngẫu nhiên liên tục

Biến ngẫu nhiên liên tục

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

Ví dụ 7

Chọn ngẫu nhiên một hợp chất hóa học và đo độ pH, X, của nó. Khi đó X là một biến ngẫu nhiên liên tục, vì mọi pH đều nằm trong khoảng từ 0 đến 14.

Hàm phân phối của biến ngẫu nhiên liên tục

Định nghĩa 7 (Hàm phân phối xác suất của b.n.n liên tục)

Hàm phân phối xác suất của một biến ngẫu nhiên rời rạc X, ký hiệu $F_X(x)$ được xác định như sau

$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(t)dt. \tag{2}$$

Hàm mật độ xác suất

Định nghĩa 8 (Hàm mật độ xác suất)

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

(i)

$$\mathbb{P}(X \in I) = \int_I f_X(x)dx, \ \forall I \subset \mathbb{R}.$$

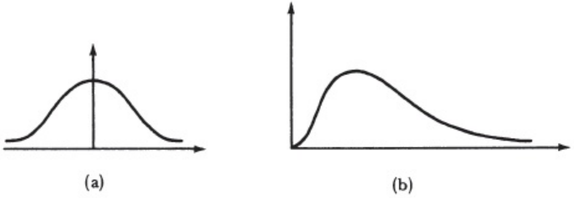
(ii)

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

hàm số $f_X(x)$ được gọi là hàm mật độ xác suất của biến ngẫu nhiên X.

Hàm mật độ xác suất

Chain/hinh37.jpg



Hình: 3

Ví dụ về hàm mật độ xác suất được cho trong hình 3; trong (a) hàm mật độ xác suất đối xứng, trong (b) hàm mật độ xác suất lệch.

Nhận xét

Nhận xét 1

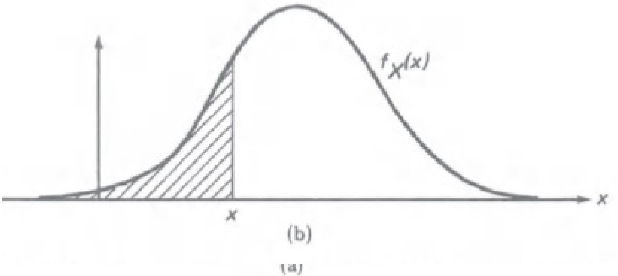
Mọi hàm $f_X(x)$ không âm và thỏa điều kiện $\int_{-\infty}^{\infty} f_X(x)dx = 1$ đều là hàm mật độ xác suất của một biến ngẫu nhiên X nào đó.

Nhận xét

Nhận xét 2

(1) *Từ định nghĩa về hàm mật độ ta có hàm phân phối xác suất của biến ngẫu nhiên X có hàm mật độ $f(x)$ là*
$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(t)dt.$

(2) $F'(x) = f_X(x).$



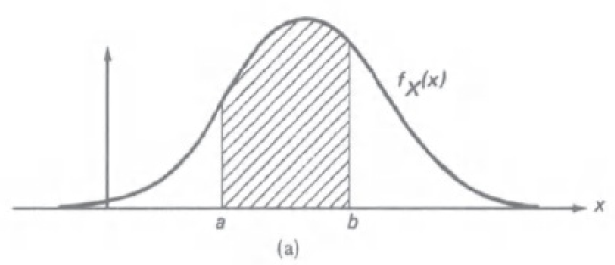
Chain/hinh38.jpg

Nhận xét

Nhận xét 3

Nếu X là biến ngẫu nhiên liên tục, với 2 số thực a, b bất kỳ, ta có

$$\mathbb{P}(a \le X \le b) = \mathbb{P}(a < X \le b) = \mathbb{P}(a \le X < b) = \mathbb{P}(a < X < b)$$



Chain/hinh38a.jpg

Trong hình trên, xác suất $a < X < b$ bằng với diện tích của miền

Phân vị

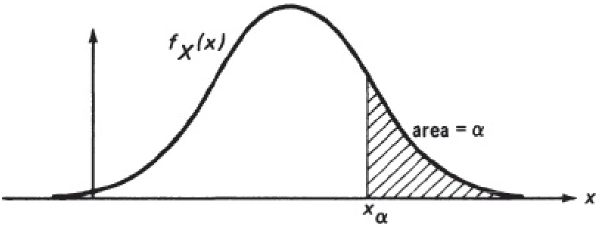
Định nghĩa 9

Nghiệm $x = x_\alpha$ của phương trình

$$F_X(x) = 1 - \alpha$$

được gọi là phân vị α của biến ngẫu nhiên X .

Chain/hinh310.jpg

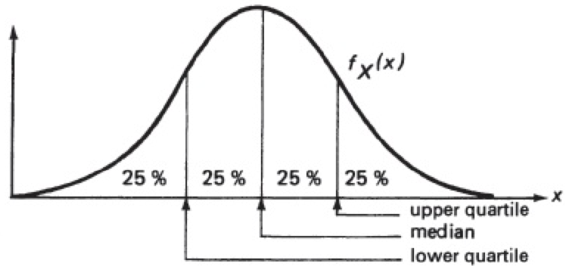


Hình: 7 Phân vị

Phân vị

Các phân vị $x_{0.25}, x_{0.50}, x_{0.75}$ được gọi là upper quartile, trung vị (median), lower quartile. Trung vị chia diện tích của miền bên dưới hàm mật độ xác suất thành hai phần diện tích bằng nhau.

Chain/hinh311.jpg



Hình: 8-Các phân vị và trung vị.

Kỳ vọng của biến ngẫu nhiên rời rạc

Định nghĩa 10 (Kỳ vọng của biến ngẫu nhiên rời rạc)

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

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

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

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

Kỳ vọng của biến ngẫu nhiên rời rạc

Ví dụ 8

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

Kỳ vọng của biến ngẫu nhiên liên tục

Định nghĩa 11 (Kỳ vọng của biến ngẫu nhiên liên tục)

Giả sử biến ngẫu nhiên liên tục X có hàm mật độ xác suất $f(x)$, kỳ vọng của X là

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} xf(x)dx. \tag{3}$$

Kỳ vọng của biến ngẫu nhiên liên tục

Ví dụ 9

Cho biến ngẫu nhiên X có hàm mật độ xác suất

$$f(x) = \begin{cases} 20e^{-20(x-12.5)} & x \geq 12.5 \\ 0 & \textit{nơi khác} \end{cases}$$

(a) *Tính* $\mathbb{P}(X > 12,60)$.

(b) *Tính kỳ vọng của* X.

Ý nghĩa của kỳ vọng

(i) Là giá trị trung bình theo xác suất của tất cả các giá trị có thể của biến ngẫu nhiên.

(ii) Kỳ vọng phản ánh giá trị trung tâm của phân phối xác suất.

Mệnh đề 1 (Tính chất của kỳ vọng)

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

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

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

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

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

Kỳ vọng của biến ngẫu nhiên

Ví dụ 10

Tung đồng thời 2 cặp đồng xu. Gọi X là số mặt sấp xuất hiện ở cặp thứ nhất, Y là số mặt ngửa xuất hiện ở cặp thứ hai. Tính $E(X^2), E(Y^2), E(XY)$.

Phương sai của biến ngẫu nhiên

Định nghĩa 12 (Phương sai)

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

$$\mathbb{V}ar(X) = \mathbb{E}\left([X - \mathbb{E}(X)]^2\right) = \mathbb{E}\left(X^2\right) - [\mathbb{E}(X)]^2 . \tag{4}$$

Phương sai của biến ngẫu nhiên

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

$$\mathbb{V}ar(X) = \sum_{x} (x-\mu)^2 p(x) = \sum_{x} x^2 p(x) - \mu^2.$$

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

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

Ý nghĩa của phương sai

- Phương sai là trung bình của bình phương các sai lệch giữa X và $\mathbb{E}(X)$. Phương sai phản ánh mức độ phân tán các giá trị của biến ngẫu nhiên xung quanh giá trị trung bình.
- Trong công nghiệp, phương sai biểu thị độ chính xác trong sản xuất. Trong canh tác, phương sai biểu thị mức độ ổn định của năng suất,...

Tính chất phương sai

Cho hai biến ngẫu nhiên X, Y và hằng số thực $C \in \mathbb{R}$, phương sai có các tính chất sau

(i) $\mathbb{V}ar(C) = 0$.

(ii) $\mathbb{V}ar(CX) = C^2\mathbb{V}ar(X)$

(iii) Nếu X, Y độc lập thì $\mathbb{V}ar(X + Y) = \mathbb{V}ar(X) + \mathbb{V}ar(Y)$.

Độ lệch chuẩn

Định nghĩa 13 (Độ lệch chuẩn)

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

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

Ví dụ 11

Cho biến ngẫu nhiên Y có hàm mật độ xác suất

$$g(y) = \begin{cases} \frac{2}{y^2} & \text{nếu } y \in [1, 2] \\ 0 & \text{nếu } y \notin [1, 2]. \end{cases}$$

Tính phương sai, độ lệch chuẩn của biến ngẫu nhiên Y

Một vài phân phối rời rạc: Two-point Distribution

Định nghĩa 14 (Two-point Distribution)

Nếu biến ngẫu nhiên X chỉ có hai giá trị a và b với xác suất tương ứng là p và q, thì X được gọi là có phân phối hai điểm (two-point).

Định lý 2

Biến ngẫu nhiên X có phân phối "hai điểm" (two-point)

- $\mathbb{E}(X) = ap + bq$,
- $\mathbb{V}(X) = (a-b)^2 pq$

Hàm xác suất là

$$p_X(a) = p, \quad p_X(b) = q$$

Một vài phân phối rời rạc:Biến ngẫu nhiên Bernoulli

Định nghĩa 15 (Biến ngẫu nhiên Bernoulli)

Thực hiện một phép thử, ta quan tâm đến biến cố A. Nếu biến cố A xảy ra (thành công) thì X nhận giá trị là 1 $(X = 1)$*, ngược lại biến ngẫu nhiên X nhận giá trị 0. Phép thử này gọi là phép thử Bernoulli. Giả sử xác suất xảy ra biến cố A là p,* $0 < p < 1$

$$\mathbb{P}(A) = \mathbb{P}(X = 1) = p$$

và

$$\mathbb{P}(\bar{A}) = \mathbb{P}(X = 0) = 1 - p = q.$$

Khi đó, biến ngẫu nhiên X được gọi là biến ngẫu nhiên có phân phối Bernoulli với tham số p, ký hiệu $X \sim B(1; p)$*.*

Một vài phân phối rời rạc:Biến ngẫu nhiên Bernoulli

Bảng phân phối xác suất

Bảng phân phối xác suất của biến ngẫu nhiên $X \sim B(1; p)$ có dạng

X	1	0
$\mathbb{P}$	p	q

với $q = 1 - p$.

Dựa vào bảng phân phối xác suất của biến ngẫu nhiên X, ta có

$$\mathbb{E}(X) = 1.p + 0.q = p$$

$$\mathbb{V}ar(X) = 0^2 \times q + 1^2 \times p - p^2 = p(1-p) = pq.$$

Một vài phân phối rời rạc:Biến ngẫu nhiên Bernoulli

Mô hình có phân phối Bernoulli: Xem xét một thí nghiệm ngẫu nhiên có hai kết quả $\Omega = \{\omega, \bar{\omega}\}$, trong đó $P(\omega) = p$.

Gọi X là số lần ω xuất hiện

$$\begin{aligned} X : \Omega &\to \mathbb{R} \\ \omega &\mapsto X(\omega) = 1 \\ \bar{\omega} &\mapsto X(\bar{\omega}) = 0 \end{aligned}$$

Ta có, $\mathbb{P}(X = 1) = \mathbb{P}(\omega) = p \qquad \mathbb{P}(X = 0) = \mathbb{P}(\bar{\omega}) = 1 - p = q$.

Vậy X có hàm mật độ

$$p_X(x) = \begin{cases} 1 - p & \text{khi } x = 0 \\ p & \text{khi } x = 1 \\ 0 & \text{nơi khác} \end{cases}$$

nghĩa là X có phân phối Bernoulli

Một vài phân phối rời rạc:Biến ngẫu nhiên Bernoulli

Ví dụ 12

Các phép thử sau đây cho kết quả là một biến ngẫu nhiên Bernoulli

- *Tung ngẫu nhiên một đồng xu:* $X = 1$ *nếu xuất hiện mặt sấp,* $X = 0$ *nếu xuất hiện mặt ngửa,* $X \sim B(1, 1/2)$.
- *Mua vé số:* $X = 0$ *nếu không trúng số,* $X = 1$ *nếu trúng số.*
- *Trả lời ngẫu nhiên một câu trắc nghiệm :* $X = 0$ *nếu trả lời đúng,* $X = 1$ *nếu trả lời sai.*

Nhận xét: *Mọi thí nghiệm ngẫu nhiên có hai kết quả đều có phân phối Bernoulli.*

Một vài phân phối rời rạc:Phân phối nhị thức

Định nghĩa 16 (Binomial distribution)

Thực hiện n phép thử Bernoulli độc lập, với xác suất thành công trong mỗi phép thử là p. Gọi X là số lần thành công (biến cố A xảy ra) trong n phép thử thì

$$X = X_1 + X_2 + \ldots + X_n$$

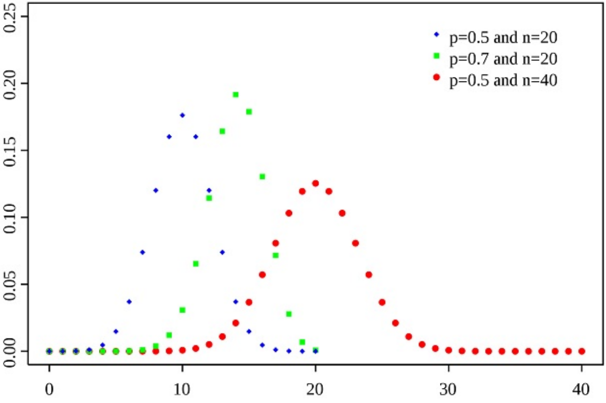
với X_i là biến ngẫu nhiên có phân phối Bernoulli với cùng tham số p, $(i = 1, 2, \ldots, n)$.Khi đó X là biến ngẫu nhiên rời rạc với miền giá trị $S = \{0, 1, \ldots., n\}$và hàm xác suất có dạng

$$p_X(k) = \mathbb{P}(X = k) = \begin{cases} C_n^k p^k q^{n-k}, & \text{với } k = 0, 1, 2, \ldots, n \\ 0 & \text{nơi khác} \end{cases}$$

trong đó, $p \in [0, 1]$. X được gọi là có phân phối nhị thức với tham số

Một vài phân phối rời rạc:Phân phối nhị thức

Hình: Hàm xác suất - Phân phối nhị thức



Một vài phân phối rời rạc:Phân phối nhị thức

Mô hình nhị thức: Xem xét một thí nghiệm ngẫu nhiên có hai kết quả $\Omega_* = \{\omega_*, \bar{\omega_*}\}$ với $P(\omega_*) = p$. Ta lập lại thí nghiệm này n lần độc lập với nhau và quan tâm đến số lần xuất hiện ω_* trong n lần quan sát đó.

Không gian mẫu của n lần thí nghiệm là

$$\Omega = \{\omega = (\omega_{(1)}, \omega_{(2)}, ..., \omega_{(n)}) : \omega_{(i)} \in \{\omega_*, \bar{\omega_*}\}, i = 1, 2, ..., n\}$$

Đặt X_i là kết quả lần quan sát thứ i

$$\begin{aligned} X_i : \Omega_* &\rightarrow \mathbb{R} \\ \omega_* &\mapsto X_i(\omega_*) = 1 \\ \bar{\omega_*} &\mapsto X_i(\bar{\omega_*}) = 0 \end{aligned}$$

Gọi X là số lần suất hiện ω_* trong n lần quan sát.

$$X : \Omega \rightarrow \mathbb{R}$$

Một vài phân phối rời rạc:Phân phối nhị thức

Với $x \in \{0, 1, 2, ..., n\}$,
nghĩa là nó chứa các kết quả của n lần thí nghiệm mà trong đó có x lần xuất hiện ω_* và $n - x$ lần xuất hiện $\bar{\omega}_*$.
Vì mỗi phép thử Bernoulli là độc lập nên với mỗi $\omega \in (X = k)$ thì

$$\mathbb{P}(\omega) = p^k(1 - p)^{n-k}.$$

Số biến sơ cấp của $(X = x)$ là C_n^x. Do đó,

$$p_X(k) = \mathbb{P}(X = k) = \sum_{\omega \in (X=x)} \mathbb{P}(\omega) = C_n^k p^k q^{n-k}.$$

Vậy X có phân phối nhị thức.

Một vài phân phối rời rạc:Phân phối nhị thức

Ví dụ 13

Một lô thuốc, có tỉ lệ hỏng $p = 0.2$ *. Ta lấy ngẫu nhiên 5 lọ. Gọi* X *là số lọ hỏng trong số lọ lấy ra. Tìm hàm xác suất của* X.

Bài giải 1

Quan sát lọ thuốc trong 5 lần độc lập.
Gọi X *là số lọ bị hỏng trong 5 lọ lấy ra.*
Vậy $X \in \{0, 1, 2, .., 5\}$ *và* $X \sim B(5; 0.2)$ *với hàm xác suất*

$$p_x(h) = \mathbb{P}(X = h) = \begin{cases} \mathbf{C}_5^h (0.2)^h (1 - 0.2)^{5-h}, & h = 0, 1, ..., 5 \\ 0 & \textit{nơi khác} \end{cases}$$

Một vài phân phối rời rạc:Phân phối nhị thức

Ta có bảng phân phối xác suất

H	0	1	2	3	4	5
$\mathbb{P}(H = h)$	0.32768	0.4096	0.2048	0.0512	0.0064	0.0032

Một vài phân phối rời rạc:Phân phối nhị thức

Ví dụ 14

Trong một nhà máy sản xuất vi mạch điện tử, biết rằng tỷ lệ vi mạch không đạt chất lượng là 5%. Kiểm tra ngẫu nhiên 15 vi mạch. Tính xác suất

- a *Có đúng 7 vi mạch không đạt chất lượng*
- b *Có ít nhất 1 vi mạch không đạt chất lượng*

Một vài phân phối rời rạc:Phân phối nhị thức

Định lý 3 (Các đặc trưng của BNN có phân phối nhị thức)

Nếu X là biến ngẫu nhiên có phân phối nhị thức $Bin(n;p)$ thì

1. $\mathbb{E}(X) = np$.
2. $\mathbb{V}ar(X) = npq$.
3. *Với x, h là hai số nguyên dương thì*

$$\begin{aligned} \mathbb{P}(x \leq X \leq x+h) &= \mathbb{P}(X=x) + \mathbb{P}(X=x+1) + \ldots \\ \ldots &+ \mathbb{P}(X=x+h) \end{aligned}$$

Định lý 4

Nếu $X \sim B(n_1, p)$ và $Y \sim B(n_2, p)$, giả sử X và Y là hai biến ngẫu nhiên độc lập, thì $X + Y \sim B(n_1 + n_2, p)$.

Một vài phân phối rời rạc:Phân phối Poisson

Định nghĩa 17 (Phân phối Poisson)

Nếu biến ngẫu nhiên X có hàm xác suất

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

thì X được gọi là có phân phối Poison.
Ký hiệu: $X \sim Po(\lambda)$ hoặc $X \sim P(\lambda)$.

Một vài phân phối rời rạc:Phân phối Poisson

Các đặc trưng của biến ngẫu nhiên có phân phối Poisson

Định lý 5

Nếu biến ngẫu nhiên X có phân phối Poisson với tham số λ, $X \sim P(\lambda)$ thì kỳ vọng và phương sai của X lần lượt bằng

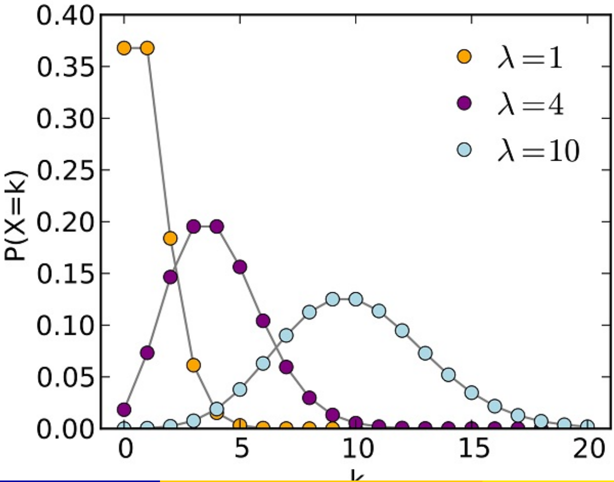
$$\mathbb{E}(X) = \lambda, \mathbb{V}ar(X) = \lambda.$$

Nhận xét 4

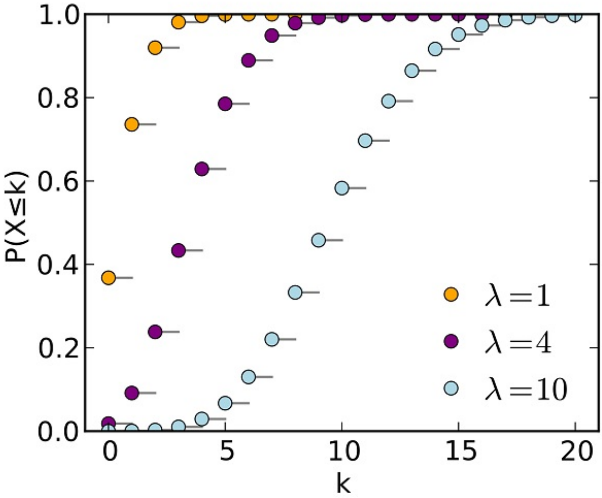
$\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{\lambda}$.

Một vài phân phối rời rạc:Phân phối Poisson

Hình: Hàm xác suất-Phân phối Poisson



Một vài phân phối rời rạc:Phân phối Poisson



Hình: Hàm phân phối-Phân phối Poisson

Một vài phân phối rời rạc:Phân phối Poisson

Xấp xỉ phân phối nhị thức bằng phân phối Poisson

Định lý 6

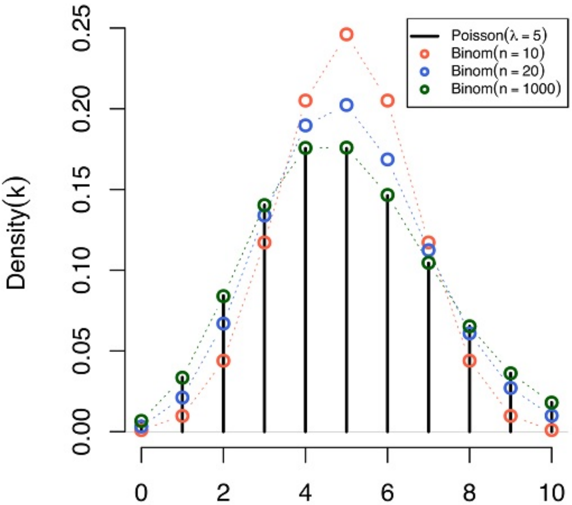
Cho $X \sim B(n,p)$*, nếu* $n \to \infty$ *và* $p \to 0$ *sao cho* $np \to m$ *thì*

$$\mathbb{P}(X = x) = \frac{e^{-m}m^x}{x!}.$$

Trong thực tế, phân phối Poisson sẽ xấp xỉ tốt cho phân phối nhị thức khi $n \geq 100$ và $np \leq 20, p \leq 0.01$.

Một vài phân phối rời rạc:Phân phối Poisson

Xấp xỉ phân phối nhị thức bằng phân phối Poisson



Một vài phân phối rời rạc:Phân phối Poisson

Mô hình Poisson: Một số biến ngẫu nhiên mô tả các sự kiện sau thường được xem là tuân theo phân phối Poisson

i. Số lỗi in trong một (hoặc một số) trang sách.
ii. Số người sống lâu trên 100 tuổi trong một cộng đồng dân cư.
iii. Số người đến một bưu điện nào đó trong một ngày.
iv. Số tai nạn hoặc sự cố giao thông xảy ra tại một điểm giao thông trong một ngày....

Các biến ngẫu nhiên được sử dụng để mô tả, ''đếm''số lần xảy ra của một biến cố, sự kiện nào đó xảy ra trong một khoảng thời gian và thỏa một số điều kiện (các điều kiện này thường thỏa trong thực tế) thường được mô tả bằng phân phối Poisson.

Một vài phân phối rời rạc:Phân phối Poisson

Ví dụ 15

Giả sử số lỗi in trong một trang sách nào đó của quyển sách có phân phối Poisson với tham số $\lambda = \frac{1}{2}$. *Tính xác suất có ít nhất một lỗi in trong trang này.*

Một vài phân phối rời rạc:Phân phối Poisson

Ví dụ 16

Số cuộc điện thoại gọi đến một tổng đài điện thoại trong một giờ có phân phối Poisson với $\lambda = 10$*. Tính xác suất*

i *có 5 cuộc điện thoại gọi đến trong một giờ.*

ii *Có nhiều nhất 3 cuộc điện thoại gọi đến trong một giờ.*

iii *Có 15 cuộc gọi đến trong hai giờ.*

Một vài phân phối rời rạc:Phân phối Poisson

Định lý 7

Nếu $X \sim P(\lambda_1)$ *và* $Y \sim P(\lambda_2)$, *trong đó* X, Y *là độc lập, chúng ta có* $X + Y \sim P(\lambda_1 + \lambda_2)$.

Ví dụ 17

Trong một đợt tiêm chủng cho trẻ em ở một khu vực, biết xác suất một trẻ bị phản ứng với thuốc sau khi tiêm là $0,001$. *Thực hiện tiêm cho* 2000 *trẻ. Tính xác suất có nhiều nhất 1 trẻ bị phản ứng với thuốc sau khi tiêm.*

Một vài phân phối rời rạc:Phân phối đều rời rạc

Định nghĩa 18 (Phân phối đều rời rạc)

Nếu biến ngẫu nhiên X có giá trị $1, 2, \ldots, m$ với xác suất bằng nhau, bằng $\frac{1}{m}$, thì X được gọi là có phân phối đều.

Hàm xác suất

$$p_X(k) = \mathbb{P}(X = k) = \frac{1}{m}, \ k = 1, 2, \ldots, m.$$

Định lý 8

Nếu biến ngẫu nhiên X có phân phối đều rời rạc thì

- $\mathbb{E}(X) = \frac{m+1}{2}$
- $\mathbb{V}(X) = \frac{m^2-1}{12}$

Một vài phân phối rời rạc: Phân phối hình học

Định nghĩa 19 (Phân phối hình học)

Nếu biến ngẫu nhiên X có hàm xác suất

$$p_X(k) = \mathbb{P}(X = k) = q^k p \quad (k = 0, 1, 2, \ldots),$$

trong đó $q = 1 - p$, thì X được gọi là có một phân phối hình học.

Định nghĩa 20 (Phân phối *fft − Distribution*)

Nếu một biến ngẫu nhiên X có hàm xác suất

$$p_X(k) = q^{k-1} p \quad (k = 1, 2, \ldots)$$

trong đó $1 = 1 - p$, thì X được goi là có phân phối fft − Distribution.

Một vài phân phối rời rạc: Phân phối siêu bội

Định nghĩa 21 (Phân phối siêu bội)

Nếu X là biến ngẫu nhiên có hàm xác suất

$$p_X(k) = \frac{C_a^k C_b^{n-k}}{C_{a+b}^n}$$

trong đó k là số nguyên sao cho $0 \leq k \leq a$, $0 \leq n-k \leq b$, thì X được gọi là có phân phối siêu bội.

b

Một vài phân phối liên tục: Phân phối đều

Định nghĩa 22 (Phân phối đều)

Nếu biến ngẫu nhiên X có hàm mật độ

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{nếu } a \le x \le b, \\ 0 & \text{ngược lại,} \end{cases}$$

thì X được gọi là có phân phối đều.

Chú ý 1

Để tránh nhầm lẫn với trường hợp biến ngẫu nhiên rời rạc có phân phối đều, chúng ta gọi biến ngẫu nhiên có phân phối đều trên (a, b) cho trường hợp biến ngẫu nhiêu liên tục có phân phối đều.

Một vài phân phối liên tục: Phân phối đều

Với hàm mật độ của biến ngẫu nhiên có phân phối đều trên (a, b) được định nghĩa như trên, ta có được hàm phân phối như bên dưới

$$F(x) = \begin{cases} 0, & \text{khi } x < a \\ \frac{x-a}{b-a}, & \text{khi } x \in [a, b]] \\ 1, & \text{khi } x > b. \end{cases}$$

Hình: 9-Hàm mật độ và hàm phân phối của một biến ngẫu nhiên có phân phối đều trên (a,b).

Một vài phân phối liên tục: Phân phối đều

Định lý 9 (Các đặc trưng của biến ngẫu nhiên có phân phối đều)

Cho X là biến ngẫu nhiên có phân phối đều trên $[a, b]$ (nghĩa là $X \sim U([a, b])$) thì

i *Kỳ vọng* $\mathbb{E}(X) = \frac{a+b}{2}$.

ii *Phương sai* $\mathbb{V}ar(X) = \frac{(b-a)^2}{12}$.

Một vài phân phối liên tục: Phân phối mũ

Định nghĩa 23 (Hàm phân phối mũ)

Nếu T là biến ngẫu nhiên có hàm mật độ

$$f_T(t) = \begin{cases} \frac{1}{m}\exp(-\frac{t}{m}) & \textit{nếu } t \geq 0, \\ 0 & \textit{nếu } t < 0, \end{cases}$$

trong đó $m > 0$, thì X được gọi là có phân phối mũ (Exponential distribution).

Ký hiệu : $T \sim Exp(\frac{1}{m})$.

Hàm phân phối của biến ngẫu nhiên X

$$F(t) = \begin{cases} 0, & \text{khi } t < 0 \\ 1 - \exp(-\frac{t}{m}), & \text{khi } t \geq 0 \end{cases}$$

Một vài phân phối liên tục: Phân phối mũ

Định lý 10

Nếu $T \sim Exp(m)$ thì kỳ vọng và phương sai của T lần lượt là $\mathbb{E}(T) = m$, $\mathbb{V}ar(T) = m^2$

Hình: Hàm mật độ xác suất-Phân phối mũ

Một vài phân phối liên tục: Phân phối mũ

Hình: Hàm phân phối-Phân phối mũ

Một vài phân phối liên tục: Phân phối chuẩn

Định nghĩa 24 (Phân phối chuẩn)

Nếu biến ngẫu nhiên X có hàm mật độ

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), -\infty < x < \infty$$

trong đó μ, σ là các hằng số và $\sigma > 0, -\infty < \mu < \infty$, thì X được gọi là có phân phối chuẩn.
ký hiệu $X \sim \mathcal{N}(\mu, \sigma^2)$.

Một vài phân phối liên tục: Phân phối chuẩn tắc(Standard Normal Distribution)

Nếu $\mu = 0,\ \sigma = 1$ chúng ta có được $X \sim \mathcal{N}(0, 1)$. Khi đó, chúng ta gọi biến ngẫu nhiên X có phân phối chuẩn tắc.

Hàm mật độ của biến ngẫu nhiên này

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

Hàm phân phối của biến ngẫu nhiên có phân phối chuẩn tắc thường được ký hiệu $\Phi(x)$ và được xác định như sau

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt.$$

Một vài phân phối liên tục: Phân phối chuẩn tắc(Standard Normal Distribution)

Hình: 1- Phân phối chuẩn tắc. (a) Hàm mật độ $X \sim \mathcal{N}(0, 1)$. (b) Hàm phân phối $X \sim \mathcal{N}(0, 1)$.

Một vài phân phối liên tục: Phân phối chuẩn tắc(Standard Normal Distribution)

Tính chất 2

- $\Phi(-x) = 1 - \Phi(x)$.
- $\mathbb{P}(a < X \leq b) = \Phi(b) - \Phi(a)$.

Với giá trị cụ thể của x, ta tra bảng để tìm giá trị $\phi(x)$.

Ví dụ 18

Cho biến ngẫu nhiên $X \sim \mathcal{N}(0, 1)$. *Tính các xác suất sau*

- $\mathbb{P}(X < 1.55)$
- $\mathbb{P}(X < -1.45)$
- $\mathbb{P}(-1 \leq X < 1.5)$

Một vài phân phối liên tục:Phân phối chuẩn

Chúng ta xem xét biến ngẫu nhiên $X \sim \mathcal{N}(\mu, \sigma^2)$.

matdo1.jpg

Một vài phân phối liên tục:Phân phối chuẩn

phanphoi1.jpg

Một vài phân phối liên tục:Phân phối chuẩn

* Đồ thị có dạng hình chuông
* Phân phối đối xứng
* Trung bình = trung vị (median)= Mode
* Vị trí của phân phối được xác định bởi kỳ vọng μ
* Độ phân tán được xác định bởi độ lệch tiêu chuẩn σ
* Xác định trên $\mathbb{R}$.

Một vài phân phối liên tục:Phân phối chuẩn

Định lý 11

Nếu $X \sim \mathcal{N}(\mu, \sigma^2)$ *và* $Y = \frac{X-\mu}{\sigma}$ *thì* $Y \sim \mathcal{N}(0,1)$

Định lý 12

Nếu $X \sim \mathcal{N}(\mu, \sigma^2)$ *thì*

$$\mathbb{E}(X) = \mu, \quad \mathbb{V}ar(X) = \sigma^2.$$

Một vài phân phối liên tục:Phân phối chuẩn

Định lý 13

Nếu $Y = aX + b$, *trong đó* $X \sim \mathcal{N}(\mu, \sigma^2)$ *thì*

$$Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$$

Hệ quả 1

Nếu $X \sim \mathcal{N}(\mu, \sigma^2)$ *thì* $\mathbb{P}(X < a) = \phi\left(\frac{a-\mu}{\sigma}\right)$.

Hệ quả 2

Nếu $X \sim \mathcal{N}(\mu, \sigma^2)$ *thì* $\mathbb{P}(a \leq X < b) = \phi\left(\frac{b-\mu}{\sigma}\right) - \phi\left(\frac{a-\mu}{\sigma}\right)$.

Xấp xỉ phân phối nhị thức bằng phân phối chuẩn

Với n đủ lớn, phân phối nhị thức có thể xấp xỉ bởi phân phối chuẩn. Khi đó $X \sim \mathcal{N}(np, npq)$. Ta có

$$\mathbb{P}(a < X \leq b) \approx \Phi\left(\frac{b - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a - np}{\sqrt{npq}}\right) \tag{5}$$

Độ chính xác nói chung có phần cải thiện nếu, b được thay bởi $b + \frac{1}{2}$ và a được thay bởi $a + \frac{1}{2}$, khi đó

$$\mathbb{P}(a < X \leq b) \approx \Phi\left(\frac{b + \frac{1}{2} - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a + \frac{1}{2} - np}{\sqrt{npq}}\right) \tag{6}$$

Chúng ta nói rằng chúng ta có một sự "hiệu chỉnh liên tục".

Chú ý 2

$npq \geq 10$.

Một vài phân phối liên tục: Phân phối Weibull

Định nghĩa 25 (Phân phối Weibull)

Nếu biến ngẫu nhiên X có hàm mật độ

$$f_x(x) = \begin{cases} \frac{c}{a}\left(\frac{x}{x}\right)^{c-1} \exp\left(-\left(\frac{x}{a}\right)^c\right), & khi \geq 0, \\ 0, & khi\ x < 0 \end{cases}$$

trong đó a và c là những số dương, thì X được gọi là có phân phối Weibull.

Hàm phân phối

$$F(x) = \begin{cases} 0, & \text{khi } x < 0 \\ 1 - \exp\left(-\left[\frac{x}{a}\right]^c\right), & \text{khi } x \geq 0 \end{cases}$$

Một vài phân phối liên tục: Phân phối Weibull

Hình: 11- Hàm mật độ của biến ngẫu nhiên có phân phối Weibull (z=1; c=0.5, 1.0, 3.0).

Một vài phân phối liên tục: Phân phối Gamma

Định nghĩa 26 (Phân phối Gamma)

Nếu biến ngẫu nhiên X có hàm mật độ

$$f_x(x) = \begin{cases} \frac{1}{a^p\Gamma(p)} x^{p-1} \exp\left(-\frac{x}{a}\right), & khi \geq 0, \\ 0, & khi\ x < 0 \end{cases}$$

trong đó $p > 0$ và $a > 0$, thì X được gọi là có phân phối Gamma.

$$\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx \quad (p > 0).$$

Một vài phân phối liên tục: Phân phối Gamma

Hình: 11- Hàm mật độ của biến ngẫu nhiên có phân phối Gamma (a=1; p=1,2,3)