Biến Ngẫu Nhiên

Khoa Công nghệ Thông tin và Truyền thông

CMC UNIVERSITY

Nội dung

1. Biến ngẫu nhiên
2. Biến ngẫu nhiên Rời rạc
   - Biến ngẫu nhiên Bernoulli
   - Biến ngẫu nhiên Nhị thức
   - Biến ngẫu nhiên Poisson
3. Biến ngẫu nhiên Liên tục
   - Biến ngẫu nhiên Đều
   - Biến ngẫu nhiên Chuẩn
4. Kỳ vọng của một Biến ngẫu nhiên
   - Biến Rời rạc
   - Biến Liên tục
   - Kỳ vọng của một Hàm số của Biến ngẫu nhiên

Biến ngẫu nhiên - Hàm phân bố

Ví dụ

- Tung hai đồng xu, gọi X là số đồng xu cho mặt sấp. Khi đó, X là một biến ngẫu nhiên.
- Thời gian sử dụng T của một viên pin là một biến ngẫu nhiên.
- Số cây đổ tại Hà Nội sau trận bão Yagi là một biến ngẫu nhiên.
- Số sinh viên trúng tuyển vào trường CMC năm sau là một biến ngẫu nhiên.

Biến ngẫu nhiên

Định nghĩa 1

Biến ngẫu nhiên là một hàm thực xác định trên không gian mẫu.

Biến ngẫu nhiên

Định nghĩa 1

Biến ngẫu nhiên là một hàm thực xác định trên không gian mẫu.

- Là một phép gán mỗi biến cố với một giá trị thực $X = f(\omega)$.
- Giá trị của biến ngẫu nhiên phụ thuộc vào biến cố.
- Với biến cố ta có khái niệm xác suất để biến cố xảy ra. Với biến ngẫu nhiên, ta có khái niệm xác suất để biến ngẫu nhiên nhận một giá trị cụ thể.

Ví dụ

Ví dụ 1

Gieo hai đồng xu cân, và gọi Y là số lần xuất hiện mặt ngửa. Khi đó, Y là một biến ngẫu nhiên nhận một trong các giá trị $0, 1, 2$ với xác suất tương ứng là

$$\mathrm{P}(Y = 0) = \mathrm{P}\left(\{S, S\}\right) = \frac{1}{4},$$
$$\mathrm{P}(Y = 1) = \mathrm{P}\left(\{S, N\}, \{N, S\}\right) = \frac{1}{2},$$
$$\mathrm{P}(Y = 2) = \mathrm{P}\left(\{N, N\}\right) = \frac{1}{4}.$$

Ví dụ

Ví dụ 1

Gieo hai đồng xu cân, và gọi Y là số lần xuất hiện mặt ngửa. Khi đó, Y là một biến ngẫu nhiên nhận một trong các giá trị $0, 1, 2$ *với xác suất tương ứng là*

$$\mathrm{P}(Y = 0) = \mathrm{P}\left(\{S, S\}\right) = \frac{1}{4},$$

$$\mathrm{P}(Y = 1) = \mathrm{P}\left(\{S, N\}, \{N, S\}\right) = \frac{1}{2},$$

$$\mathrm{P}(Y = 2) = \mathrm{P}\left(\{N, N\}\right) = \frac{1}{4}.$$

$\mathrm{P}(Y = 0) + \mathrm{P}(Y = 1) + \mathrm{P}(Y = 2) = 1.$

Hàm phân bố

Định nghĩa 2

Hàm phân bố chồng chất (cumulative distribution function - cdf) hay hàm phân bố của một biến ngẫu nhiên X, ký hiệu là $F(\cdot)$, là một hàm thực, xác định với mọi $-\infty \le a \le \infty$, và được cho bởi

$$F(a) = \mathrm{P}(X \le a).$$

Tính chất

(i) $F(a)$ là hàm không giảm.

(ii) $\lim\limits_{a \to \infty} F(a) = F(\infty) = 1.$

(iii) $\lim\limits_{a \to -\infty} F(a) = F(-\infty) = 0.$

(iv) $F(b) - F(a) = \mathrm{P}(a < X \leq b)$ với mọi $a \leq b$.

Biến ngẫu nhiên Rời rạc

Biến ngẫu nhiên Rời rạc

Định nghĩa 3

Cho X là một biến ngẫu nhiên. Khi đó, X được gọi là biến rời rạc nếu tập hợp các giá trị có thể của X là một tập hữu hạn hoặc đếm được.

Xác suất - Hàm chồng chất

Hàm khối lượng (probability mass function - pmf) của một biến rời rạc X được ký hiệu bởi $p(a)$, với

$$p(a) = \mathrm{P}(X = a).$$

Xác suất - Hàm chồng chất

Hàm khối lượng (probability mass function - pmf) của một biến rời rạc X được ký hiệu bởi $p(a)$, với

$$p(a) = \mathrm{P}(X = a).$$

Giả sử X có thể nhận các giá trị $x_1, x_2, \ldots$ Khi đó,

- $p(x_i) > 0$ với mọi $i = 1, 2, \ldots$
- $p(x) = 0$ với mọi $x \notin \{x_1, x_2, \ldots\}$.
- $p(x_1) + p(x_2) + \cdots = 1$.

Xác suất - Hàm chồng chất

Hàm khối lượng (probability mass function - pmf) của một biến rời rạc X được ký hiệu bởi $p(a)$, với

$$p(a) = \mathrm{P}(X = a).$$

Giả sử X có thể nhận các giá trị $x_1, x_2, \ldots$ Khi đó,

- $p(x_i) > 0$ với mọi $i = 1, 2, \ldots$
- $p(x) = 0$ với mọi $x \notin \{x_1, x_2, \ldots\}$.
- $p(x_1) + p(x_2) + \cdots = 1$.

Hàm phân bố chồng chất $F(a)$ được xác định bởi

$$F(a) = \sum_{x_i \leq a} p(x_i).$$

Ví dụ

Ví dụ 2

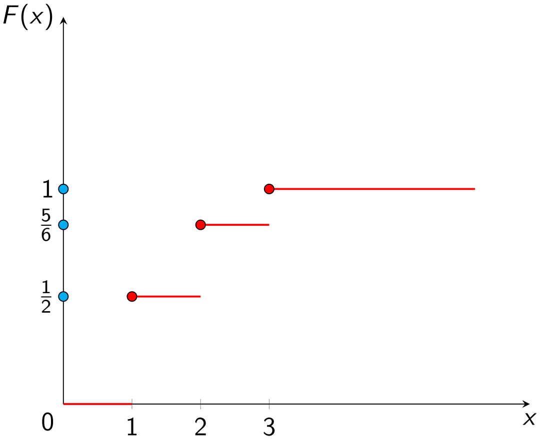
Cho X là biến ngẫu nhiên có hàm khối lượng

$$p(1) = \frac{1}{2}, \qquad p(2) = \frac{1}{3}, \qquad p(3) = \frac{1}{6}.$$

Hàm phân bố chồng chất của X được cho bởi

$$F(a) = \begin{cases} 0, & a < 1, \\ \frac{1}{2}, & 1 \leq a < 2, \\ \frac{5}{6}, & 2 \leq a < 3, \\ 1, & 3 \leq a. \end{cases}$$

Đồ thị của $F(a)$



Hình 1: Đồ thị của $F(a)$.

Ví dụ

Ví dụ 3

Cho X là biến ngẫu nhiên có hàm khối lượng

$$p(1) = \frac{1}{4}, \quad p(2) = \frac{1}{2}, \quad p(3) = p(4) = \frac{1}{8}.$$

Hãy xác định hàm phân bố chồng chất $F(a)$ của X?

Biến ngẫu nhiên Bernoulli

Định nghĩa 4

Cho $\mathcal{C}$ là một phép thử có hai kết quả "thành công" hoặc "thất bại." Nếu chúng ta gán cho X giá trị 1 trong trường hợp kết quả của phép thử là thành công, và giá trị 0 trong trường hợp còn lại, thì hàm khối lượng của X được cho bởi

$$p(0) = \mathrm{P}(X = 0) = 1 - p,$$
$$p(1) = \mathrm{P}(X = 1) = p,$$

trong đó $0 \leq p \leq 1$. Biến ngẫu nhiên X định nghĩa như trên được gọi là biến ngẫu nhiên Bernoulli với tham số p.

Biến ngẫu nhiên Nhị thức

Định nghĩa 5

Giả sử n phép thử Bernoulli với cùng tham số p được thực hiện độc lập. Gọi X là số lượng các phép thử cho kết quả thành công. Khi đó, X được gọi là một biến nhị thức (binomial) với tham số (n, p).

Biến ngẫu nhiên Nhị thức

Định nghĩa 5

Giả sử n phép thử Bernoulli với cùng tham số p được thực hiện độc lập. Gọi X là số lượng các phép thử cho kết quả thành công. Khi đó, X được gọi là một biến nhị thức (binomial) với tham số (n, p).

Hàm khối lượng của X được cho bởi

$$p(i) = \binom{n}{i} p^i (1-p)^{n-i}, \qquad \text{với } i = 0, 1, \ldots, n.$$

Biến ngẫu nhiên Nhị thức

Định nghĩa 5

Giả sử n phép thử Bernoulli với cùng tham số p được thực hiện độc lập. Gọi X là số lượng các phép thử cho kết quả thành công. Khi đó, X được gọi là một biến nhị thức *(binomial) với tham số (n, p).*

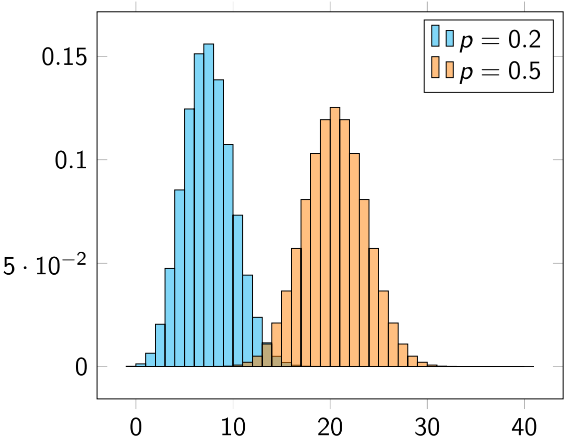
Hàm khối lượng của X được cho bởi

$$p(i) = \binom{n}{i} p^i (1-p)^{n-i}, \qquad \text{với } i = 0, 1, \ldots, n.$$

Chú ý

Nếu X là một biến ngẫu nhiên nhị thức với tham số (n, p), chúng ta nói X có phân bố nhị thức *với tham số (n, p).*

Đồ thị



Hình 2: Biến Nhị thức với $n = 40$, $p = 0.2$, và $p = 0.5$.

Ví dụ

Ví dụ 4

Giả sử số lượng người thuận tay trái chiếm 13%. Chọn ngẫu nhiên một nhóm 5 người, hãy tính xác suất để:

(α) *Người thuận tay trái là người thứ năm (những người trước đó đều thuận tay phải).*

(β) *Có đúng 3 người thuận tay trái trong nhóm đã chọn.*

(γ) *Có một số người thuận tay trái trong nhóm đã chọn.*

(δ) *Có không quá 3 người thuận tay trái trong nhóm đã chọn.*

Biến ngẫu nhiên Poisson

Định nghĩa 6

Cho X là một biến ngẫu nhiên nhận các giá trị $0, 1, \ldots$. Khi đó, X được gọi là biến ngẫu nhiên Poisson với tham số $\lambda > 0$, nếu

$$p(i) = \mathrm{P}(X = i) = e^{-\lambda}\frac{\lambda^i}{i!}, \qquad i = 0, 1, \ldots$$

Biến ngẫu nhiên Poisson

Định nghĩa 6

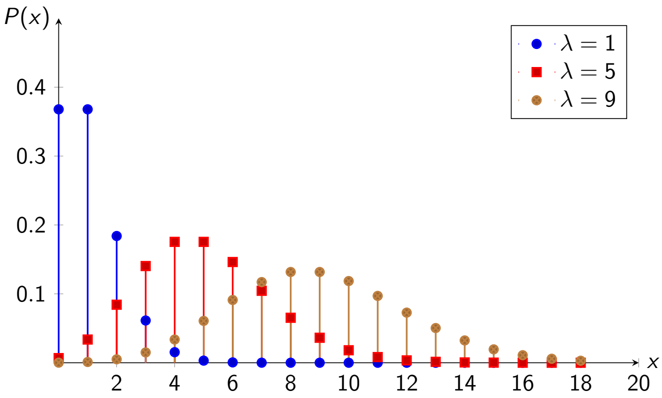
Cho X là một biến ngẫu nhiên nhận các giá trị $0, 1, \ldots$. Khi đó, X được gọi là biến ngẫu nhiên Poisson *với tham số $\lambda > 0$, nếu*

$$p(i) = \mathrm{P}(X = i) = e^{-\lambda}\frac{\lambda^i}{i!}, \qquad i = 0, 1, \ldots$$

Một số biến ngẫu nhiên có phân bố Poisson

- Số lượng lỗi in trong một trang sách của một cuốn sách.
- Số lượng khách hàng vào cửa hàng trong một ngày.
- Số lượng lỗi của một phần mềm.
- Số lượng cơn bão trong một năm.

Đồ thị



Hình 3: Phân bố Poisson với các tham số khác nhau.

Quiz

Quiz

Liệu

$$e^{-\lambda} + e^{-\lambda}\frac{\lambda}{1} + e^{-\lambda}\frac{\lambda^2}{1 \cdot 2} + e^{-\lambda}\frac{\lambda^3}{1 \cdot 2 \cdot 3} + \cdots \stackrel{?}{=} 1.$$

Ví dụ

Ví dụ 5

Khách hàng của một nhà cung cấp dịch vụ internet mở tài khoản với số lượng trung bình 10 *tài khoản mỗi ngày. Xác suất để hôm nay có nhiều hơn* 8 *khách hàng mở tài khoản là bao nhiêu?*

Ví dụ

Ví dụ 5

Khách hàng của một nhà cung cấp dịch vụ internet mở tài khoản với số lượng trung bình 10 tài khoản mỗi ngày. Xác suất để hôm nay có nhiều hơn 8 khách hàng mở tài khoản là bao nhiêu?

Số khách hàng mở tài khoản hôm nay, X, có phân bố Poisson với tham số $\lambda = 10$. Do đó,

$$\begin{aligned} \mathrm{P}(X > 8) &= 1 - F_X(8) \\ &= 1 - \sum_{n=0}^{8} e^{-10} \frac{\lambda^n}{n!} \\ &= 0.667. \end{aligned}$$

Ví dụ (tiếp)

Ví dụ 6

Xét một thí nghiệm trong đó người ta đếm số hạt α phát ra trong một giây bởi một gam chất phóng xạ X. Từ các thí nghiệm trước đó, họ biết rằng, trung bình, X phát ra 3.2 hạt α, hãy ước lượng xác suất có không quá 2 hạt α được phát ra, biết rằng số hạt phát ra có phân bố Poisson.

Biến ngẫu nhiên Liên tục

Biến ngẫu nhiên Liên tục

Định nghĩa 7

Biến ngẫu nhiên X được gọi là liên tục nếu tồn tại một hàm không âm $f(x)$, xác định với mọi $x \in (-\infty, +\infty)$ thỏa mãn tính chất: với mọi tập số thực B

$$\mathrm{P}(X \in B) = \int_B f(x)dx.$$

Hàm $f(x)$ được gọi là hàm mật độ (probability density function - pdf) của biến ngẫu nhiên X.

Biến ngẫu nhiên Liên tục

Định nghĩa 7

Biến ngẫu nhiên X được gọi là liên tục nếu tồn tại một hàm không âm $f(x)$, xác định với mọi $x \in (-\infty, +\infty)$ thỏa mãn tính chất: với mọi tập số thực B

$$\mathrm{P}(X \in B) = \int_B f(x)dx.$$

Hàm $f(x)$ được gọi là hàm mật độ (probability density function - pdf) của biến ngẫu nhiên X.

X được gọi là liên tục nếu tập các giá trị nó có thể nhận là không đếm được.

Hàm mật độ

- $f(a) \geq 0$ với mọi $a \in \mathbb{R}$.
- Hàm phân bố $F(a)$ của X được cho bởi

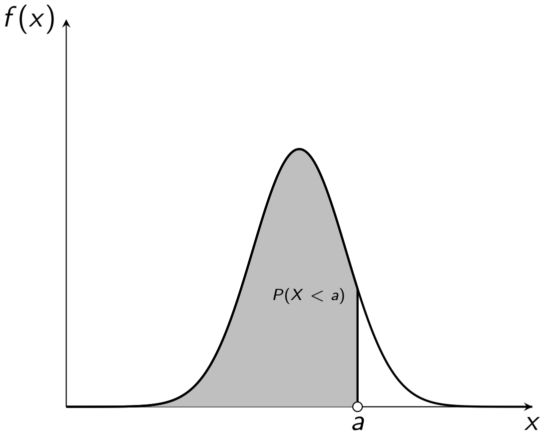
$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)dx.$$

- Hàm mật độ được suy ra từ hàm phân bố $F(x)$ bởi

$$f(x) = \frac{d}{dx}F(x).$$

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

Biểu diễn



Hình 4: $F(a)$ được biểu diễn bởi phần tô đậm.

Ví dụ

Ví dụ 7

Thời gian sử dụng, tính theo năm, của một thiết bị điện là một biến ngẫu nhiên với hàm mật độ

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

(a) *Tìm k.*

(b) *Tính xác suất để thiết bị đó dùng được lâu hơn 5 năm.*

Biến ngẫu nhiên Đều

Định nghĩa 8

X được gọi là có phân bố đều (uniformly distributed) trên khoảng (a, b), ký hiệu là $X \sim U(a, b)$, nếu hàm mật độ của nó được cho bởi

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b, \\ 0, & x \notin (a, b). \end{cases}$$

Biến ngẫu nhiên Đều

Định nghĩa 8

X được gọi là có phân bố đều (uniformly distributed) trên khoảng (a, b), ký hiệu là $X \sim U(a, b)$, nếu hàm mật độ của nó được cho bởi

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b, \\ 0, & x \notin (a, b). \end{cases}$$

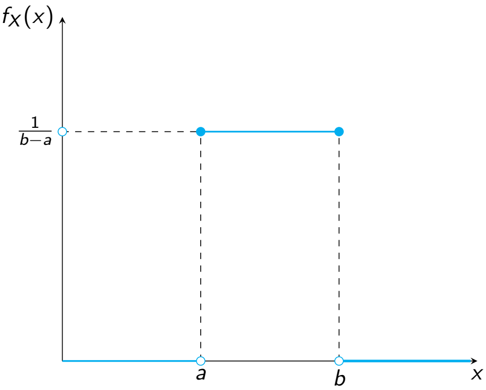
Thay a bởi 0, thay b bởi 1, ta được

Định nghĩa 9

X được gọi là có phân bố đều trên $(0, 1)$ nếu hàm mật độ của nó được cho nởi

$$f(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & x \notin (0, 1). \end{cases}$$

Đồ thị



Hình 5: Phân bố đều $U(a, b)$.

Ví dụ về phân bố đều

- Giả sử bạn đi thang máy từ Tầng 1 lên Tầng 4 mất từ 10 giây đến 1 phút. Thời gian đi thang máy có phân bố đều trong khoảng $(10, 60)$.
- Mỗi khi có ca sinh mới, thời gian sinh được ghi lại. Thời gian sinh có phân bố đều trong khoảng từ 0h:00 đến 23h:59'.

Ví dụ

Ví dụ 8

Cho X *có phân bố đều trên* $(0, 10)$. *Tính*

(a) $P(X < 3)$.

(b) $P(X > 7)$.

(c) $P(1 < X < 6)$.

Ví dụ

Ví dụ 8

Cho X *có phân bố đều trên* $(0, 10)$. *Tính*

(a) $P(X < 3)$.

(b) $P(X > 7)$.

(c) $P(1 < X < 6)$.

Ví dụ 9

Tính hàm phân bố của $X \sim U(a, b)$.

Biến ngẫu nhiên Chuẩn

Định nghĩa 10

X được gọi là có phân bố chuẩn (normally distributed) với tham số μ và σ^2 nếu hàm mật độ của nó được cho bởi

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad \text{với } -\infty \leq x \leq \infty.$$

Biến ngẫu nhiên Chuẩn

Định nghĩa 10

X được gọi là có phân bố chuẩn (normally distributed) với tham số μ và σ^2 nếu hàm mật độ của nó được cho bởi

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad \text{với } -\infty \leq x \leq \infty.$$

Đặc biệt, nếu $\mu = 0, \sigma^2 = 1$, thì X được gọi là có phân bố chuẩn tắc.

Ký hiệu của phân bố chuẩn: $X \sim \mathcal{N}(\mu, \sigma^2)$.

Biến ngẫu nhiên Chuẩn

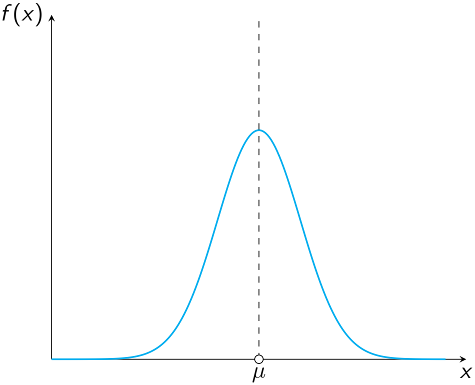
Định nghĩa 10

X được gọi là có phân bố chuẩn (normally distributed) với tham số μ và σ^2 nếu hàm mật độ của nó được cho bởi

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad \text{với } -\infty \leq x \leq \infty.$$

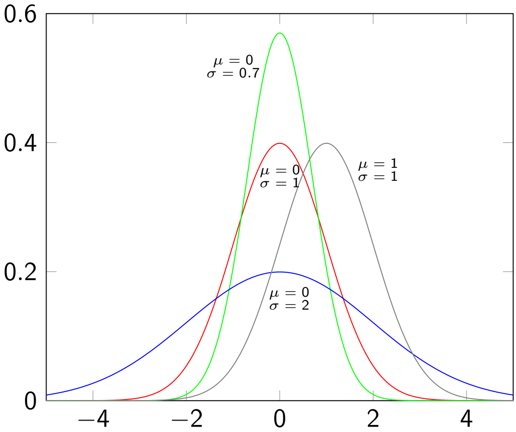
Đặc biệt, nếu $\mu = 0, \sigma^2 = 1$, thì X được gọi là có phân bố chuẩn tắc.
Ký hiệu của phân bố chuẩn: $X \sim \mathcal{N}(\mu, \sigma^2)$.
Phân bố chuẩn còn được gọi là phân bố Gauß.

Đồ thị



Hình 6: Phân bố chuẩn.

Đồ thị (tiếp)



Hình 7: Phân bố chuẩn với các tham số khác nhau.

Ví dụ về phân bố chuẩn

- Chiều cao người trưởng thành trong một vùng.
- Số đo huyết áp của người dân khỏe mạnh trong một vùng.
- Chỉ số IQ. Chỉ số trung bình là 100. Phần lớn mọi người có chỉ số dao động quanh 100, chỉ có số ít có chỉ số rất cao hoặc rất thấp.

Chuẩn tắc hóa Biến ngẫu nhiên Chuẩn

Mệnh đề 1

Nếu X có phân bố $\mathcal{N}(\mu, \sigma^2)$, thì biến ngẫu nhiên

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

có phân bố chuẩn tắc, i.e., $Z \sim \mathcal{N}(0, 1)$.

Chuẩn tắc hóa Biến ngẫu nhiên Chuẩn

Mệnh đề 1

Nếu X có phân bố $\mathcal{N}(\mu, \sigma^2)$, thì biến ngẫu nhiên

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

có phân bố chuẩn tắc, i.e., $Z \sim \mathcal{N}(0, 1)$.

$$\begin{aligned}
P(X < a) &= \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\
&= \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} d\left(\frac{x-\mu}{\sigma}\right) \\
&= \int_{-\infty}^{\frac{a-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = P\left(Z < (a-\mu)/\sigma\right).
\end{aligned}$$

Biến ngẫu nhiên Chuẩn tắc

- Chữ Z thường được dùng để ký hiệu biến ngẫu nhiên chuẩn tắc, *i.e.*, $Z \sim \mathcal{N}(0,1)$.
- Hàm phân bố được ký hiệu bởi $\Phi(x)$, *i.e.*,

$$\Phi(x) = \mathrm{P}(Z < x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz.$$

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

Bảng giá trị của $\Phi(x)$

x	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441

Bảng 1: Một số giá trị của $\Phi(x)$.

Ví dụ

Ví dụ 10

Từ Bảng giá trị của $\Phi(x)$, *ta tính được*

- $\mathrm{P}(Z < 1.35) = \Phi(1.35) = 0.9115$.
- $\mathrm{P}(-0.77 < Z < 1.35) = \Phi(1.35) - \Phi(-0.77) = 0.9115 - 0.2206 = 0.6909$.
- *Nếu* $\mathrm{P}(Z < z_0) = 0.9441$, *thì* $z_0 \approx 1.59$.

Áp dụng

Bài tập

Giả sử thu nhập của người sống ở Hà Nội có phân bố chuẩn với $\mu = 9$ triệu/tháng và $\sigma = 2$ triệu. Tính tỉ lệ tầng lớp trung lưu, i.e., những người có thu nhập trong khoảng từ 6 đến 12 triệu/tháng.

Kỳ vọng của một Biến ngẫu nhiên

Kỳ vọng của Biến Rời rạc

Định nghĩa 11

Cho X là một biến ngẫu nhiên rời rạc với hàm khối lượng $p(x)$. Khi đó, giá trị kỳ vọng (expected value) của X, ký hiệu là $E[X]$, được cho bởi

$$E[X] = x_1 p(x_1) + x_2 p(x_2) + \cdots ,$$

trong đó $x_1, x_2, \ldots$ là tất cả các giá trị mà X có thể nhận.

Kỳ vọng của Biến Rời rạc

Định nghĩa 11

Cho X là một biến ngẫu nhiên rời rạc với hàm khối lượng $p(x)$. Khi đó, giá trị kỳ vọng (expected value) của X, ký hiệu là $E[X]$, được cho bởi

$$E[X] = x_1 p(x_1) + x_2 p(x_2) + \cdots ,$$

trong đó $x_1, x_2, \ldots$ là tất cả các giá trị mà X có thể nhận.

Giá trị kỳ vọng còn được gọi là giá trị trung bình (mean value).

Kỳ vọng của Biến Rời rạc

Định nghĩa 11

Cho X là một biến ngẫu nhiên rời rạc với hàm khối lượng $p(x)$. Khi đó, giá trị kỳ vọng (expected value) của X, ký hiệu là $E[X]$, được cho bởi

$$E[X] = x_1 p(x_1) + x_2 p(x_2) + \cdots ,$$

trong đó $x_1, x_2, \ldots$ là tất cả các giá trị mà X có thể nhận.

Giá trị kỳ vọng còn được gọi là giá trị trung bình (mean value). Chữ cái Hy-Lạp μ thường được dùng để ký hiệu giá trị kỳ vọng.

Ví dụ

Ví dụ 11

Cho X là biến ngẫu nhiên có hàm khối lượng

$$p(1) = \frac{1}{2}, \quad p(2) = \frac{1}{3}, \quad p(3) = \frac{1}{6}.$$

Khi đó,

$$\begin{aligned} E[X] &= 1 \cdot p(1) + 2 \cdot p(2) + 3 \cdot p(3) \\ &= 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{6} \\ &= \frac{5}{3}. \end{aligned}$$

Ví dụ (tiếp)

Ví dụ 12

Cho X có phân bố Bernoulli với tham số p, i.e.,

$$p(0) = 1 - p, \quad p(1) = p.$$

Khi đó,

$$\begin{aligned} E[X] &= 0 \cdot (1 - p) + 1 \cdot p \\ &= p. \end{aligned}$$

Áp dụng

Ví dụ 13

Cho X là biến ngẫu nhiên có hàm khối lượng

$$p(1) = \frac{1}{4}, \quad p(2) = \frac{1}{2}, \quad p(3) = p(4) = \frac{1}{8}.$$

Tính $E[X]$.

Áp dụng

Ví dụ 13

Cho X là biến ngẫu nhiên có hàm khối lượng

$$p(1) = \frac{1}{4}, \quad p(2) = \frac{1}{2}, \quad p(3) = p(4) = \frac{1}{8}.$$

Tính $E[X]$.

Ví dụ 14

(a) *Cho X là biến ngẫu nhiên có phân bố nhị thức với tham số (n, p). Tính $E[X]$?*

(b) *Cho Y là biến ngẫu nhiên có phân bố Poisson với tham số λ. Tính $E[Y]$?*

Kỳ vọng của Biến Liên tục

Định nghĩa 12

Cho X là một biến liên tục với hàm mật độ $f(x)$. Khi đó, giá trị kỳ vọng của X được cho bởi

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx.$$

Ví dụ

Ví dụ 15

Cho X là biến ngẫu nhiên có phân bố đều trên (a, b). Tính $E[X]$?

Ví dụ

Ví dụ 15

Cho X là biến ngẫu nhiên có phân bố đều trên (a, b). Tính $E[X]$?

- X có hàm phân bố

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in (a, b), \\ 0, & x \notin (a, b). \end{cases}$$

- Theo định nghĩa,

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_a^b x \cdot \frac{1}{b-a}dx = \frac{a+b}{2}.$$

Áp dụng

Bài tập

Cho X là biến ngẫu nhiên có hàm mật độ $f(x) = 2x^{-3}, x \geq 1$. Hãy tính $E[X]$?

Áp dụng

Bài tập

Cho X là biến ngẫu nhiên có hàm mật độ $f(x) = 2x^{-3}, x \geq 1$. Hãy tính $E[X]$?

Bài tập

Cho $Z \sim \mathcal{N}(0,1)$. Tính $E[Z]$?

Mở rộng

Bài toán

Cho X là một biến ngẫu nhiên, $g(\cdot)$ là một hàm thực. Tính $E[g(X)]$?

Mở rộng

Bài toán

Cho X là một biến ngẫu nhiên, $g(\cdot)$ là một hàm thực. Tính $E[g(X)]$?

Lời giải trực giác: Dùng định nghĩa.

Mở rộng

Bài toán

Cho X là một biến ngẫu nhiên, $g(\cdot)$ là một hàm thực. Tính $E[g(X)]$?

Lời giải trực giác: Dùng định nghĩa.

- **Bước 1.** Xác định hàm khối lượng (hay hàm mật độ) của biến ngẫu nhiên $g(X)$ nếu biến X rời rạc (hay liên tục).
- **Bước 2.** Tùy thuộc vào tính rời rạc hay liên tục của $g(X)$, sử dụng các định nghĩa kỳ vọng tương ứng để tính $E[g(X)]$.

Trường hợp Rời rạc

Ví dụ 16

Cho X là biến ngẫu nhiên có hàm khối lượng xác định bởi:

$$p(0) = 0.2, \qquad p(1) = 0.5, \qquad p(2) = 0.3.$$

Tính $E[X], E[X^2]$?

Trường hợp Rời rạc

Ví dụ 16

Cho X là biến ngẫu nhiên có hàm khối lượng xác định bởi:

$$p(0) = 0.2, \qquad p(1) = 0.5, \qquad p(2) = 0.3.$$

Tính $E[X], E[X^2]$?

- Đặt $Y = X^2$. Do $X \in \{0, 1, 2\}$, nên $Y \in \{0, 1, 4\}$

$$p_Y(0) = \mathrm{P}(Y = 0^2) = \mathrm{P}(X = 0) = 0.2,$$
$$p_Y(1) = \mathrm{P}(Y = 1^2) = \mathrm{P}(X = 1) = 0.5,$$
$$p_Y(4) = \mathrm{P}(Y = 2^2) = \mathrm{P}(X = 2) = 0.3.$$

- Dùng định nghĩa để tìm kỳ vọng. Ta có

$$E[X^2] = E[Y] = 0 \cdot 0.2 + 1 \cdot 0.5 + 4 \cdot 0.3 = 1.7.$$

Trường hợp Liên tục

Ví dụ 17

Cho X là biến ngẫu nhiên có phân bố đều trên $(0, 1)$. Tính $E[X^3]$.

Trường hợp Liên tục

Ví dụ 17

Cho X là biến ngẫu nhiên có phân bố đều trên $(0, 1)$. Tính $E[X^3]$.

- **Bước 1.** Tìm hàm mật độ $f_Y(a)$ của biến ngẫu nhiên $Y = X^3$.

$$F_Y(a) = \mathrm{P}(Y \leq a) = a^{\frac{1}{3}}.$$

Lấy đạo hàm $F_Y(a)$ theo a, ta được $f_Y(a) = \frac{a^{-\frac{2}{3}}}{3}, 0 \leq a \leq 1$.

- **Bước 2.** Dùng định nghĩa để tính kỳ vọng.

$$E[X^3] = E[Y] = \int_{-\infty}^{\infty} a f_Y(a) da = \int_0^1 a \cdot \frac{a^{-\frac{2}{3}}}{3} da$$
$$= \frac{1}{4}.$$

Mệnh đề

Mệnh đề 2

(*i*) *Nếu X là một biến ngẫu nhiên rời rạc, có hàm khối lượng $p(x)$, thì với mọi hàm thực $g(x)$,*

$$E[g(X)] = g(x_1)p(x_1) + g(x_2)p(x_2) + \cdots ,$$

trong đó $x_1, x_2, \ldots$ là các giá trị mà X có thể nhận.

(*ii*) *Nếu X là một biến ngẫu nhiên liên tục, có hàm mật độ $f(x)$, thì với mọi hàm thực $g(x)$,*

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

Áp dụng vào các ví dụ trước

- Tính $E[X^2]$ trong Ví dụ 16:

$$E[X^2] = 0^2 \cdot 0.2 + 1^2 \cdot 0.5 + 2^2 \cdot 0.3 = 1.7.$$

- Tính $E[X^3]$ trong Ví dụ 17:

$$E[X^3] = \int_0^1 x^3 \cdot 1dx = \frac{1}{4}.$$

Hệ quả

Hệ quả

Nếu a, b là hai hằng số, thì

$$E[aX + b] = aE[X] + b.$$

Hệ quả

Hệ quả

Nếu a, b là hai hằng số, thì

$$E[aX + b] = aE[X] + b.$$

Đặc biệt hóa,

- Cho $b = 0$, ta được $E[aX] = aE[X]$.
- Cho $a = 0$, ta được $E[b] = b$.
- Từ hai điều trên suy ra $E\big[E[X]\big] = E[X]$.

Phương sai & Độ lệch chuẩn

Định nghĩa 13 (Phương sai - Variance)

Cho X là một biến ngẫu nhiên có kỳ vọng μ. Phương sai của X, ký hiệu là $\text{Var}(X)$, *được cho bởi*

$$\text{Var}(X) = E\left[(X-\mu)^2\right].$$

Phương sai & Độ lệch chuẩn

Định nghĩa 13 (Phương sai - Variance)

Cho X là một biến ngẫu nhiên có kỳ vọng μ. Phương sai của X, ký hiệu là $\mathrm{Var}(X)$, được cho bởi

$$\mathrm{Var}(X) = E\left[(X-\mu)^2\right].$$

Giá trị của phương sai thường được ký hiệu bởi σ^2.

Phương sai & Độ lệch chuẩn

Định nghĩa 13 (Phương sai - Variance)

Cho X là một biến ngẫu nhiên có kỳ vọng μ. Phương sai của X, ký hiệu là $\mathrm{Var}(X)$, *được cho bởi*

$$\mathrm{Var}(X) = E\left[(X-\mu)^2\right].$$

Giá trị của phương sai thường được ký hiệu bởi σ^2.

Định nghĩa 14 (Độ lệch chuẩn - Standard deviation)

Độ lệch chuẩn của một biến ngẫu nhiên X, ký hiệu là Std, *được định nghĩa bởi* $\mathrm{Std} = \sqrt{\mathrm{Var}(X)}$.

Phương sai & Độ lệch chuẩn

Định nghĩa 13 (Phương sai - Variance)

Cho X là một biến ngẫu nhiên có kỳ vọng μ. Phương sai của X, ký hiệu là $\mathrm{Var}(X)$, *được cho bởi*

$$\mathrm{Var}(X) = E\left[(X-\mu)^2\right].$$

Giá trị của phương sai thường được ký hiệu bởi σ^2.

Định nghĩa 14 (Độ lệch chuẩn - Standard deviation)

Độ lệch chuẩn của một biến ngẫu nhiên X, ký hiệu là Std, *được định nghĩa bởi* $\mathrm{Std} = \sqrt{\mathrm{Var}(X)}$.

Giá trị của độ lệch chuẩn thường được ký hiệu bởi σ.

Ví dụ minh họa

- Trong Ví dụ 16,

$$
\begin{aligned}
\mu = E[X] &= 0 \cdot 0.2 + 1 \cdot 0.5 + 2 \cdot 0.3 = 1.1, \\
\sigma^2 = \mathrm{Var}(X) &= E\left[(X-\mu)^2\right] \\
&= (0-1.1)^2 \cdot 0.2 + (1-1.1)^2 \cdot 0.5 + (2-1.1)^2 \cdot 0.3 \\
&= 0.49, \\
\sigma &= 0.7.
\end{aligned}
$$

Ví dụ minh họa

- Trong Ví dụ 16,

$$\begin{aligned}
\mu &= E[X] = 0 \cdot 0.2 + 1 \cdot 0.5 + 2 \cdot 0.3 = 1.1, \\
\sigma^2 = \mathrm{Var}(X) &= E\left[(X-\mu)^2\right] \\
&= (0-1.1)^2 \cdot 0.2 + (1-1.1)^2 \cdot 0.5 + (2-1.1)^2 \cdot 0.3 \\
&= 0.49, \\
\sigma &= 0.7.
\end{aligned}$$

- Trong Ví dụ 15, biến $X \sim U(a,b)$, $\quad \mu = \frac{a+b}{2}$,

$$\begin{aligned}
\sigma^2 &= E\left[(X-\mu)^2\right] \\
&= \int_a^b \left(x - \frac{a+b}{2}\right)^2 \cdot \frac{1}{b-a} dx \\
&= \frac{(b-a)^2}{12}.
\end{aligned}$$

Ví dụ (tiếp)

Ví dụ 18

Một biến ngẫu nhiên X *nhận các giá trị* $-0.5, 0, 1$ *với xác suất*

$$\mathrm{P}(X = -0.5) = \frac{1}{2}, \quad \mathrm{P}(X = 0) = \mathrm{P}(X = 1) = \frac{1}{4}.$$

Tính $\mathrm{Var}(X)$.

Ví dụ (tiếp)

Ví dụ 18

Một biến ngẫu nhiên X nhận các giá trị $-0.5, 0, 1$ với xác suất

$$\mathrm{P}(X = -0.5) = \frac{1}{2}, \quad \mathrm{P}(X = 0) = \mathrm{P}(X = 1) = \frac{1}{4}.$$

Tính $\mathrm{Var}(X)$.

Ví dụ 19

Tung ba đồng xu cân, đồng chất. Gọi Y là số mặt ngửa.

(a) *Xác định phân bố xác suất của Y.*

(b) *Tính $E[Y]$.*

(c) *Tính* $\mathrm{Var}(Y)$.

Tính chất

Mệnh đề 3

Cho X là một biến ngẫu nhiên có giá trị kỳ vọng μ, a, b là các hằng số. Khi đó,

(i) $\mathrm{Var}(X) = E[X^2] - \left(E[X]\right)^2 = E[X^2] - \mu^2$.

(ii) $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$.

Tính chất

Mệnh đề 3

Cho X là một biến ngẫu nhiên có giá trị kỳ vọng μ, a, b là các hằng số. Khi đó,

(*i*) $\operatorname{Var}(X) = E[X^2] - \left(E[X]\right)^2 = E[X^2] - \mu^2$.

(*ii*) $\operatorname{Var}(aX + b) = a^2 \operatorname{Var}(X)$.

Đặc biệt hóa,

- Cho $a = 0$, ta được $\operatorname{Var}(b) = 0$.
- Cho $b = 0$, ta được $\operatorname{Var}(aX) = a^2 \operatorname{Var}(X)$.

Ví dụ

- Trong Ví dụ 16, ta có

$$E[X] = 1.1,$$
$$E[X^2] = 1.7.$$

Do đó, $\mathrm{Var}(X) = E[X^2] - \big(E[X]\big)^2 = 0.49.$

Ví dụ

- Trong Ví dụ 16, ta có

$$E[X] = 1.1,$$
$$E[X^2] = 1.7.$$

Do đó, $\mathrm{Var}(X) = E[X^2] - \left(E[X]\right)^2 = 0.49.$

- Trong Ví dụ 15, có $\mu = \frac{a+b}{2}$, và

$$E[X^2] = \int_a^b x^2 \cdot \frac{1}{b-a} dx = \frac{a^2+ab+b^2}{3}.$$

Do đó,

$$\mathrm{Var}(X) = E[X^2] - \left(E[X]\right)^2 = \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

Áp dụng

Ví dụ 20

Cho biến ngẫu nhiên X có hàm mật độ

$$f(x) = 1 - \frac{x}{2}, \quad 0 < x < 2.$$

Tính $E[X], \mathrm{Var}(X)$.

Áp dụng (tiếp)

Bài tập

Tính phương sai của các biến ngẫu nhiên sau:

(i) *Biến Bernoulli với tham số p.*

(ii) *Biến nhị thức X với tham số (n, p).*

(iii) *Biến Poisson Y với tham số λ.*

(iv) *Biến chuẩn tắc Z.*

To be continued...