Chương 3: Biến ngẫu nhiên

Phan Quang Sáng

Bộ môn Toán- Khoa CNTT- VNUA

Hà Nội, Ngày 19 tháng 10 năm 2018

http://fita.vnua.edu.vn/vi/pqsang/

pqsang@vnua.edu.vn

Nội dung chính

1. Các khái niệm
   - 1.1. Định nghĩa biến ngẫu nhiên
   - 1.2. Bảng phân phối của biến ngẫu nhiên rời rạc
   - 1.3. Hàm phân phối xác suất
   - 1.4. Hàm mật độ xác suất
2. Các số đặc trưng của biến ngẫu nhiên
   - 2.1. Kỳ vọng
   - 2.2. Phương sai
   - 2.3. Độ lệch chuẩn
   - 2.4. Mode
3. Một số quy luật xác suất thường gặp
   - 3.1. Phân phối nhị thức
   - 3.2. Phân phối Poisson
   - 3.3. Phân phối chuẩn
4. Xấp xỷ phân phối nhị thức

Khi cho tương ứng các kết quả của một phép thử với các số thực ta được một biến ngẫu nhiễn.

Ví dụ 1: Gọi X là số chấm xuất hiện khi tung 1 con xúc xắc. Khi đó X là 1 Bnn nhận các giá trị $0, 1, 2, 3, 4, 5, 6$.

Ví dụ 2: Tung 1 đồng xu, tương ứng mặt sấp ta quy ước X nhận giá trị 1, mặt ngửa nhận giá trị 0.

Ví dụ 3: Một xạ thủ bắn vào bia cho tới khi nào trúng thì dừng. Gọi Y là số viên đạn cần dùng, khi đó Y là một Bnn nhận giá trị đếm được vô hạn.

Ví dụ 4: Gọi Z là nhiệt độ khi đun 1 ấm nước. Khi đó Z là một Bnn nhận giá trị liên tục.

Định nghĩa

Bảng pp xs của 1 Bnn gồm 2 dòng: 1 dòng ghi các giá trị, một dòng ghi xác suất tương ứng

X	x_1	x_2	...	x_k	...	x_n	...
P	p_1	p_2	...	p_k	...	p_n	...

Ví dụ

Chú ý: tổng Xs phải bằng 1: $\Sigma p_k = 1$

Ví dụ: đặt câu hỏi tỷ lệ thí sinh trong một kỳ thi dưới 4 điểm, 5 điểm, 7 điểm, 8 điểm... là bao nhiêu?

Định nghĩa

Hàm pp xs của Bnn X là hàm số $F(x)$ cho bởi
$F(x) = P(X < x)$, $x \in \mathbb{R}$.

Ví dụ

Một số tính chất của hàm phân phối:

1) $0 \leq F(x) \leq 1$
2) $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to +\infty} F(x) = 1$
3) $F(x)$ là hàm không giảm
4) $P(a \leq X < b) = F(b) - F(a)$
5) $F(x)$ là hàm liên tục trái.

Định nghĩa

Nếu hàm pp xs $F(x)$ của Bnn X có đạo hàm với mọi $x \in \mathbb{R}$ thì hàm đạo hàm $f(x) = F'(x)$ được gọi là hàm mật độ xác suất của X.

Chú ý: Bnn rời rạc không có hàm mật độ Xs

Một số tính chất của hàm mật độ:

1) $f(x) \geq 0$ với mọi $x \in \mathbb{R}$
2) $F(x) = \int_{-\infty}^{x} f(t)dt$
3) $P(a \leq X < b) = \int_a^b f(x)dx$
4) Nếu X có hàm mật độ thì $P(X = a) = 0$
5) $\int_{-\infty}^{+\infty} f(x)dx = 1$

Định nghĩa

Kỳ vọng (hay trung bình) của Bnn X, ký hiệu là $E(X)$ hoặc $M(X)$, định nghĩa bởi,

1. Nếu X là Bnn rời rạc thì
$$E(X) = \sum x_k p_k,$$
và nếu X là vô hạn thì thêm điều kiện $\sum |x_k p_k|$ hội tụ.
2. Nếu X là Bnn liên tục có hàm mật độ $f(x)$ thì
$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx,$$
với Đk $\int_{-\infty}^{+\infty} |xf(x)| dx$ hội tụ.

Một số tính chất của kỳ vọng:

1) $E(C) = C$, với C là một hằng số
2) $E(kX) = kE(X)$, với $k \in \mathbb{R}$
3) $E(X + Y) = E(X) + E(Y)$
4) Nếu X, Y là các Bnn độc lập thì $E(XY) = E(X)E(Y)$
5) Nếu $Y = \varphi(X)$ thì $E(Y) = \int_{-\infty}^{+\infty} \phi(x)f(x)dx$

Định nghĩa

Giả sử X là bnn có kỳ vọng $E(X)$. Khi đó phương sai của X là số

$$D(X) = E[X - E(X)]^2.$$

Ta cũng có

$$D(X) = E(X^2) - (E(X))^2.$$

Ví dụ...

Tính chất của phương sai:

1) $D(C) = 0$, với C là một hằng số
2) $D(kX) = k^2 D(X)$, với $k \in \mathbb{R}$
3) Nếu X, Y là các Bnn độc lập thì

$$D(X+Y) = D(X) + D(Y)$$

Định nghĩa

Độ lệch chuẩn của Bnn X là số

$$\sigma(X) = \sqrt{D(X)}.$$

Chú ý: $\sigma(X)$ có đơn vị như đơn vị của X.

Định nghĩa

Mod(X)

1. Nếu X là Bnn rời rạc thì $Mod(X)$ là điểm x_k tại đó Xs $P(X = x_i)$ lớn nhất.
2. Nếu X là Bnn liên tục thì $Mod(X)$ là điểm tại đó hàm mật độ $f(x)$ đạt giá trị lớn nhất.

Định nghĩa

Bnn X được gọi là có phân phối nhị thức, ký hiệu $X \sim B(n, p)$, nếu nó có bảng pp Xs

X	0	1	...	k	...	n
P	$C_n^0 p^0 q^n$	$C_n^1 pq^{n-1}$	...	$C_n^k p^k q^{n-k}$	...	$C_n^n p^n q^0$

với $0 < p < 1$, $q = 1 - p$.

Mệnh đề

Nếu $X \sim B(n, p)$ thì

$$E(X) = np, \ \ D(X) = npq$$

Định nghĩa

Bnn X được gọi là có phân phối Poisson với tham số thực $\lambda > 0$, ký hiệu $X \sim P_\lambda$, nếu nó có bảng pp Xs

X	0	1	...	k	...	n	...
P	$e^{-\lambda}$	$e^{-\lambda}\frac{\lambda}{1!}$	...	$e^{-\lambda}\frac{\lambda^k}{k!}$	...	$e^{-\lambda}\frac{\lambda^n}{n!}$	...

Định nghĩa

Bnn X được gọi là có phân phối chuẩn với hai tham số μ và σ^2 $(\sigma > 0)$, ký hiệu $X \sim N(\mu, \sigma^2)$, nếu nó hàm mật độ Xs

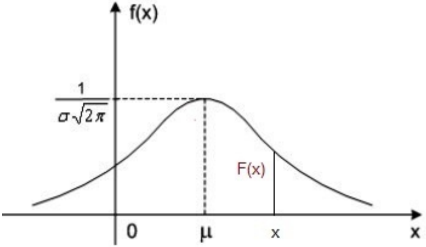
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

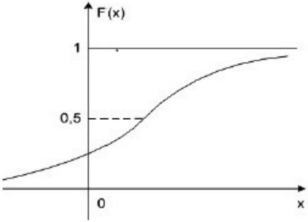
Mệnh đề

Khi đó

$$E(X) = \mu, \ D(X) = \sigma^2$$

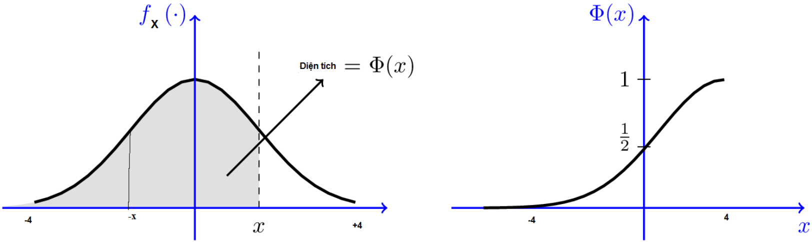
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, \ F(x) = \int_{-\infty}^{x} f(t)dt$$





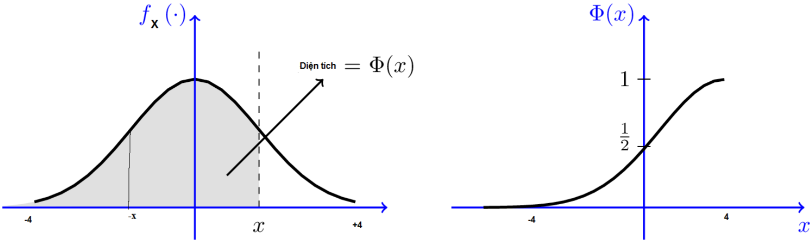
Hàm mật độ và hàm phân phối của pp chuẩn

Phân phối chuẩn tắc: $X \sim N(0,1)$ ($\mu = 0$ và $\sigma = 1$), khi đó viết $f(x) = \varphi(x)$ (hàm Gauss) và $F(x) = \Phi(x)$ (hàm Laplace).



Hàm mật độ và hàm phân phối của pp chuẩn tắc

Phân phối chuẩn tắc: $X \sim N(0,1)$ ($\mu = 0$ và $\sigma = 1$), khi đó viết $f(x) = \varphi(x)$ (hàm Gauss) và $F(x) = \Phi(x)$ (hàm Laplace).



Hàm mật độ và hàm phân phối của pp chuẩn tắc

Một số tính chất:

1. Hàm φ là hàm chẵn $\varphi(-x) = \varphi(x)$, đơn điệu giảm trên $(0, +\infty)$.
2. Nếu $x \geq 4$ hoặc $x \leq -4$ thì $\varphi \approx 0$.
3. $\Phi(-x) = 1 - \Phi(x)$
4. Nếu $x \geq 4$ thì $\Phi \approx 1$; Nếu $x \leq -4$ thì $\Phi \approx 0$.

Áp dụng tính Xs với phân phối chuẩn

Giả sử $X \sim N(\mu, \sigma^2)$. Chú ý rằng ta có thể viết

$$f(x) = \frac{1}{\sigma}\varphi(\frac{x-\mu}{\sigma}),$$

$$F(x) = \Phi(\frac{x-\mu}{\sigma}).$$

Do đó:

$$P(X \leq b) = F(b) = \Phi(\frac{b-\mu}{\sigma})$$

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

Ví dụ: Giả sử $X \sim N(3, 4)$. Tính $P(X \leq 5)$ và $P(2 \leq X \leq 4)$.

Nhắc lại: giả sử $X \sim B(n, p)$, với $0 < p < 1$, $q = 1 - p$.

X	0	1	...	k	...	n
P	$C_n^0 p^0 q^n$	$C_n^1 p q^{n-1}$	...	$C_n^k p^k q^{n-k}$	...	$C_n^n p^n q^0$

Nhắc lại: giả sử $X \sim B(n, p)$, với $0 < p < 1$, $q = 1 - p$.

X	0	1	...	k	...	n
P	$C_n^0 p^0 q^n$	$C_n^1 p q^{n-1}$	...	$C_n^k p^k q^{n-k}$	...	$C_n^n p^n q^0$

1) Khi n lớn,

Nhắc lại: giả sử $X \sim B(n,p)$, với $0 < p < 1$, $q = 1 - p$.

X	0	1	...	k	...	n
P	$C_n^0 p^0 q^n$	$C_n^1 p q^{n-1}$	...	$C_n^k p^k q^{n-k}$	...	$C_n^n p^n q^0$

1) Khi n lớn, $X \approx N(\mu, \sigma^2)$, với $\mu = np$, $\sigma^2 = npq$. Do đó

$$P(X = k) = \approx \frac{1}{\sigma}\varphi(\frac{x-\mu}{\sigma}),$$

$$P(X \leq k_2) \approx \Phi(\frac{k_2-\mu}{\sigma}),$$

$$P(k_1 \leq X \leq k_2) \approx \Phi(\frac{k_2-\mu}{\sigma}) - \Phi(\frac{k_1-\mu}{\sigma})$$

Nhắc lại: giả sử $X \sim B(n,p)$, với $0 < p < 1$, $q = 1 - p$.

X	0	1	...	k	...	n
P	$C_n^0 p^0 q^n$	$C_n^1 pq^{n-1}$	...	$C_n^k p^k q^{n-k}$	...	$C_n^n p^n q^0$

1) Khi n lớn, $X \approx N(\mu, \sigma^2)$, với $\mu = np$, $\sigma^2 = npq$. Do đó

$$P(X = k) =\approx \frac{1}{\sigma}\varphi(\frac{x-\mu}{\sigma}),$$

$$P(X \le k_2) \approx \Phi(\frac{k_2-\mu}{\sigma}),$$

$$P(k_1 \le X \le k_2) \approx \Phi(\frac{k_2-\mu}{\sigma}) - \Phi(\frac{k_1-\mu}{\sigma})$$

2) Nếu p gần 0 hoặc gần 1,

Nhắc lại: giả sử $X \sim B(n, p)$, với $0 < p < 1$, $q = 1 - p$.

X	0	1	...	k	...	n
P	$C_n^0 p^0 q^n$	$C_n^1 pq^{n-1}$	...	$C_n^k p^k q^{n-k}$	...	$C_n^n p^n q^0$

1) Khi n lớn, $X \approx N(\mu, \sigma^2)$, với $\mu = np$, $\sigma^2 = npq$. Do đó

$$P(X = k) =\approx \frac{1}{\sigma}\varphi(\frac{x-\mu}{\sigma}),$$

$$P(X \leq k_2) \approx \Phi(\frac{k_2-\mu}{\sigma}),$$

$$P(k_1 \leq X \leq k_2) \approx \Phi(\frac{k_2-\mu}{\sigma}) - \Phi(\frac{k_1-\mu}{\sigma})$$

2) Nếu p gần 0 hoặc gần 1, khi đó X được xấp xỷ bởi pp Poisson $X \approx P_\lambda$, với $\lambda = np$. Do đó

$$P(X = k) \approx e^{-\lambda}\frac{\lambda^k}{k!}$$

Ví dụ 1: Trồng 400 cây biết xác suất để mỗi cây sống là $0,8$.

(a) Tính xác suất có đúng 320 cây sống.

(b) Tính xác suất có từ 300 đến 360 cây sống.

Ví dụ 2: Tỉ lệ người có ký sinh trùng sốt rét trong máu của một vùng là 0,01. Lấy mẫu máu của 100 người. Tính xác suất có 5 mẫu có ký sinh trùng sốt rét.