MỘT SỐ PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

BỘ MÔN THỐNG KÊ TOÁN HỌC
KHOA TOÁN - TIN HỌC
ĐẠI HỌC KHOA HỌC TỰ NHIÊN TP.HCM

Tháng 9 năm 2021

Outline

Outline

Phân phối Bernoulli

Định nghĩa 1

Cho b.n.n X rời rạc lấy hai trị số 0, 1. Ta nói X có phân phối Bernoulli khi hàm xác suất có dạng:

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

Kí hiệu: $X \sim B(1, p)$ trong đó $p \in (0, 1)$.

Đặc trưng

Kì vọng: $EX = 0(1-p) + 1.p = p$.
Phương sai: $Var(X) = 0^2(1-p) + 1^2 p - p^2 = p(1-p)$.

Phân phối Bernoulli: Mô hình

Coi 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

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

Ta có

$$\begin{aligned} P(X = 1) &= P(\omega) = p \\ P(X = 0) &= P(\bar{\omega}) = 1 - p \end{aligned}$$

Vậy X có mật độ

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

nghĩa là X có phân phối Bernoulli.

Phân phối Bernoulli: Ví dụ

Nhận xét 2

Mọi thí nghiệm ngẫu nhiên có hai kết quả đều có phân phối Bernoulli.

Ví dụ 3

Tung đồng xu 1 lần, lưu ý mặt ngửa. Đặt

$$X = \begin{cases} 1 & \text{nếu ngửa} \\ 0 & \text{nếu sấp} \end{cases}$$

thì $X \sim B(1, 1/2)$.

Phân phối Bernoulli: Ví dụ (tt)

Ví dụ 4

Tung con xúc sắc, lưu ý mặt 6. Đặt

$$Y = \begin{cases} 1 & \text{nếu mặt 6 xuất hiện} \\ 0 & \text{nếu là mặt khác} \end{cases}$$

thì $Y \sim B(1, 1/6)$.

Ví dụ 5

Quan sát giới tính trong một lần sanh. Đặt

$$Z = \begin{cases} 1 & \text{nếu con trai} \\ 0 & \text{nếu con gái} \end{cases}$$

thì $Z \sim B(1, 1/2)$.

Outline

Phân phối nhị thức

Định nghĩa 6

Cho biến ngẫu nhiên rời rạc X nhận các giá trị $0, 1, 2, \ldots, n$. X có phân phối nhị thức, kí hiệu $X \sim B(n, p)$, khi hàm xác suất có dạng

$$f(x) = \begin{cases} C_n^x p^x (1-p)^{n-x} & \text{với } x = 0, 1, 2, \ldots, n \\ 0 & \text{nơi khác} \end{cases}$$

trong đó $0 < p < 1$.

Phân phối nhị thức: Mô hình

Coi 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 và quan tâm đến số lần xuất hiện ω_* trong n lần quan sát đó.
Không gian mẫu của n lần thí nghiệm là

$$\Omega = \{\omega = (\omega_{(1)}, \omega_{(2)}, \ldots, \omega_{(n)}) \colon \omega_{(i)} \in \{\omega_*, \bar{\omega}_*\}, i = 1, 2, \ldots, n\}$$

Đặt X_i là kết quả lần quan sát thứ i

$$\begin{aligned} X_i \colon \Omega_* &\longrightarrow \mathbb{R} \\ \omega_* &\longmapsto X_i(\omega_*) = 1 \\ \bar{\omega}_* &\longmapsto X_i(\bar{\omega}_*) = 0 \end{aligned}$$

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

$$\begin{aligned} X \colon \Omega &\longrightarrow \mathbb{R} \\ \omega = (\omega_{(1)}, \omega_{(2)}, \ldots, \omega_{(n)}) &\longmapsto X(\omega) = X_1(\omega_{(1)}) + \cdots + X_n(\omega_{(n)}) \end{aligned}$$

Phân phối nhị thức: Mô hình (tt)

Với $x \in \{0, 1, \ldots, n\}$, $(X = x) = \{\omega \in \Omega \colon \exists I \subset \{1, 2, \ldots, n\}, |I| = x, \omega_{(i)} = \omega_* \forall i \in I, \omega_{(i)} = \bar{\omega}_* \forall i \notin I\}$ 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 $\omega \in (X = x)$ thì

$$P(\omega) = p^x (1-p)^{n-x}$$

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

$$f(x) = P(X = x) = \sum_{\omega \in (X=x)} P(\omega) = C_n^x p^x (1-p)^{n-x}$$

Vậy X có phân phối nhị thức.

Phân phối nhị thức - Một số ví dụ

Ví dụ 7

Trong một gia đình có 6 người con. Tính xác suất gia đình này

1. có đúng 3 con trai.
2. có nhiều nhất 3 con trai
3. có ít nhất 3 con trai.

Gợi ý 8

Quan sát sinh con trai trong 6 lần độc lập. $P(\omega) = P(trai) = 1/2$.
Gọi X là số con trai trong 6 lần sinh. $X \in \{0, 1, 2 \ldots, 6\}$ và $X \sim B(6, 1/2)$ với hàm mật độ

$$f(x) = \begin{cases} C_6^x (1/2)^x (1/2)^{6-x} & x = 0, 1, 2, \ldots, 6 \\ 0 & \text{nơi khác} \end{cases}$$

Phân phối nhị thức - Một số ví dụ

Ta có bảng phân phối

X	0	1	2	3	4	5	6
$P(X=k)$	0.016	0.093	0.24	0.32	0.24	0.093	0.016

1. Xác suất để gia đình này có đúng 3 con trai:
$$P(X=3)=0.32$$
2. Xác suất để gia đình này có nhiều nhất là 3 con trai
$$P(X\le 3)=P(X=0)+P(X=1)+P(X=2)+P(X=3)=0.67$$
3. Xác suất để gia đình này có ít nhất 3 con trai
$$P(X\ge 3)=P(X=3)+P(X=4)+P(X=5)+P(X=6)=0.67$$

Phân phối nhị thức - Một số ví dụ

Ví dụ 9

Tại một địa phương tỉ lệ sốt rét là 25% dân số. Chọn ngẫu nhiên 6 người. Tính khả năng để có 4 người bị sốt rét.

Ví dụ 10

Một lô thuốc (rất nhiều), 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 mật độ xác suất của X.

Phân phối nhị thức - Một số ví dụ

Ví dụ 11

Một bài thi trắc nghiệm gồm 10 câu hỏi, mỗi câu hỏi có 4 phương án trả lời trong đó chỉ có 1 phương án đúng. Giả sử mỗi câu trả lời đúng được 4 điểm, mỗi câu trả lời sai trừ 2 điểm. Một sinh viên làm bài bằng cách chọn ngẫu nhiên đáp án cho các câu hỏi. Tính xác suất:

1. Để sinh viên được 4 điểm.
2. Để sinh viên được điểm âm

Phân phối nhị thức - Các đặc trưng

Định lí 12

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

1. $\mathbb{E}(X)=np,\ \mathbb{V}ar(X)=npq,\quad$ *với* $q=1-p$.
2. $Mod(X)$ *là (các) số nguyên thỏa* $np-q\le Mod(X)\le np+p$.

Chứng minh

1. Ta có,
$$E[X^k]=\sum_{i=0}^{n} i^k C_n^i p^i(1-p)^{n-i}=\sum_{i=1}^{n} i^k C_n^i p^i(1-p)^{n-i}$$
Sử dụng đẳng thức $iC_n^i=nC_{n-1}^{i-1}$, ta viết lại

Chứng minh (tt)

$$\begin{aligned} E[X^k] &= np\sum_{i=1}^{n} i^{k-1}C_{n-1}^{i-1}p^{i-1}(1-p)^{n-i} \\ &\overset{\text{đặt } j=i-1}{=} np\sum_{j=0}^{n}(j+1)^{k-1}C_{n-1}^{j}p^{j}(1-p)^{n-1-j} \\ &= npE(Y+1)^{k-1} \quad \text{với } Y \sim B(n-1,p) \end{aligned}$$

Với $k=1$, $EX=np$.
Với $k=2$, $E[X^2]=npE(Y+1)=np((n-1)p+1)$.
Do đó, $Var(X)=E(X^2)-(EX)^2=np(1-p)$.

Chứng minh (tt)

(ii) Ta xét tỉ số

$$\begin{aligned} \frac{P(X=k)}{P(X=k-1)} &= \frac{\frac{n!}{k!(n-k)!}p^k(1-p)^k}{\frac{n!}{(k-1)!(n-k+1)!}p^{k-1}(1-p)^{n-k+1}} \\ &= \frac{(n-k+1)p}{k(1-p)} \end{aligned}$$

Do đó $P(X=k)\geq P(X=k-1)$ nếu và chỉ nếu $(n-k+1)p\geq k(1-p)$, tức là $k\leq np+p$.

Phân phối nhị thức - Ví dụ

Ví dụ 13

Hàng đóng thành kiện, mỗi kiện 10 sản phẩm, trong đó có 3 phế phẩm. Khi kiện hàng được giao cho khách hàng, khách hàng sẽ lấy ngẫu nhiên ra 2 sản phẩm trong kiện để kiểm tra. Nếu cả hai sản phẩm đều tốt, kiện hàng sẽ được nhận, ngược lại kiện hàng sẽ bị trả lại. Gọi X là số kiện hàng được nhận trong số 100 kiện hàng giao cho khách hàng. Tìm $\mathbb{E}(X)$, $\mathbb{V}ar(X)$ và $Mod(X)$.

Outline

Phân phối siêu bội

Định nghĩa 14

Cho biến ngẫu nhiên rời rạc X nhận các giá trị $0, 1, \ldots, n$. X có phân phối siêu bội, kí hiệu $X \sim H(n, M, N)$, khi hàm xác suất có dạng

$$f(x) = \begin{cases} \dfrac{C_M^x C_{N-M}^{n-x}}{C_N^n} & \text{nếu } x = 0, 1, \ldots, n \\ 0 & \text{nếu khác} \end{cases}$$

Phân phối siêu bội

Nhận xét 15

Bởi vì ta quy ước rằng C_r^k bằng 0 khi $k < 0$ hoặc $k > r$ nên $f(x)$ sẽ bằng 0 nếu x không thỏa

$$\begin{cases} 0 \le x \le M \\ 0 \le n - x \le N - M \end{cases}$$

tức là

$$\max\{0, n - (N - M)\} \le x \le \min\{n, M\}$$

Phân phối siêu bội - Mô hình và các đặc trưng

Mô hình siêu bội

Từ một hộp có M bi đỏ, $N - M$ bi đen lấy ngẫu nhiên không hoàn lại n bi. Gọi X là số bi đỏ trong n bi lấy ra. Khi đó $X \sim H(n, M, N)$.

Chứng minh.

Dễ dàng có được. □

Định lí 16

Cho $X \sim H(n, M, N)$ và đặt $p = \dfrac{M}{N}$, $q = 1 - p$. Khi đó

1. $E(X) = np$
2. $Var(X) = npq\dfrac{N-n}{N-1}$

Chứng minh

$$E(X^k) = \sum_{i=0}^{n} i^k P(X = i) = \sum_{i=1}^{n} i^k C_M^i C_{N-M}^{n-i} / C_N^n$$

Sử dụng hệ thức

$$iC_M^i = MC_{M-1}^{i-1} \text{ và } nC_N^n = NC_{N-1}^{n-1}$$

Chứng minh ...

Ta viết lại

$$E(X^k) = \frac{nM}{N}\sum_{i=1}^{n} i^{k-1} C_{M-1}^{i-1} C_{N-M}^{n-i} / C_{N-1}^{n-1}$$
$$= \frac{nM}{N}\sum_{j=0}^{n-1} (j+1)^{k-1} C_{M-1}^{j} C_{N-M}^{n-1-j} / C_{N-1}^{n-1}$$
$$= \frac{nM}{N} E[(Y+1)^{k-1}] \quad \text{với } Y \sim H(n-1, M-1, N-1).$$

Do đó, với $k = 1$

$$EX = \frac{nM}{N} = np$$

Chứng minh ...

với $k = 2$,

$$E(X^2) = \frac{nM}{N}E(Y+1) = \frac{nM}{N}\left[\frac{(n-1)(M-1)}{N-1} + 1\right]$$

Từ đó,

$$Var(X) = E(X^2) - (EX)^2$$
$$= np\left[\frac{(n-1)(M-1)}{N-1} + 1 - np\right]$$
$$= np\left[\frac{(n-1)(Np-1)}{N-1} + 1 - np\right]$$
$$= npq\frac{N-n}{N-1}$$

Phân phối siêu bội - Ví dụ

Ví dụ 17

Một lớp có 50 sinh viên trong đó có 30 nữ. Cần chọn ra 10 bạn để tham gia vào công tác chuẩn bị cho 1 hoạt động sắp tới của trường. Nếu ta chọn các bạn trên một cách ngẫu nhiên, xác suất để số sinh viên nữ được chọn không quá 3 là bao nhiêu? Xác suất để chọn được ít nhất 1 sinh viên nữ là bao nhiêu?

Gợi ý

Gọi X là số sinh viên nữ trong số 10 sinh viên được chọn.

$$X \sim H(50, 30, 10)$$

Xác suất để số sinh viên nữ được chọn không quá 3 là

$$P(X \le 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$$
$$= \frac{C_{30}^0 C_{20}^{10}}{C_{50}^{10}} + \frac{C_{30}^1 C_{20}^{9}}{C_{50}^{10}} + \frac{C_{30}^2 C_{20}^{8}}{C_{50}^{10}} + \frac{C_{30}^3 C_{20}^{7}}{C_{50}^{10}} = 0.0365$$

Xác suất để có ít nhất 1 nữ là

Gợi ý . . .

$$\begin{aligned} P(X \geq 1) &= 1 - P(X < 1) \\ &= 1 - P(X = 0) \\ &= 1 - \frac{C_{30}^{0} C_{20}^{10}}{C_{50}^{10}} \approx 1 \end{aligned}$$

Outline

Phân phối Poisson

Định nghĩa 18 (Phân phối Poisson)

Cho biến ngẫu nhiên rời rạc X nhận các giá trị $0, 1, 2, \ldots$ X có phân phối Poisson, kí hiệu $X \sim P(\lambda)$, khi hàm xác suất có dạng

$$f(x) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!} & x = 0, 1, 2, \ldots \\ 0 & \text{nơi khác} \end{cases} \qquad \text{với } \lambda > 0$$

Định lí 19 (Các đặc trưng của biến ngẫu nhiên có phân phối Poisson)

Nếu b.n.n X có phân phối Poisson với tham số λ, $X \sim P(\lambda)$, thì

1. *Kỳ vọng $\mathbb{E}(X) = \lambda$.*
2. *Phương sai $\mathbb{V}ar(X) = \lambda$.*

Chứng minh

Lưu ý rằng $\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{\lambda}$

1. $$\begin{aligned} \mathbb{E}X &= \sum_{x=1}^{\infty} x f(x) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \\ &\overset{\text{đặt } t=x-1}{=} \lambda e^{-\lambda} \sum_{t=0}^{\infty} \frac{\lambda^t}{t!} = \lambda \end{aligned}$$

Chứng minh (tt)

(ii)

$$\begin{aligned}\mathbb{E}(X^2) &= \sum_{x=0}^{\infty} x^2 e^{-\lambda}\frac{\lambda^x}{x!} = \sum_{x=0}^{\infty}[x(x-1)+x]e^{-\lambda}\frac{\lambda^x}{x!}\\ &= e^{-\lambda}\sum_{x=2}^{\infty}\frac{\lambda^x}{(x-2)!} + e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^x}{(x-1)!}\\ &= \lambda^2 e^{-\lambda}\sum_{t=0}^{\infty}\frac{\lambda^t}{t!} + \lambda = \lambda^2+\lambda\end{aligned}$$

Do đó, $\mathbb{V}ar(X) = \mathbb{E}(X^2) - (\mathbb{E}X)^2 = \lambda$.

Định lí 20 (giới hạn Poisson)

Cho $X \sim B(n;p)$ và đặt $\lambda = np$. Khi đó

$$\lim_{\substack{n\to\infty\\ p\to 0}} P(X=x) = e^{-\lambda}\frac{\lambda^x}{x!} \tag{1}$$

Chứng minh

Lưu ý rằng $\lim_{n\to\infty}\left(1+\frac{\alpha}{n}\right)^n = e^{\alpha}$.

$$\begin{aligned}&P(X=x)\\ =\ & C^x p^x(1-p)^{n-x} = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}\\ =\ & \frac{(n-x+1)(n-x+2)\cdots(n-1)n}{x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}\\ =\ & \left(1-\frac{x-1}{n}\right)\left(1-\frac{x-2}{n}\right)\cdots\left(1-\frac{1}{n}\right)1.\frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^{n-x}\end{aligned}$$

Chứng minh (tt)

Cho $n\to\infty$,

$$1-\frac{x-i}{n}\to 1 \quad \forall i = 1,\ldots,x-1$$

$$\left(1-\frac{\lambda}{n}\right)^{n-x} \to e^{-\lambda}$$

Vậy (1) được chứng minh.

Nhận xét 21

Định lí trên cho thấy trong phân phối nhị thức nếu n lớn, p nhỏ, $np=\lambda$ thì ta có thể tính các xác suất xấp xỉ theo luật Poisson và vì vậy việc tính toán sẽ dễ dàng hơn. Để an toàn, xấp xỉ này được dùng khi $n\geq 100$, $p\leq 0.01$ và $np\leq 20$.

Phân phối Poisson - Mô hình

Đó là những quan sát mà số lần lặp lại lớn (n lớn) mà xác suất biến cố ta lưu tâm $P(\omega) = p$ thì nhỏ.
Chẳng hạn ta lưu ý đến những biến cố hiếm, xảy ra trong một thời gian, không gian nhất định:

- Số trẻ em sinh đôi trong một năm tại 1 bệnh viện X
- Số tai nạn giao thông tại 1 ngã tư trong 1 năm
- Số hồng cầu trong mỗi ô của hồng cầu kế.
- Số chữ in sai trong một trang.
- Số người sống lâu trên 100 tuổi trong một cộng đồng dân cư.
- Số người đến một bưu điện nào đó trong một ngày.

Phân phối Poisson - Ví dụ

Ví dụ 22

Giả sử số lỗi in trong một trang 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.

Ví dụ 23

Giả sử xác suất tử vong của bệnh sốt xuất huyết là 0.007. Tính xác suất để có 5 người chết do sốt xuất huyết trong một nhóm 400 người.

Phân phối Poisson - Ví dụ

Ví dụ 24

Tỉ lệ thuốc hỏng một lô thuốc (rất nhiều) là $p = 0.05$. Ta lấy ngẫu nhiên $n = 20$ lọ. Gọi X là số lọ hỏng. Tìm hàm mật độ của X và so sánh với giá trị xấp xỉ bởi phân phối Poisson.

Ví dụ 25

Một trung tâm bưu điện nhận trung bình 150 cuộc điện thoại trong một giờ, tìm xác suất để trung tâm bưu điện này nhận không quá hai cuộc gọi trong một phút.

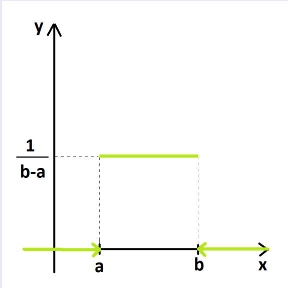
Outline

Phân phối đều

Định nghĩa 26 (Phân phối đều)

Biến ngẫu nhiên liên tục X được gọi là có phân phối đều trên đoạn $[a; b]$, ký hiệu $X \sim U[a; b]$, nếu hàm mật độ xác suất của X có dạng

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{khi } x \in [a, b] \\ 0 & \text{nơi khác} \end{cases}$$

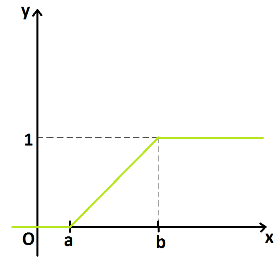


Hình 1: Hàm mật độ của phân phối đều trên khoảng $[a, b]$

Phân phối đều

Từ định nghĩa trên ta có được hàm phân phối xác suất của $X \sim U[a; b]$

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



Hình 2: Hàm phân phối xác suất của phân phối đều trên khoảng $[a, b]$

Phân phối đều

Định lí 27 (Các đặc trưng của biến ngẫu nhiên có phân phối đều)

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

1. *Kỳ vọng* $\mathbb{E}(X) = \frac{a+b}{2}$.
2. *Phương sai* $\mathbb{V}ar(X) = \frac{(a-b)^2}{12}$.

Chứng minh.

Dễ dàng có được. □

Phân phối đều

Ví dụ 28

Tại một trạm xe buýt khoảng cách giữa các chuyến liên tiếp của một tuyến xe buýt T là 15 phút. Chuyến đầu tiên đến trạm lúc 7 giờ sáng. Nếu một hành khách tới trạm xe buýt vào một thời điểm có phân phối đều từ 7 giờ tới 7 giờ 30 để đi tuyến xe buýt T Tính xác suất để anh ta đợi:

1. ít hơn hoặc bằng 5 phút
2. ít hơn hoặc bằng 10 phút
3. từ 6 đến 12 phút

Outline

Phân phối mũ

Định nghĩa 29

Cho biến ngẫu nhiên X liên tục, với $\lambda > 0$, X có phân phối mũ, kí hiệu $X \sim Exp(\lambda)$, khi hàm mật độ có dạng:

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

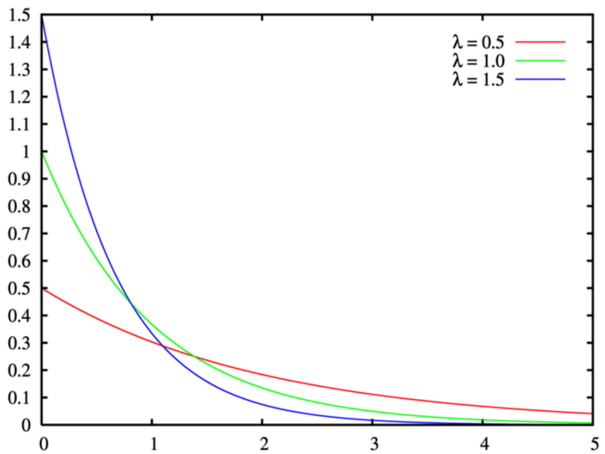
Định lí 30

Cho B.N.N $X \sim Exp(\lambda)$. Khi đó,

1. *Kỳ vọng $\mathbb{E}X = 1/\lambda$,*
2. *Phương sai $\mathbb{V}ar(X) = 1/\lambda^2$, và*
3. *Tính không nhớ*

$$P(X > s + t \mid X > t) = P(X > s) \quad \text{với mọi } s, t \geq 0.$$

Phân phối mũ - Minh họa



Hình 3: Hàm mật độ của $Exp(\lambda)$

Phân phối mũ - Ví dụ

Ví dụ 31

Thời gian giữa hai cuộc gọi đến tổng đài trong khoảng thời gian từ 14h00-16h00 là một biến ngẫu nhiên có phân phối mũ với trung bình 2 phút. Giả sử vừa có một cuộc gọi đến tổng đài. Hỏi xác suất để trong 3 phút tiếp theo không có cuộc gọi đến tổng đài.

Outline

Phân phối chuẩn hóa (Standard normal distribution)

Định nghĩa 32

Cho biến ngẫu nhiên Z liên tục, Z có phân phối chuẩn hóa (hay chuẩn tắc), kí hiệu $Z \sim N(0,1)$, khi hàm mật độ có dạng:

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

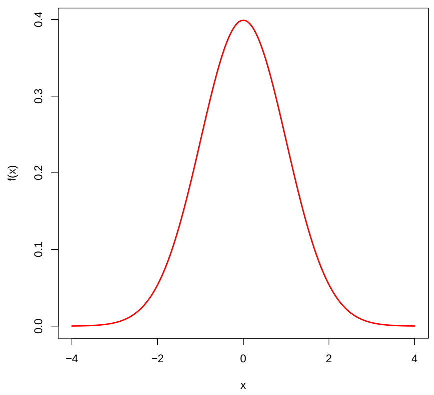
Định lí 33

B.N.N $Z \sim N(0,1)$ có kì vọng $\mathbb{E}Z = 0$ và phương sai $\mathbb{V}ar(Z) = 1$.

Chứng minh.

Chú ý rằng $\int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{\pi}$. □

Phân phối chuẩn hóa - Minh họa



Hình 4: Hàm mật độ của $N(0,1)$

Phân phối chuẩn hóa - Hàm phân phối

Hàm phân phối

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}} du$$

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

Tính chất

a.
$$\Phi(z) + \Phi(-z) = 1$$

b.
$$P(-a \le Z \le a) = 2\Phi(a) - 1$$

Phân phối chuẩn hóa - Ví dụ

Ví dụ 34

Cho biến ngẫu nhiên $Z \sim N(0,1)$. Tính các xác suất sau

1. $P(Z \leq 1.55)$
2. $P(Z \leq -1.45)$
3. $P(-1 < Z \leq 1.5)$

Phân phối chuẩn (Normal distribution)

Định nghĩa 35

Cho biến ngẫu nhiên X liên tục, với $\sigma > 0$, μ là hai tham số, X có phân phối chuẩn, kí hiệu $X \sim N(\mu, \sigma^2)$, khi hàm mật độ có dạng

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{với } x \in \mathbb{R}$$

Định lí 36

B.N.N $X \sim N(\mu, \sigma^2)$ *có kì vọng* $\mathbb{E}X = \mu$ *và phương sai* $\mathbb{V}ar(X) = \sigma^2$.

Chứng minh.

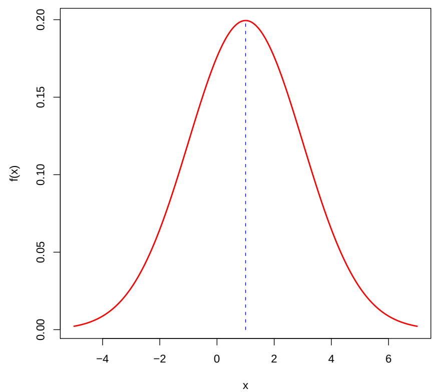
Dễ dàng có được bằng cách đổi biến và chú ý rằng

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1$$

□

Vậy trong phân phối chuẩn thì tham số μ và σ chính là trung bình và độ lệch chuẩn.

Phân phối chuẩn - Minh họa



Hình 5: Hàm mật độ của $N(1,4)$

Phân phối chuẩn - Phân phối chuẩn hóa

Định lí 37

Nếu $X \sim N(\mu, \sigma^2)$ *thì* $\dfrac{X-\mu}{\sigma} \sim N(0,1)$.

Chứng minh.

Đặt $Y = \dfrac{X-\mu}{\sigma}$. Ta có,

$$F_Y(y) = P(Y \le y) = P\left(\frac{X-\mu}{\sigma} \le y\right) = P(X \le \sigma y + \mu) = F_X(\sigma y + \mu)$$

Do đó,

$$f_Y(y) = F_Y'(y) = \sigma f_X(\sigma y + \mu) = \sigma \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\sigma y + \mu - \mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

Vậy $Y \sim N(0,1)$. □

Phân phối chuẩn

Nhận xét 38

Định lý 37 cho phép chúng ta đưa một biến ngẫu nhiên có phân phối chuẩn bất kỳ về phân phối chuẩn hóa.

Hệ quả 39

Nếu $X \sim N(\mu, \sigma^2)$ *thì*

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

Hệ quả 40

Nếu $X \sim N(\mu, \sigma^2)$ *thì*

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

Phân phối chuẩn

Quy tắc $k\sigma$

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

(i) $P(|X-\mu| < \sigma) = 0.68$

(ii) $P(|X-\mu| < 2\sigma) = 0.955$

(iii) $P(|X-\mu| < 3\sigma) = 0.997$

Ví dụ 41

Chỉ số thông minh (IQ), được đo bằng bài kiểm tra IQ Stanford-Binet, có phân phối chuẩn trong một tổng thể nào đó. IQ trung bình là 100 điểm, và độ lệch chuẩn là 16 điểm. Hỏi phần trăm số người trong tổng thể có IQ

a. từ 140 trở lên?
b. từ 80 trở xuống?
c. giữa 80 và 140?

Xấp xỉ phân phối nhị thức bằng phân phối chuẩn

Định lí 42 (Moivre - Laplace)

Cho X *là một biến ngẫu nhiên nhị thức với các tham số* n *và* p*. Khi đó với các số* a, b *bất kì,* $a < b$*,*

$$\lim_{n\to\infty} P\left(a < \frac{X - np}{\sqrt{np(1-p)}} < b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt$$

Chú ý rằng $EX = np$, $\sigma_X = \sqrt{np(1-p)}$.

Áp dụng

Định lí nói rằng khi n lớn ta có thể xấp xỉ phân phối nhị thức $B(n,p)$ bằng phân phối chuẩn $N(np, np(1-p))$.

Xấp xỉ phân phối nhị thức bằng phân phối chuẩn

Điều kiện áp dụng

- Xác suất p không quá gần 0 hoặc 1, sao cho $0.1 < p < 0.9$.
- $np \geq 5$ và $np(1-p) \geq 5$.

Hiệu chỉnh liên tục (Correction for continuity)

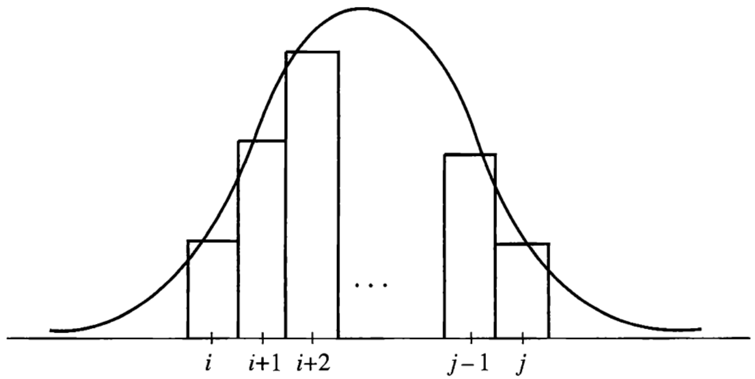
Vì X trong phân phối nhị thức là rời rạc nên khi tính xấp xỉ các giá trị xác suất của X bằng phân phối chuẩn ta đã chuyển sang một biến mới liên tục nên trong thực hành phải thực hiện phép hiệu chỉnh liên tục như sau:

$$P(X \leq x) = P(X < x + 0.5) \approx \Phi\left(\frac{x + 0.5 - \mu}{\sigma}\right)$$

$$P(X < x) = P(X < x - 0.5) \approx \Phi\left(\frac{x - 0.5 - \mu}{\sigma}\right)$$

Hiệu chỉnh liên tục

Minh họa



Xấp xỉ phân phối nhị thức bằng phân phối chuẩn

Ví dụ 43

Một xạ thủ có xác suất bắn trúng mục tiêu ở mỗi lần bắn là 0.8. Xạ thủ này bắn 64 phát vào bia. Tính xác suất

a) Có 50 phát trúng bia.
b) Có từ 45 đến 52 phát trúng bia.
c) Có không quá 51 phát trúng bia.

Một số ví dụ về phân phối chuẩn

Theo Borel nếu một biến ngẫu nhiên là kết quả của nhiều nguyên nhân, mỗi nguyên nhân tác động một ít và không có nguyên nhân nào là quyết định, thì biến ngẫu nhiên đó có phân phối chuẩn.
Vậy:

- Các số đo về đặc tính sinh học: chiều cao, cân nặng, huyết áp, nồng độ,... hầu như có phân phối chuẩn.
- Trong xã hội: lợi tức hàng năm, sản lượng một vụ mùa,... tuân theo phân phối chuẩn.
- Sai số trong đo lường về vật lí cũng có phân phối chuẩn.

Outline

Phân phối Gamma

Định nghĩa 44 (Hàm Gamma)

Với $\alpha > 0$, đặt $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$.

Tính chất

- $\Gamma(1) = 1$
- $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$
- $\Gamma(n+1) = n!$
- $\Gamma(1/2) = \sqrt{\pi}$

Chứng minh.

Dễ dàng có được. □

Phân phối Gamma

Định nghĩa 45 (Phân phối Gamma)

B.N.N X được gọi là có phân phối Gamma với hai tham số dương α và β, kí hiệu $X \sim G(\alpha, \beta)$, nếu hàm mật độ của X có dạng

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1}e^{-x/\beta} & x > 0 \\ 0 & x \le 0 \end{cases}$$

Các đặc trưng

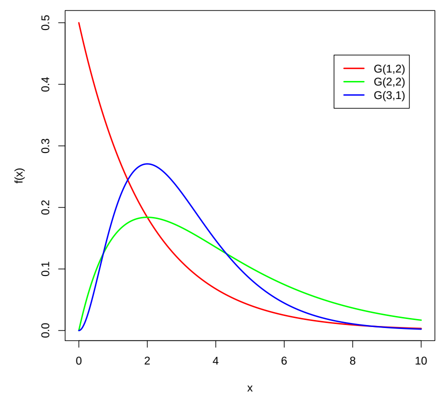
- Trung bình: $\mathbb{E}X = \alpha\beta$.
- Phương sai: $\mathbb{V}ar(X) = \alpha\beta^2$.

Chứng minh.

Dễ dàng có được. □

Phân phối Gamma

Minh họa



Hình 6: Hàm mật độ của $G(1,2)$, $G(2,2)$, $G(3,1)$

Outline

Phân phối Chi bình phương

Định nghĩa 46 (Phân phối Chi bình phương: $X \sim \chi^2(r), r = 1, 2, 3, \ldots$)

$X \sim \chi^2(r)$ nếu $X \sim G(r/2, 2)$

Hàm mật độ của X

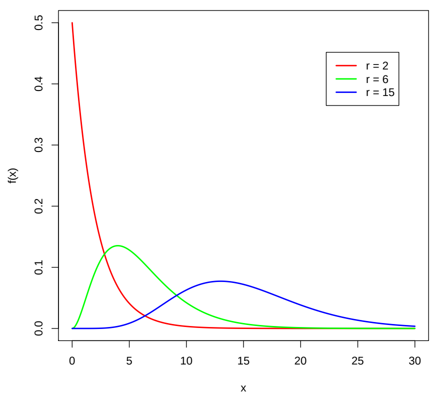
$$f(x) = \begin{cases} \frac{1}{\Gamma(1/2)2^{r/2}} x^{\frac{r}{2}-1} e^{\frac{x}{2}} & \text{nếu } x > 0 \\ 0 & \text{nơi khác} \end{cases}$$

Các đặc trưng

- Kì vọng: $\mathbb{E}X = \frac{r}{2}2 = r$.
- Phương sai: $\mathbb{V}ar = \frac{r}{2}2^2 = 2r$.

Phân phối Chi bình phương

Minh họa



Hình 7: Hàm mật độ của $\chi^2(2), \chi^2(6), \chi^2(15)$

Định lí 47

Nếu $X \sim N(0,1)$ thì $Y = X^2 \sim \chi^2(1)$.

Chứng minh.

Biến ngẫu nhiên $Y \geq 0$, ta tính hàm phân phối của Y.

$$\begin{aligned} G(y) &= P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= \Phi(\sqrt{y}) - \Phi(-\sqrt{y}) \quad \text{vì } X \sim N(0,1) \\ &= 2\Phi(\sqrt{y}) - 1 \end{aligned}$$

Hàm mật độ của Y,

$$\begin{aligned} g(y) &= G'(y) = 2\Phi'(\sqrt{y}) = 2.\frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}}.\frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{\pi}2^{\frac{1}{2}}} y^{-\frac{1}{2}} e^{-\frac{y}{2}} \\ &= \frac{1}{\Gamma(1/2)2^{1/2}}.y^{\frac{1}{2}-1}.e^{-\frac{y}{2}} \end{aligned}$$

chính là hàm mật độ của $\chi^2(1)$. Vậy $Y \sim \chi^2(1)$. $\square$

Định lí 48

Nếu $X \sim \chi^2(r)$, $Y \sim \chi^2(s)$, X và Y độc lập thì $Z = X + Y \sim \chi^2(r+s)$

Chứng minh.

Có thể sử dụng hàm đặc trưng để chứng minh. □

Hệ quả 49

Nếu $X_i \sim \chi^2(r_i)$ với mọi $i = 1, \ldots, n$ và các X_i độc lập, thì

$$X_1 + X_2 + \cdots + X_n \sim \chi^2(r_1 + r_2 + \cdots + r_n)$$

Hệ quả 50

Nếu $X_1, X_2, \ldots, X_r$ độc lập và có cùng phân phối chuẩn $N(0,1)$ thì

$$X_1^2 + X_2^2 + \cdots + X_r^2 \sim \chi^2(r)$$

Outline

Phân phối Student

Định nghĩa 51

Xét hai biến ngẫu nhiên độc lập $X \sim N(0,1)$, $Y \sim \chi^2(n)$. Đặt $T = \frac{X}{\sqrt{\frac{Y}{n}}}$. Khi đó, phân phối của BNN T được gọi là phân phối Student bậc tự do n. Kí hiệu $T \sim T(n)$.

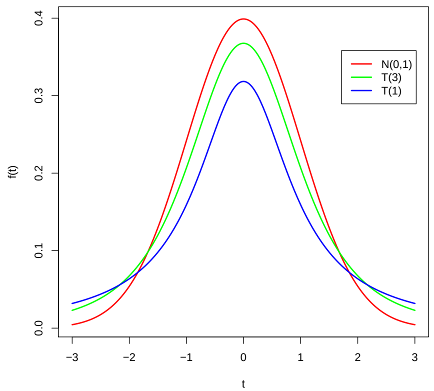
Định lí 52

B.N.N $T \sim T(n)$ có hàm mật độ

$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)} \cdot \frac{1}{\left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}}, \quad t \in \mathbb{R}$$

và $\mathbb{E}(T) = 0, \mathbb{V}ar(T) = \frac{n}{n-2}$. Khi $n \geq 30$, phân phối $T(n)$ gần trùng với phân phối chuẩn tắc $N(0,1)$.

Phân phối Student
Minh họa



Hình 8: Hàm mật độ của $T(1)$, $T(3)$ và $N(0,1)$

Outline

Phân phối Fisher

Định nghĩa 53

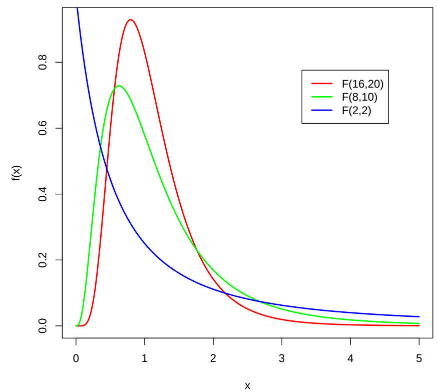
Xét hai biến ngẫu nhiên X, Y độc lập: $X \sim \chi^2(n), Y \sim \chi^2(m)$. Đặt $F = \frac{X/n}{Y/m}$. Khi đó, phân phối của B.N.N F được gọi là phân phối Fisher bậc tự do n, m. Kí hiệu $F \sim F(n,m)$.

Định lí 54

B.N.N $F \sim F(n,m)$ có hàm mật độ

$$h(f) = \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{n}{2}).\Gamma(\frac{m}{2})}.\left(\frac{m}{n}\right)^{\frac{m}{2}}.\frac{f^{\frac{m}{2}-1}}{(1+\frac{m}{n}f)^{\frac{m+n}{2}}}, \quad f \geq 0$$

Phân phối Fisher
Minh họa



Hình 9: Hàm mật độ của $F(16,20)$, $F(8,10)$ và $F(2,2)$