Trường Đại Học Khoa Học Tự Nhiên TP.HCM

XÁC SUẤT THỐNG KÊ

Lưu hành nội bộ
2020

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc

Phân phối Bernoulli $\mathcal{B}(1,p)$

Định nghĩa:

Cho biến ngẫu nhiên (b.n.n.) X rời rạc nhận hai trị số 0, 1. Ta nói X có phân phối Bernoulli khi:

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

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

Đặc trưng

Kì vọng: $\mathbb{E}[X]=0(1-p)+1.p=p$.
Phương sai: $\mathbb{V}(X)=p(1-p)$.

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n,m)$

Phân phối Bernoulli: Mô hình

Coi một thí nghiệm ngẫu nhiên có hai kết quả $\Omega=\{\omega,\bar\omega\}$, trong đó $\mathbb{P}(\omega)=p$.
Gọi X là số lần ω xuất hiện

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

Bảng phân phối xác suất của X:

X	1	0
p	p	$1-p$

Vậy X có phân phối Bernoulli $\mathcal{B}(1,p)$.
Nhận xét: *Mọi thí nghiệm ngẫu nhiên chỉ có hai kết quả đều có phân phối Bernouilli.*

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n,m)$

Phân phối Bernoulli: Ví dụ

VD 4.1: Tung đồng xu (đồng nhất) một lần, chúng ta quan tâm mặt ngửa.

$$\text{Đặt}\quad X=\begin{cases}1 & \text{nếu xuất hiện mặt ngửa}\\ 0 & \text{nếu xuất hiện mặt sấp}\end{cases}\quad\text{thì}\quad X\sim\mathcal{B}(1,1/2).$$

VD 4.2: Quan sát giới tính trong một lần sanh.

$$\text{Đặt}\quad Y=\begin{cases}1 & \text{nếu con trai}\\ 0 & \text{nếu con gái}\end{cases}\quad\text{thì}\quad Y\sim B(1,1/2).$$

VD 4.3: Tung con xúc sắc (đồng nhất) một lần, chúng ta quan tâm mặt 6 chấm.

$$\text{Đặt}\quad Z=\begin{cases}1 & \text{nếu mặt 6 xuất hiện}\\ 0 & \text{nếu là mặt khác}\end{cases}\quad\text{thì}\quad Z\sim B(1,1/6).$$

Phân phối nhị thức $\mathcal{B}(n, p)$

Định nghĩa:

Biến ngẫu nhiên rời rạc X nhận các giá trị $k = 0, 1, \ldots, n$ với xác suất

$$\mathbb{P}(X = k) = C_n^k p^k (1 - p)^{n-k}$$

đgl có luật phân phối nhị thức với tham số n và p. Kí hiệu: $X \sim \mathcal{B}(n, p)$ trong đó $n \geq 0$ và $0 \leq p \leq 1$.

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc

Phân phối nhị thức: Mô hình

Nhận xét :

- Khi có n phép thử Bernoulli độc lập, ở mỗi phép thử có xác suất thành công là p, thì biến ngẫu nhiên X chỉ số lần thành công sẽ có luật phân phối nhị thức $\mathcal{B}(n, p)$. Ta gọi đó là mô hình nhị thức.
- Biến ngẫu nhiên rời rạc X có phân phối nhị thức $\mathcal{B}(n, p)$ có thể biểu diễn dưới dạng tổng của n biến ngẫu nhiên độc lập X_i có phân phối Bernoulli $\mathcal{B}(1, p)$:

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc

Phân phối nhị thức : Ví dụ

VD 4.4 : 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:

(i) Để sinh viên được 4 điểm.
(ii) Để sinh viên được điểm âm.

VD 4.5 : Một bài trắc nghiệm của mệt game show trên truyền hình có 6 câu hỏi, mỗi câu hỏi có 5 phương án trả lời trong đó chỉ có 1 phương án trả lời đúng. Một người làm bài trắc nghiệm này bằng cách chọn ngẫu nhiên 1 trong 5 phương án trả lời cho câu hỏi. Tính xác suất:

(i) trả lời đúng 3 câu.
(ii) trả lời đúng ít nhất 3 câu.

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc

Phân phối nhị thức : Ví dụ (tt)

VD 4.6 : Trong một gia đình có 6 người con. Tính xác suất gia đình này

(i) có đúng 3 con trai.
(ii) có nhiều nhất 3 con trai
(iii) có ít nhất 3 con trai.

VD 4.7 : 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.

VD 4.8 : 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 : hướng dẫn giải (tt)

VD 4.4 : *Gọi X là số câu trả lời đúng của người này. Ta có* $X \sim \mathcal{B}(6, 0.2)$. *Vì ta xem người này trả lời 6 câu hỏi như thực hiện 6 phép thử độc lập, trong mỗi phép thử có :*

- *hoặc là trả lời đúng với xác suất là* $p = 1/5 = 0.2$;
- *hoặc là trả lời sai với xác suất là* $1 - p = 0.8$.

1. *Xác suất để người này trả lời đúng 3 câu là :*

$$P(X = 3) = C_6^3 (0.2)^3 (0.8)^3$$

2. *Xác suất để người này trả lời đúng ít nhất 3 câu là :*

$$P(X \geq 3) = 1 - [P(X = 0) + P(X = 1) + P(X = 2)] = 0.0988$$

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1, p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n, p)$
4.1.3 Phân phối siêu bội $H(N, M, n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha, \beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n, m)$

Phân phối nhị thức : hướng dẫn giải (tt)

VD 4.6 : *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.* $X \in \{0, 1, 2 \dots, 6\}$ *và* $X \sim B(6, 1/2)$ *với hàm mật độ*

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

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*

$$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 0.67$$

3. *Xác suất để gia đình này có ít nhất 3 con trai*

$$P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) = 0.67$$

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1, p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n, p)$
4.1.3 Phân phối siêu bội $H(N, M, n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha, \beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n, m)$

Phân phối nhị thức - Các đặc trưng

Định lý :

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

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

Chứng minh

1. Ta có,

$$E[X^k] = \sum_{i=0}^{n} i^k C_n^i p^i (1-p)^{n-i} = \sum_{i=1}^{n} i^k C_n^i p^i (1-p)^{n-i}$$

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1, p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n, p)$
4.1.3 Phân phối siêu bội $H(N, M, n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha, \beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n, m)$

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

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

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc

Phân phối nhị thức - Ví dụ

VD 4.9 : 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 $Mod(X)$.

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc

Phân phối siêu bội $H(N, M, n)$

Định nghĩa :

Biến ngẫu nhiên rời rạc X nhận các giá trị $k = 0, 1, \ldots, n$ với xác suất

$$\mathbb{P}(X=k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

với $\max\{0, n-(N-M)\} \leq k \leq \min\{n, M\}$ đgl tuân theo luật phân phối siêu bội.
Kí hiệu : $X \sim \mathcal{H}(N, M, n)$.

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc

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 $X \sim \mathcal{H}(n, M, N)$.

Đặc trưng :

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

- (i) $\mathbb{E}[X] = np$
- (ii) $Var(X) = npq\dfrac{N-n}{N-1}$

Chứng minh :

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

Sử dụng hệ thức

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

Ta viết lại

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T\sim T(n)$
4.1.4 Phân phối Fisher $F\sim F(n,m)$

Do đó, với $k=1$

$$\mathbb{E}[X]=\frac{nM}{N}=np$$

với $k=2$,

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

Từ đó,

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T\sim T(n)$
4.1.4 Phân phối Fisher $F\sim F(n,m)$

Phân phối siêu bội - Ví dụ

VD 4.10 : 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?

VD 4.11 : Trong cửa hàng có bán 100 bóng đèn trong đó có 5 bóng hư mà không kiểm tra thì không thể xác định được. Một người khách chọn ngẫu nhiên 2 bóng đèn để mua. Tìm xác suất để người này mua được cả 2 bóng đèn đều tốt.

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T\sim T(n)$
4.1.4 Phân phối Fisher $F\sim F(n,m)$

Gợi ý

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

$$X\sim\mathcal{H}(50,30,10)$$

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

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

Xác suất để có ít nhất 1 nữ là

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

Phân phối Poisson $\mathcal{P}(\lambda)$

Định nghĩa

Cho biến ngẫu nhiên rời rạc X nhận các giá trị $k = 0, 1, 2, \dots,$ với xác suất

$$\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad \text{với } \lambda > 0$$

đgl tuân theo phân phối Poison với tham số $\lambda > 0$. Kí hiệu : $X \sim \mathcal{P}(\lambda)$.

Đặc trưng

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

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1, p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n, p)$
4.1.3 Phân phối siêu bội $H(N, M, n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha, \beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n, m)$

Chứng minh : lưu ý rằng $\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{\lambda}$

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

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

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1, p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n, p)$
4.1.3 Phân phối siêu bội $H(N, M, n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha, \beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n, m)$

Định lý giới hạn Poisson

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

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

Nhận xét :

- Định lý trên cho thấy trong pp nhị thức nếu n lớn, p nhỏ, $np = \lambda$ thì ta có thể tính xác suất xấp xỉ theo luật Poisson và vì vậy việc tính toán sẽ dễ dàng hơn. Chú ý rằng xấp xỉ này được dùng khi $n \geq 100$, $p \leq 0.01$ và $np \leq 20$.
- Khi $n \geq 100$, $p \leq 0.01$ và $np \leq 20$ thì mô hình nhị thức tương đương với mô hình Poisson.

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1, p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n, p)$
4.1.3 Phân phối siêu bội $H(N, M, n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha, \beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n, m)$

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

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

Cho $n \to \infty$, $1 - \frac{k-i}{n} \to 1$,

$$\forall i = 1, \dots, k-1, \quad \left(1 - \frac{\lambda}{n}\right)^{n-k} \to e^{-\lambda}$$

Phân phối Poisson - Mô hình

Đó là những quan sát mà số lần lặp lại lớn (n lớn) mà xác suất biến cố ta quan tâm $P(\omega) = p$ thì nhỏ.
Ví dụ ta quan tâm 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 1 cộng đồng dân cư.
- Số người đến một bưu điện nào đó trong 1 ngày.

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1, p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n, p)$
4.1.3 Phân phối siêu bội $H(N, M, n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha, \beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n, m)$

Phân phối Poisson - Ví dụ

VD 4.12 : 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.

VD 4.13 : 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

VD 4.14 : 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.

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1, p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n, p)$
4.1.3 Phân phối siêu bội $H(N, M, n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha, \beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n, m)$

Gợi ý :

Gọi X là số cuộc gọi đến trung tâm bưu điện trong một giờ thì $X \sim \mathcal{P}(150)$,
Gọi Y là số cuộc gọi đến trung tâm bưu điện trong một phút thì $Y \sim \mathcal{P}(2.5)$.
Khi đó,

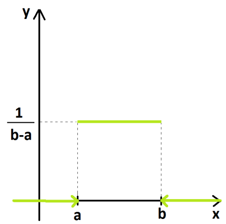
$$\begin{aligned}\mathbb{P}(Y \leq 2) &= \mathbb{P}(Y = 0) + \mathbb{P}(Y = 1) + \mathbb{P}(Y = 2) \\ &= e^{-2.5}\left(\frac{2.5^0}{0!} + \frac{2.5^1}{1!} + \frac{2.5^2}{2!}\right) = 0.5438.\end{aligned}$$

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1, p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n, p)$
4.1.3 Phân phối siêu bội $H(N, M, n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha, \beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n, m)$

Phân phối đều $\mathcal{U}[a; b]$

Định nghĩa:
Biến ngẫu nhiên liên tục X đgl tuân theo phân phối đều trên đoạn $[a; b]$, ký hiệu $X \sim U[a; b]$, nếu X có hàm mật độ xác suất

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

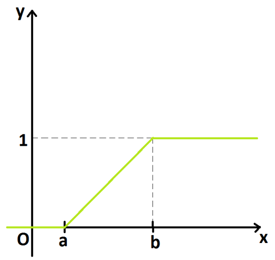


Hình: Hàm mật độ của phân phối đều trên khoảng $[a, b]$

Phân phối đều $\mathcal{U}[a;b]$

Hàm phân phối xác suất của $X \sim U[a;b]$ là

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



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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T\sim T(n)$
4.1.4 Phân phối Fisher $F\sim F(n,m)$

Phân phối đều $\mathcal{U}[a;b]$

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T\sim T(n)$
4.1.4 Phân phối Fisher $F\sim F(n,m)$

Phân phối đều $\mathcal{U}[a;b]$

VD 4.12: 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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T\sim T(n)$
4.1.4 Phân phối Fisher $F\sim F(n,m)$

Phân phối chuẩn hóa (Standard normal distribution)

Định nghĩa:

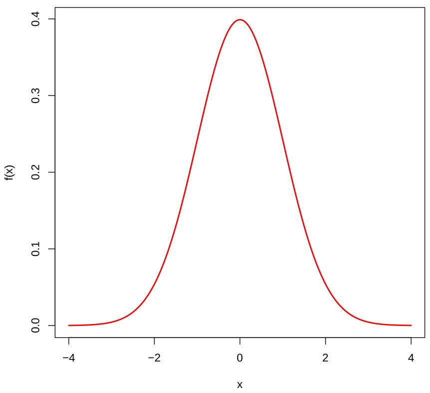
Biến ngẫu nhiên liên tục Z đgl tuân theo phân phối chuẩn (hay chuẩn tắc), kí hiệu $Z \sim \mathcal{N}(0,1)$ nếu Z có hàm mật độ:

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

Đặc trưng:

1. Kì vọng: $\mathbb{E}[Z] = 0$
2. phương sai $\mathbb{V}(Z) = 1$

Phân phối chuẩn hóa $\mathcal{N}(0,1)$



Hình: Hàm mật độ xác suất của $\mathcal{N}(0,1)$

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n,m)$

Phân phối chuẩn hóa $\mathcal{N}(0,1)$

Hàm phân phối tích lũy

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

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

Tính chất:

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n,m)$

Phân phối chuẩn hóa $\mathcal{N}(0,1)$

VD 4.13 : Cho biến ngẫu nhiên $Z \sim N(0,1)$. Tính các xác suất sau:

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n,m)$

Phân phối chuẩn hóa $\mathcal{N}(0,1)$

Giá trị h.p.p. của phân phối N(0,1): $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-t^2/2)\, dt \quad (0 \le x < 4)$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5190	0,5239	0,5279	0,5319	0,0359
0,1	5398	5438	5478	5517	5557	5596	5636	5675	5714	0753
0,2	5793	5832	5871	5910	5948	5987	6026	6064	6103	6141
0,3	6179	6217	6255	6293	6331	6368	6406	6443	6480	6517
0,4	6554	6591	6628	6664	6700	6736	6772	6808	6844	6879
0,5	6915	6950	6985	7019	7054	7088	7123	7157	7190	7224
0,6	7257	7291	7324	7357	7389	7422	7454	7486	7517	7549
0,7	7580	7611	7642	7673	7704	7734	7764	7794	7823	7852
0,8	7881	7910	7939	7967	7995	8023	8051	8078	8106	8133
0,9	8159	8186	8212	8238	8264	8289	8315	8340	8365	8389
1,0	8413	8438	8461	8485	8508	8531	8554	8577	8599	8621
1,1	8643	8665	8686	8708	8729	8749	8770	8790	8810	8830
1,2	8849	8869	8888	8907	8925	8944	8962	8980	8997	9015
1,3	9032	9049	9066	9082	9099	9115	9131	9147	9162	9177
1,4	9192	9207	9222	9236	9251	9265	9279	9292	9306	9319
1,5	9332	9345	9357	9370	9382	9394	9406	9418	9429	9441
1,6	9452	9463	9474	9484	9495	9505	9515	9525	9535	9545
1,7	9554	9564	9573	9582	9591	9599	9608	9616	9625	9633
1,8	9641	9649	9656	9664	9671	9678	9686	9693	9699	9706
1,9	9713	9719	9726	9732	9738	9744	9750	9756	9761	9767
2,0	9772	9778	9783	9788	9793	9798	9803	9808	9812	9817
2,1	9821	9826	9830	9834	9838	9842	9846	9850	9854	9857
2,2	9861	9864	9868	9871	9875	9878	9881	9884	9887	9890
2,3	9893	9896	9898	9901	9904	9906	9909	9911	9913	9916
2,4	9918	9920	9922	9925	9927	9929	9931	9932	9934	9936
2,5	9938	9940	9941	9943	9945	9945	9948	9949	9951	9952

Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$ (Normal distribution)

Định nghĩa:

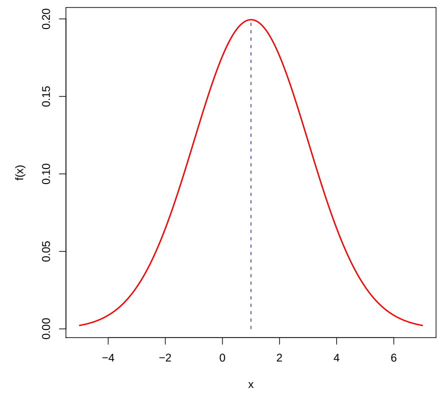
Cho biến ngẫu nhiên X liên tục, với $\sigma > 0$, μ là hai tham số, X có phân phối chuẩn, kí hiệu $X \sim \mathcal{N}(\mu,\sigma^2)$, khi hàm mật độ có dạng

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

Đặc trưng:

1. Kì vọng: $\mathbb{E}[X] = \mu$
2. Phương sai $\mathbb{V}(Z) = \sigma^2$

Phân phối chuẩn $\mathcal{N}(1,2^2)$



Hình: Hàm mật độ xác suất của $\mathcal{N}(1,4)$

Phân phối chuẩn - Phân phối chuẩn hóa

Định lý:

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

CM: 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 $\mathcal{N}(\mu,\sigma^2)$

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

Hệ quả 1:

Nếu $X \sim \mathcal{N}(\mu,\sigma^2)$ thì $P(X \le x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$.

Hệ quả 2:

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

Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$

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

VD 4.14 : 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?

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T\sim T(n)$
4.1.4 Phân phối Fisher $F\sim F(n,m)$

Xấp xỉ phân phối nhị thức bằng phân phối chuẩn

Định lý 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 $\mathbb{E}[X]=np$, $\sigma_X=\sqrt{np(1-p)}$.

Áp dụng: Định lý nói rằng khi n lớn ta có thể xấp xỉ phân phối nhị thức $B(n,p)$ bằng phân phối chuẩn $\mathcal{N}(np,np(1-p))$.

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T\sim T(n)$
4.1.4 Phân phối Fisher $F\sim F(n,m)$

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 $np(1-p)\geq 5$.

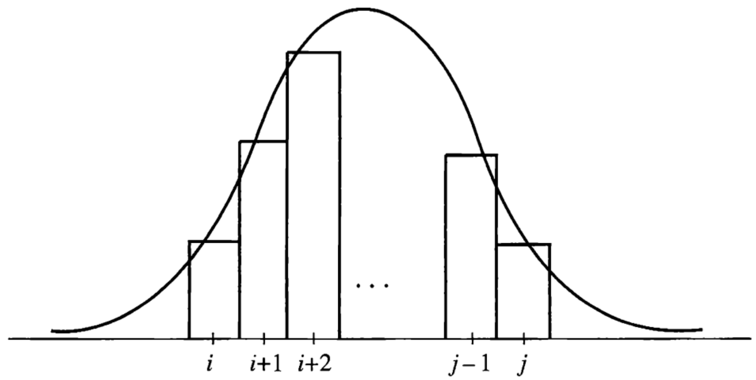
Hiệu chỉnh liên tục (Correction for continuity): Vì X trong phân phối nhị thức là rời rạc nên khi tính xấp xỉ các giá trị xác suất của X bằng phân phối chuẩn ta đã chuyển sang một biến mới liên tục nên trong thực hành phải thực hiện phép hiệu chỉnh liên tục như sau:

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

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T\sim T(n)$
4.1.4 Phân phối Fisher $F\sim F(n,m)$

Hiệu chỉnh liên tục



Xấp xỉ phân phối nhị thức bằng phân phối chuẩn

VD 4.12 : 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.

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n,m)$

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.

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n,m)$

Phân phối Gamma $G(\alpha,\beta)$

Hàm Gamma

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

Tính chất:

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n,m)$

Phân phối Gamma $G(\alpha,\beta)$

Định nghĩa:

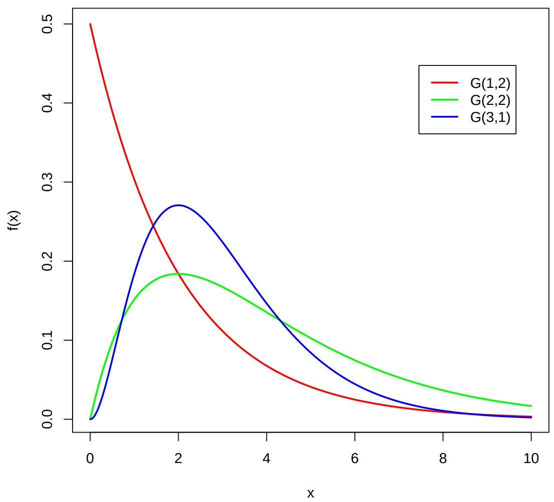
B.N.N X đgl có phân phối Gamma với hai tham số dương α và β, kí hiệu $X \sim G(\alpha,\beta)$, nếu hàm mật độ xác suất của X có dạng

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

Các đặc trưng

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

Phân phối Gamma



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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1, p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n, p)$
4.1.3 Phân phối siêu bội $H(N, M, n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha, \beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n, m)$

Phân phối Chi bình phương $\chi^2(r)$

Định nghĩa:

Phân phối Chi bình phương $X \sim \chi^2(r), r = 1, 2, \ldots,$ $X \sim \chi^2(r)$ nếu $X \sim G(r/2, 2)$,

Hàm mật độ xác suất của X

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

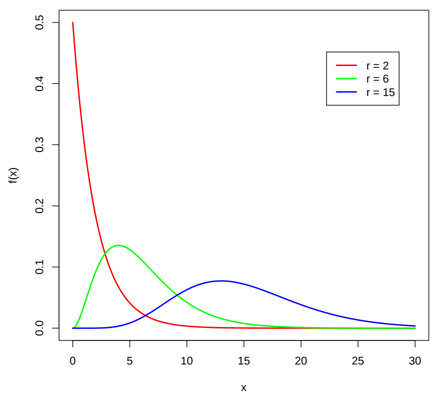
Các đặc trưng

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1, p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n, p)$
4.1.3 Phân phối siêu bội $H(N, M, n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha, \beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n, m)$

Phân phối Chi bình phương

Minh họa



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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1, p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n, p)$
4.1.3 Phân phối siêu bội $H(N, M, n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha, \beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n, m)$

Định lý:

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

CM: 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 $Y \sim \chi^2(1)$.

Định lý:

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

Hệ quả 1:

Nếu $X_i \sim \chi^2(r_i)$ với mọi $i = 1,\dots,n$ 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 2:

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

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T\sim T(n)$
4.1.4 Phân phối Fisher $F\sim F(n,m)$

Phân phối Student $T \sim T(n)$

Định nghĩa:

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T\sim T(n)$
4.1.4 Phân phối Fisher $F\sim F(n,m)$

Phân phối Student $T \sim T(n)$

Định lý:

B.N.N $T \sim T(n)$ có hàm mật độ

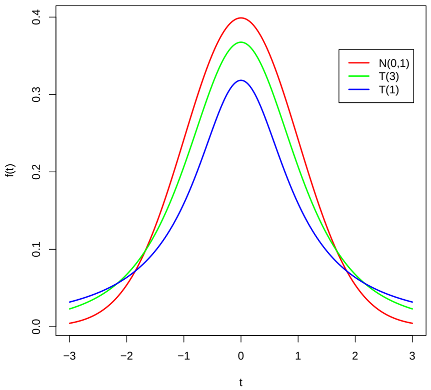
$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)} \cdot \frac{1}{\left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}}, \quad t \in \mathbb{R}$$

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

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc
4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1,p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n,p)$
4.1.3 Phân phối siêu bội $H(N,M,n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu,\sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha,\beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T\sim T(n)$
4.1.4 Phân phối Fisher $F\sim F(n,m)$

Phân phối Student

Minh họa



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

Phân phối Fisher $F \sim F(n, m)$

Định nghĩa:

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 lý:

B.N.N $F \sim F(n, m)$ có hàm mật độ

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

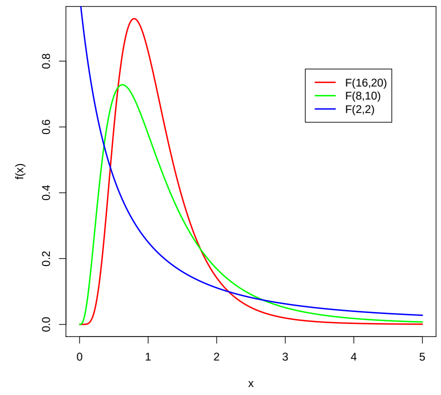
.

XÁC SUẤT THỐNG KÊ
N.T.M.Ngọc

4. Các phân phối xác suất thông dụng
4.1 Phân phối rời rạc
4.1.1 Phân phối Bernoulli $\mathcal{B}(1, p)$
4.1.2 Phân phối nhị thức $\mathcal{B}(n, p)$
4.1.3 Phân phối siêu bội $H(N, M, n)$
4.1.4 Phân phối Poisson $\mathcal{P}(\lambda)$
4.2 Phân phối liên tục
4.1.1 Phân phối đều
4.1.2 Phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$
4.1.3 Phân phối Gamma $G(\alpha, \beta)$
PP Chi bình phương $\chi^2(r)$
4.1.4 Phân phối Student $T \sim T(n)$
4.1.4 Phân phối Fisher $F \sim F(n, m)$

Phân phối Fisher

Minh họa



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