

BÀI THỰC HÀNH SỐ 4

1. Phân phối siêu bội được xấp xỉ bởi phân phối nhị thức

- *Biến ngẫu nhiên nhị thức*: Là số lần thành công của n phép thử Bernoulli với xác suất là:

$$\mathbb{P}(X = x) = C_n^x p^x (1 - p)^{n-x} \quad \text{với} \quad x = \{0, 1, \dots, n\}$$

- *Biến ngẫu nhiên siêu bội*: Ta xét tập hợp N đối tượng bao gồm K đối tượng được phân loại là “thành công” và $(N - K)$ đối tượng là “thất bại”. Lấy một mẫu ngẫu nhiên (không hoàn lại) gồm n đối tượng từ N đối tượng ban đầu ($K < N$ và $n < N$). Khi đó, biến ngẫu nhiên siêu bội là số đối tượng “thành công” trong mẫu và có xác suất là:

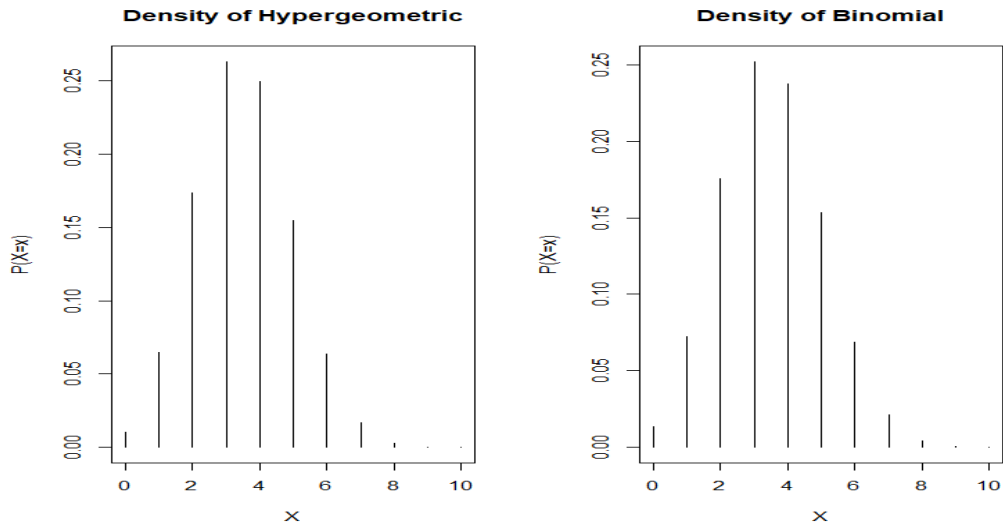
$$\mathbb{P}(X = x) = \frac{C_K^x C_{N-K}^{n-x}}{C_N^n} \quad \text{với} \quad x = \{\max(0, n + K - N), \dots, \min(K, n)\}$$

Khi n nhỏ hơn N rất nhiều thì hệ số hiệu chỉnh $\frac{N-n}{N-1}$ nhỏ và phân phối siêu bội $X \sim H(N, K, n)$ có thể được xấp xỉ bởi phân phối nhị thức $X \sim B(n, K/N)$

```
n = 10
N = 100
K = 35

par(mfrow=c(1,2))
plot(0:n,dhyper(0:n,K,N-K,n),type="h",main="Density of
Hypergeometric",xlab="X",ylab="P(X=x)")

plot(0:n,dbinom(0:n,n,K/N),type="h",main="Density of
Binomial",xlab="X",ylab="P(X=x)")
```



2. Phân phối nhị thức xấp xỉ bằng phân phối Poisson

- *Biến ngẫu nhiên Poisson*: Mô hình bằng phân phối Poisson thỏa các điều kiện: các biến cố xảy ra một cách ngẫu nhiên và độc lập giữa những khoảng thời gian hoặc không gian nhất định; số lượng trung bình λ các biến cố xảy ra trong những khoảng là đồng nhất và hữu hạn. Khi đó, xác suất số lượng các biến cố xảy ra trong những khoảng thời gian là:

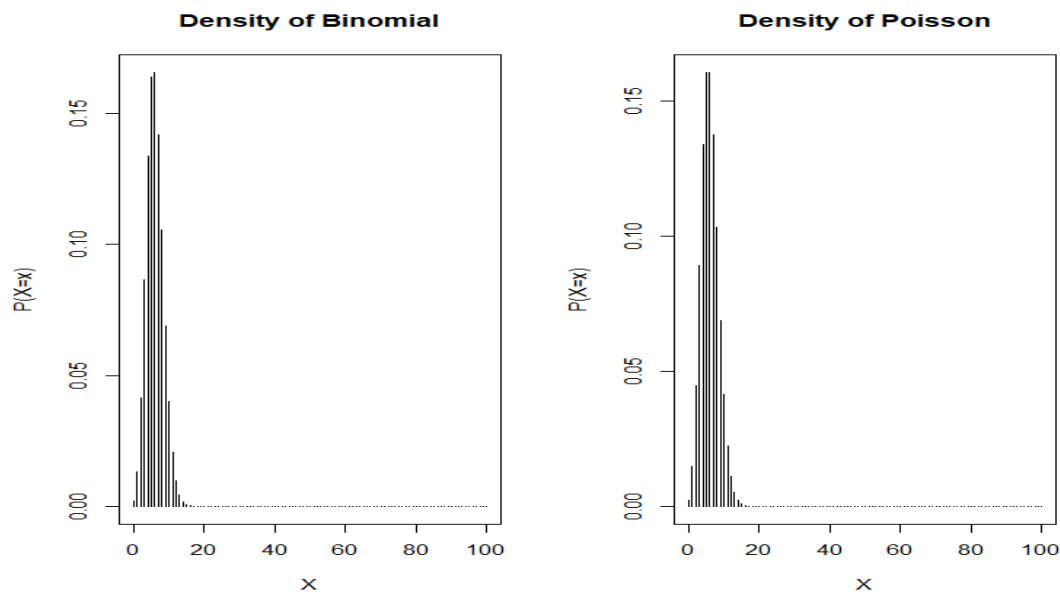
$$\mathbb{P}(X = x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad \text{với } x \in \mathbb{N}$$

Khi $X \sim B(n, p)$ với p nhỏ và n lớn thì X có thể được xấp xỉ bởi phân phối Poisson $X \sim P(np)$

```
size = 100
prob = 0.06

par(mfrow=c(1,2))
plot(0:size, dbinom(0:size, size, prob), type = "h", main =
"Density of Binomial", xlab = "X", ylab = "P(X=x)")

plot(0:size, dpois(0:size, size*prob), type = "h", main =
"Density of Poisson", xlab = "X", ylab = "P(X=x)")
```



3. Phân phối nhị thức xấp xỉ bằng phân phối chuẩn

- *Biến ngẫu nhiên chuẩn*: X được gọi là phân phối chuẩn nếu hàm mật độ xác suất có dạng

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

Nếu $\mu = 0$ và $\sigma^2 = 1$ thì $Z \sim \mathcal{N}(0,1)$ là biến ngẫu nhiên có *phân phối chuẩn tắc* và hàm phân phối tích lũy $\Phi(z) = \mathbb{P}(Z \leq z)$.

Khi $X \sim B(n, p)$ với np đủ lớn và $n(1 - p)$ đủ lớn thì X được xấp xỉ bởi phân phối chuẩn $X \sim \mathcal{N}(np, np(1 - p))$

```
size = 100
prob = 0.4

par(mfrow=c(1,2))
plot(0:size,dbinom(0:size,size,prob),type = "h",main =
"Density of Binomial",xlab = "X", ylab = "P(X=x)")

curve(dnorm(x,size*prob,size*prob*(1-prob)),from =
(size*prob)-100, to=(size*prob)+100,main = "Density of
Normal",xlab = "X", ylab = "f_X(x)")
```

