

## Tuần 6: LT Bài 5.pdf

- Điểm danh trên gg classroom của lớp.
- Mở R, thiết lập thư mục làm việc cho R.
- Mở file “Bai 5.pdf”: đọc và thực hành các ví dụ.

Thực hành xác suất thống kê

# Bài 5: PHÂN PHỐI MẪU VÀ ĐỊNH LÝ GIỚI HẠN TRUNG TÂM

# 1. Mẫu ngẫu nhiên phân phối chuẩn

## 1.1 Mô phỏng 1

**Định lý 1.1:** Cho  $X_1, X_2, \dots, X_n$  là mẫu ngẫu nhiên lấy từ phân phối chuẩn  $N(\mu; \sigma^2)$ . Ta có

$$\bar{X} \sim N\left(\mu; \frac{\sigma^2}{n}\right).$$

# 1. Mẫu ngẫu nhiên phân phối chuẩn

## 1.1 Mô phỏng 1

**Ý tưởng:**

**B1.** Tạo mẫu  $X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow$  tính  $\bar{X}$ .

**B2.** Lập lại bước trên  $m$  lần, ta được

$$X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow \text{tính } \bar{X}_1$$

$$X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow \text{tính } \bar{X}_2$$

...

$$X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow \text{tính } \bar{X}_m$$

Ta có bộ dữ liệu  $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_m$ .

**B3.** Kiểm tra phân phối

$$\bar{X} \sim N\left(\mu; \frac{\sigma^2}{n}\right)$$

# 1. Mẫu ngẫu nhiên phân phối chuẩn

## 1.1 Mô phỏng 1

**Ý tưởng:**

**B1.** Tạo mẫu  $X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow$  tính  $\bar{X}$ .

**B2.** Lập lại bước trên  $m$  lần, ta được

$$X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow \text{tính } \bar{X}_1$$

$$X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow \text{tính } \bar{X}_2$$

...

$$X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow \text{tính } \bar{X}_m$$

Ta có bộ dữ liệu  $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_m$ .

**B3.** Kiểm tra phân phối

$$\bar{X} \sim N\left(\mu; \frac{\sigma^2}{n}\right)$$

Đối với mẫu ngẫu nhiên lấy từ phân phối chuẩn, ta thành lập mẫu ngẫu nhiên kích thước  $n = 10$ , cụ thể ta sẽ lấy mẫu từ  $N(2; 2^2)$ , tức  $\mu = 2$  và  $\sigma = 2$ .

```
# Mô phỏng 1: Cho  $X_1, \dots, X_n \sim N(\mu, \sigma)$ . Khi  
# do  $\bar{X} \sim N(\mu, \sigma^2/n)$ 
```

```
# Gan các giá trị  
n = 10; mu = 2; sigma = 2
```

```
# Tạo vectơ  $X = (X_1, \dots, X_n) \sim N(\mu, \sigma)$   
VecX <- function(n) rnorm(n, mu, sigma)
```

```
# Tính giá trị trung bình mẫu  
MeanX <- function() mean(VecX(n))
```

```
# Lap lại quá trình trên m lần  
SampleMeanX <- function(m) replicate(m, MeanX())
```

```
# Vẽ biểu đồ tần số cho dữ liệu vừa tạo  
hist(SampleMeanX(1000), freq = 0, breaks = 40)
```

```
# Kiểm tra lại bằng cách vẽ hàm mật độ pp  
#  $N(\mu, \sigma^2/n)$   
curve(dnorm(x, mu, sigma/sqrt(n)), col = "blue", lty  
= 1, lwd = 2, add = T)
```

# 1. Mẫu ngẫu nhiên phân phối chuẩn

## 1.2 Mô phỏng 2

**Định lý 1.2:** Cho  $X_1, X_2, \dots, X_n$  là mẫu ngẫu nhiên lấy từ phân phối chuẩn  $N(\mu; \sigma^2)$ . Ta có

$$\frac{(n-1)S_X^2}{\sigma^2} \sim \chi^2(n-1).$$

# 1. Mẫu ngẫu nhiên phân phối chuẩn

## 1.2 Mô phỏng 2

**Ý tưởng:**

**B1.** Tạo mẫu  $X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow$  tính  $S = \frac{(n-1)S_X^2}{\sigma^2}$ .

**B2.** Lập lại bước trên  $m$  lần, ta được

$$X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow \text{tính } S_1$$

$$X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow \text{tính } S_2$$

...

$$X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow \text{tính } S_m$$

Ta có bộ dữ liệu  $S_1, S_2, \dots, S_m$ .

**B3.** Kiểm tra phân phối

$$S \sim \chi^2(n-1)$$



# 1. Mẫu ngẫu nhiên phân phối chuẩn

## 1.2 Mô phỏng 2

**Ý tưởng:**

**B1.** Tạo mẫu

$X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow$  tính  $S = \frac{(n-1)S_X^2}{\sigma^2}$ .

**B2.** Lập lại bước trên  $m$  lần, ta được

$X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow$  tính  $S_1$

$X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow$  tính  $S_2$

...

$X_1, X_2, \dots, X_n \sim N(\mu; \sigma^2) \rightarrow$  tính  $S_m$

Ta có bộ dữ liệu  $S_1, S_2, \dots, S_m$ .

**B3.** Kiểm tra phân phối

$$S \sim \chi^2(n-1)$$

Đối với mẫu ngẫu nhiên lấy từ phân phối chuẩn, ta thành lập mẫu ngẫu nhiên kích thước  $n = 10$ , cụ thể ta sẽ lấy mẫu từ  $N(2; 2^2)$ , tức  $\mu = 2$  và  $\sigma = 2$ .

```
# Mô phỏng 1.2: Cho  $X_1, \dots, X_n \sim N(\mu, \sigma)$ .
```

```
# Khi đó  $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$ 
```

```
# Tính giá trị biến  $S = (n-1) \cdot \text{var}(\text{VecX}(n)) / \sigma^2$ 
```

```
S <- function() (n-1)*var(VecX(n))/sigma^2
```

```
# Lap lai qua trinh tren m lan
```

```
SampleS <- function(m) replicate(m, S())
```

```
# Ve bieu do tan so cho du lieu vua tao
```

```
hist(SampleS(1000), freq = 0, breaks = 40)
```

```
# Ve ham mat do cua pp  $\chi^2(n-1)$ 
```

```
curve(dchisq(x, n-1), col = "red", lty = 1, lwd = 2, add = T)
```

## 2. Bài tập

**Bài 1.** Cho là mẫu ngẫu nhiên kích thước 2 lấy từ phân phối chuẩn  $N(0; 1)$ . Dùng hàm `rnorm()` phát sinh  $X_1, X_2$  và  $Y = X_1^2 + X_2^2$ . Xây dựng hàm `MauY` phát sinh mẫu ngẫu nhiên kích thước  $n$  cho  $Y$ . Lần lượt phát sinh mẫu ngẫu nhiên kích thước 100, 1000, 10000 cho  $Y$ , vẽ biểu đồ tần suất và đồ thị hàm mật độ xác suất của phân phối Chi-bình phương với 2 bậc tự do cho từng trường hợp.

Liên hệ với lý thuyết mẫu.

```
# Bai 1
# Viet ham tao Y
m = 0
s = 1
Y <- function(){
  X1 <- rnorm(1,m,s)
  X2 <- rnorm(1,m,s)
  Y = (X1)^2 + (X2)^2
}
MauY <- function(n) replicate(n,Y())
draw <- function(m){
  # Ve bieu do tan suat
  hist(MauY(m),freq=0,breaks=40)
  # Ve ham mat do cua phan phoi Chi binh phuong voi 2
  #bac tu do
  # vao cung 1 bieu do (add = TRUE)
  curve(dchisq(x,2),col="blue",lty=1,lwd=2,add=TRUE)}
par(mfrow = c(1,3))
draw(100)
draw(1000)
draw(10000)
```

## 2. Bài tập

```
# Bai 1
# Viet ham tao MauY
m = 0
s = 1
Y <- function(){
  X1 <- rnorm(1,m,s)
  X2 <- rnorm(1,m,s)
  Y = (X1)^2 + (X2)^2
}
MauY <- function(n) replicate(n,Y())
draw <- function(m){
# Ve bieu do tan suat
hist(MauY(m),freq=0,breaks=40)
# Ve ham mat do cua phan phoi Chi binh phuong voi 2 bac tu do
# vao cung 1 bieu do (add = TRUE)
curve(dchisq(x,2),col="blue",lty=1,lwd=2,add=TRUE)}
par(mfrow = c(1,3))
draw(100)
draw(1000)
draw(10000)
```

### 3. Bài tập – Bài 6

Bài 1. Tạo vec-tơ:  $x = [1, 2, 5, 7, -3, 0, 5, 1, 5, 6]$  và  $y = [2, 2, 0, -5, 7, 8, 11, 9, 3, 2]$

- Tính  $x+y$ ,  $x*y$ ,  $x-y$ .
- Tạo  $z = [\text{Những phần tử chẵn của } x]$ ,  $t = [\text{Những phần tử lẻ của } y]$
- Trích những phần tử lớn hơn 0 của  $x$  và  $y$ .
- Tính trung bình, độ lệch tiêu chuẩn, sai số chuẩn của  $x$  và  $y$ .
- Tìm phần tử lớn nhất, bé nhất của  $x$ ,  $y$ .
- Sắp xếp  $x$  tăng dần,  $y$  giảm dần.
- Lưu  $x$  và  $y$ .

### 3. Bài tập – Bài 6

```
# Bai 1
# Tao mau x,y
x <- c(1,2,5,7,-3,0,5,1,5,6)
y <- c(2,2,0,-5,7,8,11,9,3,2)
# a. tinh tong, tích, hieu
x+y;x*y; x-y
# b. Tach gia tri cua x chan gan vao z, gia tri cua y le cho t
z <- x[x%%2 == 0];z
t <- y[y%%2 == 1];t
# c. Tach gia tri duong cua x, y
x[x>0]; y[y>0]
# d.
# trung binh mau cua x
mean(x)
# do lech tieu chuan cua x
sd(x)
# sai so chuan cua x
sd(x)/sqrt(length(x))
```