

### Bài 3.

## Biến ngẫu nhiên

### 1. Tính toán

#### Ví dụ 1

a) Với  $k = 0, 1, \dots, 8$ , tính các xác suất  $P(X = k) = C_8^k 0.3^k 0.7^{8-k}$ .

b) Cho biến ngẫu nhiên  $X$  có hàm mật độ như sau:

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

với  $\mu$  và  $\sigma > 0$  là hai tham số. Viết hàm pdf của  $X$  để tính giá trị  $f(x)$  tại một điểm bất kỳ.

Giải:

a)

```
k = 0:8;
p = function(k) choose(8,k) * 0.3^k * 0.7^(8-k);
p(k)
[1] 0.05764801 0.19765032 0.29647548 0.25412184 0.13613670
0.04667544
[7] 0.01000188 0.00122472 0.00006561
```

Ta có thể kiểm tra lại rằng tổng các xác suất bằng 1 với lệnh

```
sum(p(k))
[1] 1
```

b)

```
f = function(x, mu=0, sigma=1){
  1/sqrt(2*pi*sigma^2) * exp(-(x-mu)^2/(2*sigma^2))
}
```

Kiểm tra tích phân có bằng 1?

```
integrate(function(x) f(x, 0, 1), lower=-Inf, upper=Inf)
1 with absolute error < 9.4e-05
```

Tính tại điểm  $x = 0$ :

```
f(0)
[1] 0.3989423
```

### 2. Biểu diễn bằng đồ thị

#### Ví dụ 2

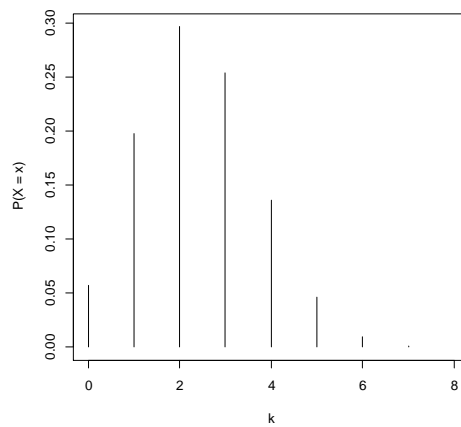
a) Vẽ đồ thị hàm xác suất trong Ví dụ 1a):  $P(X = k) = C_8^k 0.3^k 0.7^{8-k}$  với  $k = 0, 1, \dots, 8$ .

b) Vẽ hàm mật độ xác suất trong Ví dụ 1b):  $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty$

Giải:

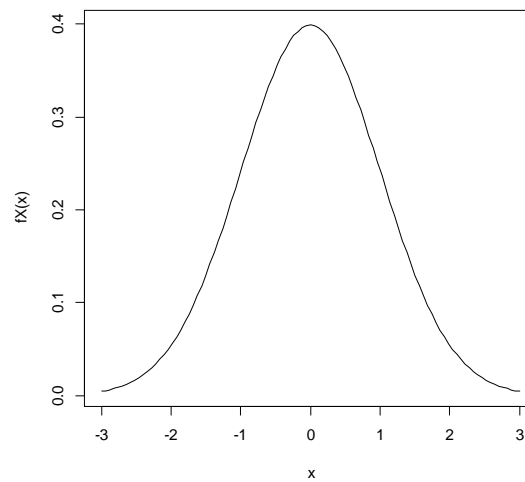
a)

```
plot(k, p(k), type = "h", ylab = "P(X = x)")
```



b)

```
curve(f(x,0,1),from=-3,to=3, ylab = "fX(x)")
```



### 3. Hàm phân phối

#### 3.1 Định nghĩa

$$F_X(x) = P(X \leq x)$$

\* X rời rạc,

$$F_X(x) = \sum_{k \in X(\Omega) \cap (-\infty, x]} P(X = k)$$

\* X liên tục,

$$F_X(x) = \int_{-\infty}^x f_X(t) dt$$

#### 3.2 Tính toán

##### Ví dụ 3

a) Xét tiếp Ví dụ 1a). Tính  $F_X(4) = P(X \leq 4) = \sum_{k=0}^4 P(X = k)$ .

b) Xét tiếp Ví dụ 1b). Tính  $F_X(1.96) = P(X \leq 1.96) = \int_{-\infty}^{1.96} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$ .

Giải:

a)

```
F = function(k) sum(p(0:k))
F = Vectorize(F)
F(4)
[1] 0.9420323
```

b)

```
F2 = function(a,mu = 0, sigma = 1){
  integrate(function(x) f(x,mu,sigma), lower = -Inf, upper =
a)$value
}
F2 = Vectorize(F2)
F2(1.96)
0.9750021
```

### 3.3 Vẽ đồ thị

Ví dụ 4

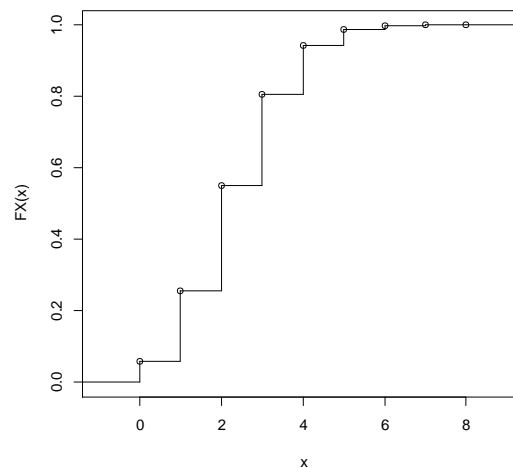
a) Xét tiếp Ví dụ 1a). Vẽ hàm phân phối xác suất của  $X$ .

b) Xét tiếp Ví dụ 1b). Vẽ hàm phân phối.

Giải:

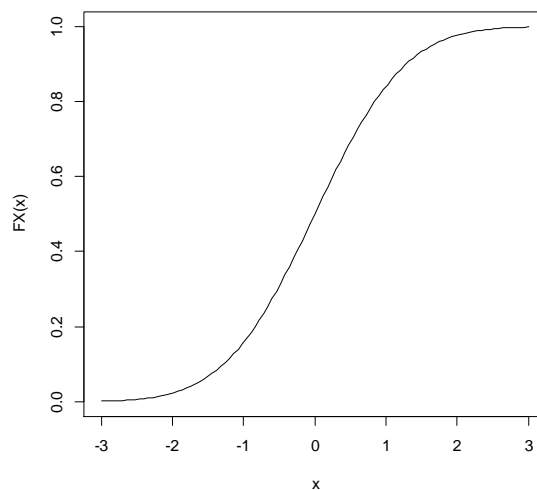
a)

```
plot(stepfun(k, c(0, F(k))), ylab = "FX(x)", main = "")
```



b)

```
curve(F2(x), from = -3, to = 3, ylab = "FX(x)")
```



## 4. Phân vị

### 4.1 Định nghĩa

Cho  $p \in (0,1)$  và  $X$  là một biến ngẫu nhiên.

\* Nếu  $X$  rời rạc, thì phân vị mức  $p$  của  $X$ , ký hiệu là  $x_p$ , là

$$x_p = \inf\{k \in \mathbb{Z}; F_X(k) \geq p\}$$

\* Nếu  $X$  liên tục, thì phân vị mức  $p$  của  $X$ , ký hiệu là  $x_p$ , là giá trị thỏa  $F_X(x_p) = p$ .

### 4.2 Tính toán

Ví dụ 5

a) Xét tiếp Ví dụ 1a). Tính phân vị mức 0.25 của  $X$ .

b) Xét tiếp Ví dụ 1b). Tính phân vị mức 0.975 của  $X$ .

Giải:

a)

```
K = k[F(k) >= 0.25];
```

```
K[1]
```

```
[1] 1
```

Kiểm tra lại bằng lệnh:

```
F(0)
```

```
[1] 0.05764801
```

```
F(1)
```

```
[1] 0.2552983
```

b)

```
uniroot(function(x) F2(x)-0.975, c(-3,3))$root
```

```
[1] 1.959992
```

Kiểm tra lại bằng lệnh

```
F2(1.959992)
```

```
[1] 0.9750016
```

## 5. Mô phỏng

### 5.1 Lệnh sample

a) Sử dụng cơ bản

Lệnh cơ bản là: `sample(x)`, với `x` là vector số, logic hoặc chuỗi ký tự.

Lệnh này trả về một vector có các phần tử được rút ra từ `x` mà không hoàn lại.

Tạo một hoán vị các phần tử của `x` ta làm:

```
x = 1:7
sample(x)
[1] 2 4 1 7 5 6 3
```

Ta chạy lại cùng lệnh này:

```
sample(x)
[1] 1 7 6 2 3 4 5
```

Ta nhận được các kết quả khác nhau; thứ tự của hoán vị thay đổi mỗi lần chạy.

Ta có thể giữ lại kết quả của một phép hoán vị trong một vector để sử dụng sau này:

```
y = sample(c("red", "green", "blue", "white", "black"))
y
[1] "black" "white" "green" "red," "blue"
```

### 5.2 Các tùy chọn

Có một số tùy chọn trong lệnh `sample` để thay đổi mẫu cho phù hợp với thực tế. Nó được làm bằng cách thêm một hoặc vài tham biến trong lệnh `sample`.

Ví dụ:

```
sample(1:3, size = 2, replace = TRUE, prob = c(25 / 100, 20 / 100, 55 / 100))
```

Các tùy chọn được trình bày bên dưới

a) *Size (cỡ mẫu)*: `size = "n"`, với `n` là một số nguyên, xác định số lần `n` ta thực hiện rút ra. Mặc định, số được rút ra bằng chiều dài `x`.

```
sample(1:10, size = 3)
[1] 10 7 3
```

Trong cú pháp `sample`, tham biến `size` là thứ hai. Ta có thể viết tắt như sau :

```
sample(1:10, 3).
```

b) *Replace (hoàn lại)*: `replace = L`, với `L` là `TRUE` hoặc `FALSE`, trong đó rút có thay thế (hoàn lại) nếu `L = TRUE` và không hoàn lại nếu khác.

Ta làm :

```
sample(1:5, size = 9, replace = TRUE)
[1] 5 1 4 1 3 1 2 1 3
```

c) *Prob (xác suất)* : Tham biến `prob` cho phép ta gán các giá trị xác suất cho từng giá trị trong vector `x`.

$$P(X = x_j) = p_j, j \in \{1, \dots, k\}, p_j \in [0, 1], \sum_{j=1}^k p_j = 1.$$

Do đó, `prob = vec`, với `vec` là vector chứa các xác suất  $(p_1, \dots, p_k)$ .

Nếu `prob` không được xác định, mặc định  $p_j = 1/k$  (các giá trị bằng nhau).

Ta làm:

```
y = c("red", "green", "blue", "white", "black")
sample(y, 2, prob = c(10 / 100, 30 / 100, 10 / 100, 30 / 100, 20 / 100))
[1] "red" "green"
```

## Bài tập

### Bài 1

1. Biến ngẫu nhiên  $P$ , nhận các giá trị giữa 0 và 1, có hàm mật độ xác suất  $f_p(p) = 0.07p^{-0.93}$ .

a) Tính xác suất  $P(P \leq 0.2)$  bằng cách sử dụng hàm: 1)  $f = \text{function}(p)\{.07*p**(-0.93)\}$ , và 2)  $\text{integrate}(f, \text{lower}=0, \text{upper}=.2)$ .

b) Kiểm tra  $f_p(p)$  là hàm mật độ xác suất bằng cách tính diện tích dưới đường cong từ 0 đến 1.

### Bài 2

Sử dụng lệnh `x=sample(1:5,100,TRUE,c(0.1,0.2,0.4,0.2,0.1))` để rút một mẫu ngẫu nhiên cỡ 100 từ phân phối với hàm xác suất

x	1	2	3	4	5
P(x)	0.1	0.2	0.4	0.2	0.1

và chứa mẫu này trong `x`. Tiếp theo vẽ biểu đồ cột của các xác suất thực nghiệm (mẫu) bằng lệnh `table(x)/100`