

# THỐNG KÊ MÁY TÍNH

(Computational Statistics)

Trường Đại học Nha Trang

Khoa Công nghệ thông tin

Bộ môn Hệ thống thông tin

Giảng viên: TS.Nguyễn Khắc Cường

# CHƯƠNG 4

## XÁC SUẤT

## 4.1. Khái niệm xác suất cơ bản

- Deterministic experiment (Phép thử xác định)
  - Phép thử chắc chắn dự đoán đúng kết quả
  - VD:
    - Phép cộng  $2+3$
    - Phản ứng  $H_2 + O_2$
    - ...
- Random experiment (Phép thử ngẫu nhiên)
  - Phép thử không thể chắc chắn dự đoán đúng kết quả
  - VD:
    - Thả đồng xu
    - Thả xúc xắc
    - ...

## 4.1. Khái niệm xác suất cơ bản

- Simple event (Biến cố sơ cấp)
  - Là 1 kết quả xảy ra khi 1 lần thực hiện phép thử
  - Biểu diễn bằng tập hợp có 1 phần tử
  - VD:
    - Thả đồng xu:  $\{1\}, \{2\}$
    - Thả xúc xắc:  $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$
- Random event (Biến cố ngẫu nhiên)
  - Là kết quả có thể xảy ra hoặc không xảy ra khi thực hiện phép thử
  - Là tập hợp các biến cố sơ cấp có cùng một đặc tính
- VD: khi thực hiện phép thử “thả xúc xắc”
  - Biến cố A: “xuất hiện mặt chẵn”  $\rightarrow B = \{2, 4, 6\}$

## 4.1. Khái niệm xác suất cơ bản

- Sample space (Không gian mẫu)
  - Tập hợp tất cả biến cố sơ cấp có thể xảy ra của phép thử
  - VD:
    - Thả 1 đồng xu có 2 mặt H (head), T (tail)  
→ sample space  $S = \{H, T\}$
    - Thả 2 đồng xu có 2 mặt H (head), T (tail)  
→ sample space  $S = \{(H1, H2), (H1, T2), (T1, H2), (T1, T2)\}$
  - VD: xem sample space trong R
    - Cài đặt package “prob”
    - Lệnh

```
> tosscoin(1)
toss1
1      H
2      T
```

```
> tosscoin(3)
toss1 toss2 toss3
1      H      H      H
2      T      H      H
3      H      T      H
4      T      T      H
5      H      H      T
6      T      H      T
7      H      T      T
8      T      T      T
```

## 4.2. Quan hệ giữa các biến cố ngẫu nhiên

Các biến cố ngẫu nhiên  $A, B, C, \dots$  thuộc phép thử nào đó, chúng có thể có các mối quan hệ sau:

- Biến cố đồng nhất:
  - $A = B$
  - Khi thực hiện phép thử,  $A$  và  $B$  có thể
    - cùng xảy ra
    - hoặc không cùng xảy ra
- Biến cố đối lập:
  - $A^c$  hoặc  $\bar{A}$
  - Khi  $A$  xảy ra thì  $A^c$  không xảy ra

## 4.2. Quan hệ giữa các biến cố ngẫu nhiên


- Biến cố tích:
  - $A \cap B$  hoặc  $AB$
  - $A \cap B$  chỉ xảy ra khi cả A và B đồng thời xảy ra
  - $\bigcap_{i=1}^n A_i$  hoặc  $A_1 A_2 \dots A_n$  hoặc  $\prod_{i=1}^n A_i$
  - $\bigcap_{i=1}^n A_i$  xảy ra khi tất cả n biến cố thành phần xảy ra
- Biến cố xung khắc:
  - $A \cap B = \emptyset$
  - A và B không đồng thời xảy ra trong một phép thử
- Biến cố độc lập:
  - Biến cố A xuất hiện không ảnh hưởng gì đến biến cố B
  - và ngược lại



## 4.2. Quan hệ giữa các biến cố ngẫu nhiên

- Biến cố tổng:
  - $A \cup B$  hoặc  $A+B$
  - $A \cup B$  xảy ra khi có ít nhất một biến cố A hoặc B xảy ra
  - $\bigcup_{i=1}^n A_i$  hoặc  $A_1 + A_2 + \dots + A_n$  hoặc  $\sum_{i=1}^n A_i$  là biến cố tổng mở rộng, chỉ xảy ra khi tất cả n biến cố xảy ra
- Biến cố kéo theo:
  - $A \subset B$
  - A xảy ra kéo theo B cũng xảy ra
- Biến cố chắc chắn:
  - $\Omega$ .
  - Biến cố chắc chắn xảy ra khi thực hiện phép thử



## 4.2. Quan hệ giữa các biến cố ngẫu nhiên

- Hệ đầy đủ các biến cố
    - $\{A_1, A_2, \dots, A_n\}$  hoặc  $\sum_{i=1}^n A_i = \Omega$ .
    - Các biến cố từng đôi một xung khắc
  - VD:
    - Thực hiện phép thử thả 2 xúc xắc đều, đồng chất, các biến cố ngẫu nhiên A, B, C, D, E có thể xảy ra như sau:
      - A: “Tổng các nút xuất hiện trên 2 xúc xắc là chẵn”
      - B: “Tổng các nút xuất hiện trên 2 xúc xắc là lẻ”
      - C: “Số nút xuất hiện trên mỗi xúc xắc là lẻ” 
      - D: “Số nút xuất hiện trên mỗi xúc xắc là chẵn”
      - E: “Số nút xuất hiện trên 2 xúc xắc là cùng chẵn hoặc cùng lẻ”
- Một số mối quan hệ giữa các biến cố là

$$A=E, A^c=B, AB = \emptyset, A=C+D, D \subset A, ..$$


## 4.2. Quan hệ giữa các biến cố ngẫu nhiên

- Dùng R thực hiện các phép toán quan hệ tập hợp

- Tạo tập con:

- Cú pháp: `subset(<tập cha>, <điều kiện>)`

- VD:

- Khởi động package “prob”: `> library(prob)`

- Tạo tập S

```
> S<-tosscoin(3)
> S
      toss1 toss2 toss3
1         H      H      H
2         T      H      H
3         H      T      H
4         T      T      H
5         H      H      T
6         T      H      T
7         H      T      T
8         T      T      T
```

## 4.2. Quan hệ giữa các biến cố ngẫu nhiên

- Dùng R thực hiện các phép toán quan hệ tập hợp

- Tạo tập con:

- VD: Tạo tập con của S

```
> S_sub<-subset(S, toss1=="T")
> S_sub
  toss1 toss2 toss3
2      T      H      H
4      T      T      H
6      T      H      T
8      T      T      T
```

```
> S<-tosscoin(3)
> S
  toss1 toss2 toss3
1      H      H      H
2      T      H      H
3      H      T      H
4      T      T      H
5      H      H      T
6      T      H      T
7      H      T      T
8      T      T      T
```

- Tìm tập con:

- Cú pháp: <vector 1> %in% <vector 2>

- Tìm phần tử nào của vector 1 có trong vector 2

- VD:

```
> x<-1:10
> x
[1] 1 2 3 4 5 6 7 8 9 10
> y<-8:12
> y
[1] 8 9 10 11 12
> y %in% x
[1] TRUE TRUE TRUE FALSE FALSE
```

## 4.2. Quan hệ giữa các biến cố ngẫu nhiên

- Dùng R thực hiện các phép toán quan hệ tập hợp

- Tìm tập con:

- Cú pháp: `isin (<vector 1> ,<vector 2>)`
- Cho biết vector 2 có phải là vector con của vector 1 không

- VD:

```
> a<-2:7
> a
[1] 2 3 4 5 6 7
> b<-1:16
> b
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
> isin(a,b)
[1] FALSE
> isin(b,a)
[1] TRUE
```

- Cú pháp tương đương

```
> all(b%in%a)
[1] FALSE
> all(a%in%b)
[1] TRUE
```

## 4.2. Quan hệ giữa các biến cố ngẫu nhiên

- Dùng R thực hiện các phép toán quan hệ tập hợp
- Hợp, giao, hiệu 2 tập hợp:

- Cú pháp:

Tên hàm	Ký hiệu	Cú pháp trong R
Hợp	$A \cup B$	<code>union(A,B)</code>
Giao	$A \cap B$	<code>intersect(A,B)</code>
Hiệu	$A \setminus B$	<code>setdiff(A,B)</code>

- VD:
  - Tạo các tập hợp A, B
  - Thực hiện các hàm trên

## 4.2. Quan hệ giữa các biến cố ngẫu nhiên

- Dùng R thực hiện các phép đếm

- Giai thừa = Số lượng hoán vị

- Cú pháp: `factorial(n)` hay `prod(n:1)`

- VD: tính 3!

```
> factorial(3)
[1] 6
> prod(3:1)
[1] 6
```



- Số lượng tổ hợp:

- Tổ hợp:

- Tổ hợp chập k của một tập hợp n phần tử là một tập con k phần tử của tập n phần tử đã cho

- Số lượng tổ hợp:

$$C_n^k = \frac{n!}{k!(n-k)!}$$

## 4.2. Quan hệ giữa các biến cố ngẫu nhiên

- Dùng R thực hiện các phép đếm

- Số lượng tổ hợp:

- Số lượng tổ hợp:

$$C_n^k = \frac{n!}{k! (n - k)!}$$

- Hàm:

- choose(n,k)

- VD:

```
> choose(5,2)
[1] 10
```

Q / A