Bài giảng xác suất thống kê

Chương 1

NHỮNG KHÁI NIỆM CƠ BẢN VỀ XÁC SUẤT

1. BỔ TÚC VỀ GIẢI TÍCH TỔ HỢP

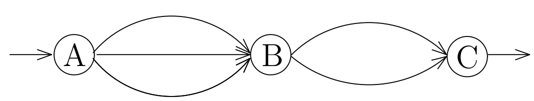
1.1 Qui tắc nhân

Giả sử một công việc nào đó được chia thành k giai đoạn. Có n_1 cách thực hiện giai đoạn thứ nhất, n_2 cách thực hiện giai đoạn thứ hai,...,n_k cách thực hiện giai đoạn thứ k. Khi đó ta có

$$n = n_1.n_2 \dots n_k$$

cách thực hiện công việc.

• **Ví dụ 1** *Giả sử để đi từ A đến C ta bắt buộc phải đi qua điểm B. Có 3 đường khác nhau để đi từ A đến B và có 2 đường khác nhau để đi từ B đến C. Vậy có $n = 3.2$ cách khác nhau để đi từ A đến C.*



1.2 Chỉnh hợp

□ **Định nghĩa 1** *Chỉnh hợp chập k của n phần tử ($k \leq n$) là một nhóm (bộ) có thứ tự gồm k phần tử khác nhau chọn từ n phần tử đã cho.*

Số chỉnh hợp chập k của n phần tử kí hiệu là A_n^k.

⊙ **Công thức tính:** $\boxed{A_n^k = \dfrac{n!}{(n-k)!} = n(n-1)\dots(n-k+1)}$

• **Ví dụ 2** *Một buổi họp gồm 12 người tham dự. Hỏi có mấy cách chọn một chủ tọa và một thư ký?*

Giải

Mỗi cách chọn một chủ tọa và một thư ký từ 12 người tham dự buổi họp là một chỉnh hợp chập k của 12 phần tử.

Do đó số cách chọn là $A_{12}^2 = 12.11 = 132$.

• **Ví dụ 3** *Với các chữ số 0,1,2,3,4,5 có thể lập được bao nhiêu số khác nhau gồm 4 chữ số.*

Giải

Các số bắt đầu bằng chữ số 0 (0123, 0234,...) không phải là số gồm 4 chữ số.

Chữ số đầu tiên phải chọn trong các chữ số 1,2,3,4,5. Do đó có 5 cách chọn chữ số đầu tiên.

Ba chữ số kế tiếp có thể chọn tùy ý trong 5 chữ số còn lại. Có A_5^3 cách chọn.

Vậy số cách chọn là $5.A_5^3 = 5.(5.4.3) = 300$

1.3 Chỉnh hợp lặp

□ **Định nghĩa 2** *Chỉnh hợp lặp chập k của n phần tử là một nhóm có thứ tự gồm k phần tử chọn từ n phần tử đã cho, trong đó mỗi phần tử có thể có mặt 1,2,...,k lần trong nhóm.*

Số chỉnh hợp lặp chặp k của n phần tử được kí hiệu B_n^k.

⊙ **Công thức tính**

$$\boxed{B_n^k = n^k}$$

• **Ví dụ 4** *Xếp 5 cuốn sách vào 3 ngăn. Hỏi có bao nhiêu cách xếp ?*

Giải

Mỗi cách xếp 5 cuốn sách vào 3 ngăn là một chỉnh hợp lặp chập 5 của 3 (Mỗi lần xếp 1 cuốn sách vào 1 ngăn xem như chọn 1 ngăn trong 3 ngăn. Do có 5 cuốn sách nên việc chọn ngăn được tiến hành 5 lần).

Vậy số cách xếp là $B_3^5 = 3^5 = 243$.

1.4 Hoán vị

□ **Định nghĩa 3** *Hoán vị của m phần tử là một nhóm có thứ tự gồm đủ mặt m phần tử đã cho.*

Số hoán vị của m phần tử được kí hiệu là P_m.

⊙ **Công thức tính**

$$\boxed{P_m = m!}$$

• **Ví dụ 5** *Một bàn có 4 học sinh. Hỏi có mấy cách xếp chỗ ngồi ?*

Giải

Mỗi cách xếp chỗ của 4 học sinh ở một bàn là một hoán vị của 4 phần tử. Do đó số cách xếp là $P_4 = 4! = 24$.

1.5 Tổ hợp

□ **Định nghĩa 4** *Tổ hợp chập k của n phần tử ($k \leq n$) là một nhóm không phân biệt thứ tự, gồm k phần tử khác nhau chọn từ n phần tử đã cho.*

Số tổ hợp chập k của n phần tử kí hiệu là C_n^k.

⊙ **Công thức tính**

$$\boxed{C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{k!}}$$

⊙ **Chú ý**

i) Qui ước $0! = 1$.
ii) $C_n^k = C_n^{n-k}$.
iii) $C_n^k = C_{n-1}^{k-1} + C_{n-1}^k$.

• **Ví dụ 6** *Mỗi đề thi gồm 3 câu hỏi lấy trong 25 câu hỏi cho trước. Hỏi có thể lập nên bao nhiêu đề thi khác nhau ?*

Giải

Số đề thi có thể lập nên là $C_{25}^3 = \dfrac{25!}{3!.(22)!} = \dfrac{25.24.23}{1.2.3} = 2.300$.

• **Ví dụ 7** *Một máy tính có 16 cổng. Giả sử tại mỗi thời điểm bất kỳ mỗi cổng hoặc trong sử dụng hoặc không trong sử dụng nhưng có thể hoạt động hoặc không thể hoạt động. Hỏi có bao nhiêu cấu hình (cách chọn) trong đó 10 cổng trong sử dụng, 4 không trong sử dụng nhưng có thể hoạt động và 2 không hoạt động?*

Giải

Để xác định số cách chọn ta qua 3 bước:

Bước 1: Chọn 10 cổng sử dụng: có $C_{16}^{10} = 8008$ cách.

Bước 2: Chọn 4 cổng không trong sử dụng nhưng có thể hoạt động trong 6 cổng còn lại: có $C_6^4 = 15$ cách.

Bước 3: Chọn 2 cổng không thể hoạt động: có $C_2^2 = 1$ cách.

Theo qui tắc nhân, ta có $C_{16}^{10}.C_6^4.C_2^2 = (8008).(15).(1) = 120.120$ cách.

1.6 Nhị thức Newton

Ở phổ thông ta đã biết các hằng đẳng thức đáng nhớ

$$\begin{aligned} a+b &= a^1 + b^1 \\ (a+b)^2 &= a^2 + 2a^1b^1 + b^2 \\ (a+b)^3 &= a^3 + 3a^2b^1 + 3a^1b^2 + b^3 \end{aligned}$$

Các hệ số trong các hằng đẳng thức trên có thể xác định từ tam giác Pascal

$$\begin{array}{ccccc} 1 & 1 & & & \\ 1 & 2 & 1 & & \\ 1 & 3 & 3 & 1 & \\ 1 & 4 & 6 & 4 & 1 \end{array}$$

$$C_n^0 \quad C_n^1 \quad C_n^2 \quad C_n^3 \quad C_n^4 \quad \ldots \quad C_n^{n-1} \quad C_n^n$$

Newton đã chứng minh được công thức tổng quát sau (*Nhị thức Newton*):

$$\begin{aligned}(a+b)^n &= \sum_{k=o}^{n} C_n^k a^{n-k} b^k \\ &= C_n^0 a^n + C_n^1 a^{n-1} b + C_n^2 a^{n-2} b^2 + \ldots + C_n^k a^{n-k} b^k + \ldots + C_n^{n-1} a b^{n-1} + C_n^n b^n\end{aligned}$$

(a,b là các số thực; n là số tự nhiên)

2. BIẾN CỐ VÀ QUAN HỆ GIỮA CÁC BIẾN CỐ

2.1 Phép thử và biến cố

Việc thực hiện một nhóm các điều kiện cơ bản để quan sát một hiện tượng nào đó được gọi một phép thử. Các kết quả có thể xảy ra của phép thử được gọi là biến cố (sự kiện).

• **Ví dụ 8**

i) Tung đồng tiền lên là một phép thử. Đồng tiền lật mặt nào đó (xấp, ngửa) là một biến cố.

ii) Bắn một phát súng vào một cái bia là một phép thử. Việc viên đạn trúng (trật) bia là một biến cố.

2.2 Các biến cố và quan hệ giữa các biến cố

i) Quan hệ kéo theo

Biến cố A được gọi là kéo theo biến cố B, kí hiệu $A \subset B$, nếu A xảy ra thì B xảy ra.

ii) Quan hệ tương đương

Hai biến cố A và B được gọi là tương đương với nhau nếu $A \subset B$ và $B \subset A$, kí hiệu $A = B$.

iii) Biến cố sơ cấp

Biến cố sơ cấp là biến cố không thể phân tích được nữa được nưa.

iv) Biến cố chắc chắn

Là biến cố nhất định sẽ xảy ra khi thực hiện phép thử. Kí hiệu Ω.

• **Ví dụ 9** *Tung một con xúc xắc. Biến cố mặt con xúc xắc có số chấm bé hơn 7 là biến cố chắc chắn.*

v) Biến cố không thể

Là biến cố nhất định không xảy ra khi thực hiện phép thử. Kí hiệu $\emptyset$.

⊕ **Nhận xét** Biến cố không thể $\emptyset$ không bao hàm một biến cố sơ cấp nào, nghĩa là không có biến cố sơ cấp nào thuận lợi cho biên cố không thể.

vi) Biến cố ngẫu nhiên

Là biến cố có thể xảy ra hoặc không xảy ra khi thực hiện phép thử. Phép thử mà các kết quả của nó là các biến cố ngẫu nhiên được gọi là phép thử ngẫu nhiên.

vii) Biến cố tổng

Biến cố C được gọi là tổng của hai biến cố A và B, kí hiệu $C = A + B$, nếu C xảy ra khi và chỉ khi ít nhất một trong hai biến cố A và B xảy ra.

• **Ví dụ 10** *Hai người thợ săn cùng bắn vào một con thú. Nếu gọi A là biến cố người thứ nhất bắn trúng con thú và B là biến cố người thứ hai bắn trúng con thú thì $C = A+B$ là biến cố con thú bị bắn trúng.*

⊙ **Chú ý**

i) Mọi biến cố ngẫu nhiên A đều biểu diễn được dưới dạng tổng của một số biến cố sơ cấp nào đó. Các biến cố sơ cấp trong tổng này được gọi là *các biến cố thuận lợi* cho biến cố A.

ii) Biến cố chắc chắn Ω là tổng của mọi biến cố sơ cấp có thể, nghĩa là mọi biến cố sơ cấp đều thuận lợi cho Ω. Do đó Ω còn được gọi là *không gian các biến cố sơ cấp.*

• **Ví dụ 11** *Tung một con xúc xắc. Ta có 6 biến cố sơ cấp $A_1, A_2, A_3, A_4, A_5, A_6$, trong đó A_j là biến cố xuất hiện mặt j chấm $j = 1, 2, \ldots, 6$.*

Gọi A là biến cố xuất hiện mặt với số chấm chẵn thì A có 3 biến cố thuận lợi là A_2, A_4, A_6.

Ta có $A = A_2 + A_4 + A_6$
Gọi B là biến cố xuất hiện mặt với số chấm chia hết cho 3 thì B có 2 biến cố thuận lợi là A_3, A_6.

Ta có $B = A_3 + A_6$

viii) Biến cố tích

Biến cố C được gọi là tích của hai biến cố A và B, kí hiệu AB, nếu C xảy ra khi và chỉ khi cả A lẫn B cùng xảy ra.

• **Ví dụ 12** *Hai người cùng bắn vào một con thú.*

Gọi A là biến cố người thứ nhất bắn trượt, B là biến cố người thứ hai bắn trượt thì $C = AB$ là biến cố con thú không bị bắn trúng.

ix) Biến cố hiệu

Hiệu của biến cố A và biến cố B, kí hiệu $A \setminus B$ là biến cố xảy ra khi và chỉ khi A xảy ra nhưng B không xảy ra.

x) Biến cố xung khắc

Hai biến cố A và B được gọi là hai biến cố xung khắc nếu chúng không đồng thời xảy ra trong một phép thử.

• **Ví dụ 13** *Tung một đồng tiền.*

Gọi A là biến cố xuất hiện mặt xấp, B là biến cố xuất hiện mặt ngửa thì $AB = \emptyset$.

xi) Biến cố đối lập

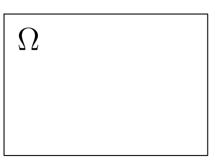
Biến cố *không xảy ra biến cố A* được gọi là biến cố đối lập với biến cố A. Kí hiệu $\overline{A}$. Ta có

$$A + \overline{A} = \Omega, \qquad A\overline{A} = \emptyset$$

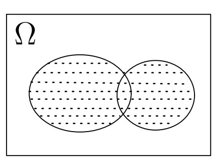
⊕ **Nhận xét**

Qua các khái niệm trên ta thấy các biến cố tổng, tích, hiệu, đối lập tương ứng với tập hợp, giao, hiệu, phần bù của lý thuyết tập hợp. Do đó ta có thể sử dụng các phép toán trên các tập hợp cho các phép toán trên các biến cố.

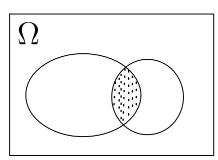
Ta có thể dùng biểu đồ Venn để miêu tả các biến cố.



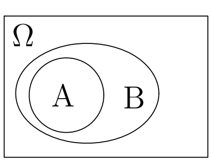
Bc chắc chắn



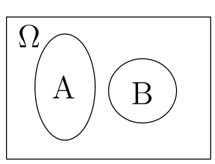
A+B



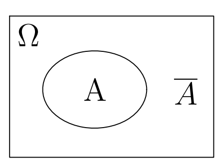
AB



A⟹B



A,B xung khắc



Đối lập $\overline{A}$

3. XÁC SUẤT

3.1 Định nghĩa xác suất theo lối cổ điển

□ **Định nghĩa 5** *Giả sử phép thử có n biến cố đồng khả năng có thể xảy ra, trong đó có m biến cố đồng khả năng thuận lợi cho biến cố A (A là tổng của m biến cố sơ cấp này). Khi đó xác suất của biến cố A, kí hiệu $P(A)$ được định nghĩa bằng công thức sau:*

$$\boxed{P(A) = \frac{m}{n} = \frac{\text{Số trường hợp thuận lợi cho } A}{\text{Số trường hợp có thể xảy ra}}}$$

• **Ví dụ 14** *Gieo một con xúc xắc cân đối, đồng chất. Tính xác suất xuất hiện mặt chẵn.*

Giải

Gọi A_i là biến cố xuất hiện mặt i chấm và A là biến cố xuất hiện mặt chẵn thì

$$A = A_2 + A_4 + A_6$$

Ta thấy phép thử có 6 biến cố sơ cấp đồng khả năng có thể xảy ra trong đó có 3 biến cố thuận lợi cho A.

$$P(A) = \frac{3}{6} = \frac{1}{2}$$

• **Ví dụ 15** *Một người gọi điện thoại nhưng lại quên 2 số cuối của số điện thoại cần gọi mà chỉ nhớ là 2 số đó khác nhau. Tìm xác suất để người đó quay ngẫu nhiên một lần trúng số cần gọi.*

Giải

Gọi A là biến cố người đó quay ngẫu nhiên một lần trúng số cần gọi.

Số biến cố sơ cấp đồng khả năng có thể xảy ra (số cách gọi 2 số cuối) là $n = A_{10}^2 = 90$.

Số biến cố thuận lợi cho A là $m = 1$.

Vậy $P(A) = \frac{1}{90}$.

• **Ví dụ 16** *Trong hộp có 6 bi trắng, 4 bi đen. Tìm xác suất để lấy từ hộp ra được*
i) 1 viên bi đen.
ii) 2 viên bi trắng.

Giải

Gọi A là biến cố lấy từ hộp ra được 1 viên bi đen và B là biến cố lấy từ hộp ra 2 viên bi trắng.

Ta có

i) $P(A) = \dfrac{C_4^1}{C_{10}^1} = \dfrac{2}{5}$

ii) $P(B) = \dfrac{C_6^2}{C_{10}^2} = \dfrac{1}{3}$

• **Ví dụ 17** *Rút ngẫu nhiên từ một cỗ bài tú lơ khơ 52 lá ra 5 lá. Tìm xác suất sao cho trong 5 lá rút ra có*

a) 3 lá đỏ và 2 lá đen.
b) 2 con cơ, 1 con rô, 2 con chuồn.

Giải

Gọi A là biến cố rút ra được 3 lá đỏ và 2 lá đen.
B là biến cố rút ra được 2 con cơ, 1 con rô, 2 con chuồn.

Số biến cố có thể xảy ra khi rút 5 lá bài là C_{52}^5.

a) Số biến cố thuận lợi cho A là $C_{26}^3.C_{26}^2$.

$$P(A) = \frac{C_{26}^3.C_{26}^2}{C_{52}^5} = \frac{845000}{2598960} = 0,3251$$

b) Số biến cố thuận lợi cho B là $C_{13}^2.C_{13}^1.C_{13}^2$

$$P(B) = \frac{C_{13}^2.C_{13}^1.C_{13}^2}{C_{52}^5} = \frac{79092}{2598960} = 0,30432$$

• **Ví dụ 18** (Bài toán ngày sinh) *Một nhóm gồn n người. Tìm xác suất để có ít nhất hai người có cùng ngày sinh (cùng ngày và cùng tháng).*

Giải

Gọi S là tập hợp các danh sách ngày sinh có thể của n người và E là biến cố có ít nhất hai người trong nhóm có cùng ngày sinh trong năm.

Ta có $\overline{E}$ là biến cố không có hai người bất kỳ trong nhóm có cùng ngày sinh.

Số các trường hợp của S là

$$n(S) = \underbrace{365.365\ldots365}_{n} = 365^n$$

Số trường hợp thuận lợi cho $\overline{E}$ là

$$\begin{aligned} n(\overline{E}) &= 365.364.363.\ldots.[365-(n-1)] \\ &= \frac{[365.364.363.\ldots.(366-n)](365-n)!}{(365-n)!} \\ &= \tfrac{365!}{(365-n)!} \end{aligned}$$

Vì các biên cố đồng khả năng nên

$$P(\overline{E}) = \frac{n(\overline{E})}{n(S)} = \frac{\frac{365!}{(365-n)!}}{365^n} = \frac{365!}{365^n.(365-n)!}$$

Do đó xác suất để ít nhất có hai người có cùng ngày sinh là

$$P(E) = 1 - P(\overline{E}) = 1 - \frac{\frac{365!}{(365-n)!}}{365^n} = \frac{365!}{365^n.(365-n)!}$$

Số người trong nhóm n	*Xác suất có ít nhất 2 người có cùng ngày sinh* $P(E)$
5	0,027
10	0,117
15	0,253
20	0,411
23	0,507
30	0,706
40	0,891
50	0,970
60	0,994
70	0,999

Bảng bài toán ngày sinh

⊙ **Chú ý** Định nghĩa xác suất theo lối cổ điển có một số hạn chế:

i) Nó chỉ xét cho hệ hữu hạn các biến cố sơ cấp.

ii) Không phải lúc nào việc "đồng khả năng" cũng xảy ra.

3.2 Định nghĩa xác suất theo lối thống kê

□ **Định nghĩa 6** *Thực hiện phép thử n lần. Giả sử biến cố A xuất hiện m lần. Khi đó m được gọi là tần số của biến cố A và tỷ số* $\frac{m}{n}$ *được gọi là tần suất xuất hiện biến cố A trong loạt phép thử.*

Cho số phép thử tăng lên vô hạn, tần suất xuất hiện biến cố A dần về một số xác định gọi là xác suất của biến cố A.

$$\boxed{P(A) = \lim_{n\to\infty} \frac{m}{n}}$$

• **Ví dụ 19** *Một xạ thủ bắn 1000 viên đạn vào bia. Có xấp xỉ 50 viên trúng bia. Khi đó xác suất để xạ thủ bắn trúng bia là* $\frac{50}{1000} = 5\%$.

• **Ví dụ 20** *Để nghiên cứu khả năng xuất hiện mặt sấp khi tung một đồng tiền, người ta tiến hành tung đồng tiền nhiều lần và thu được kết quả cho ở bảng dưới đây:*

Người làm thí nghiệm	Số lần tung	Số lần được mặt sấp	Tần suất $f(A)$
Buyffon	4040	2.048	0,5069
Pearson	12.000	6.019	0,5016
Pearson	24.000	12.012	0,5005

3.3 Định nghĩa xác suất theo quan điểm hình học

□ **Định nghĩa 7** *Xét một phép thử có không gian các biến cố sơ cấp Ω được biểu diễn bởi miền hình học Ω có độ đo (độ dài, diện tích, thể tích) hữu hạn khác 0, biến cố A được biểu diễn bởi miền hình học A. Khi đó xác suất của biến cố A được xác định bởi:*

$$\boxed{P(A) = \frac{\textit{Độ đo của miền } A}{\textit{Độ đo của miền } \Omega}}$$

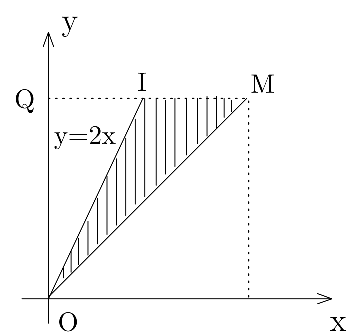
• **Ví dụ 21** *Trên đoạn thẳng OA ta gieo ngẫu nhiên hai điểm B và C có tọa độ tương ứng $OB = x$, $OC = y$ $(y \geq x)$. Tìm xác suất sao cho độ dài của đoạn BC bé hơn độ dài của đoạn OB.*

Giải

Giả sử $OA = l$. Các tọa độ x và y phải thỏa mãn các điều kiện:

$$0 \leq x \leq l, \quad 0 \leq y \leq l, \quad y \geq x \qquad (*)$$

Biểu diễn x và y lên hệ trục tọa độ vuông góc. Các điểm có tọa độ thỏa mãn (*) thuộc tam giác OMQ (có thể xem như biến cố chắc chắn).



Mặt khác, theo yêu cầu bài toán ta phải có $y - x < x$ hay $y < 2x$ (**). Những điểm có tọa độ thỏa mãn (*) và (**) thuộc miền có gạch. Miền thuận lợi cho biến cố cần tìm là tam giác OMI. Vậy xác suất cần tính

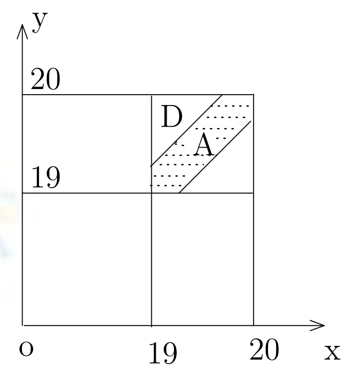
$$p = \frac{\text{diện tích } OMI}{\text{diện tích } OMQ} = \frac{1}{2}$$

• **Ví dụ 22** `(Bài toán hai người gặp nhau)`

Hai người hẹn gặp nhau ở một địa điểm xác định vào khoảng từ 19 giờ đến 20 giờ. Mỗi người đến (chắc chắn sẽ đến) điểm hẹn trong khoảng thời gian trên một cách độc lập với nhau, chờ trong 20 phút, nếu không thấy người kia đến sẽ bỏ đi. Tìm xác suất để hai người gặp nhau.

Giải

Gọi x, y là thời gian đến điểm hẹn của mỗi người và A là biến cố hai người gặp nhau. Rõ ràng x, y là một điểm ngẫu nhiên trong khoảng $[19, 20]$, ta có $19 \le x \le 20$;
$19 \le y \le 20$.



Để hai người gặp nhau thì

$|x - y| \le 20$ phút $= \frac{1}{3}$ giờ.

Do đó

$$\Omega = \{(x, y) :\ 19 \le x20,\ 19 \le y \le 20\}$$

$$A = \{(x, y) :\ |x - y| \le \frac{1}{3}\}$$

Diện tích của miền Ω bằng 1.

Diện tích của miền A bằng $1 - 2.\frac{1}{2}.\frac{2}{3}.\frac{2}{3} = \frac{5}{9}$

Vậy $P(A) = \dfrac{\text{diện tích A}}{\text{diện tích } \Omega} = \dfrac{5/9}{1} = 0,555.$

3.4 Định nghĩa xác suất theo tiên đề

Giả sử Ω là biến cố chắc chắn. Gọi $\mathcal{A}$ là họ các tập con của Ω thỏa các điều kiện sau:

i) $\mathcal{A}$ chứa Ω.

ii) Nếu $A, B \in \mathcal{A}$ thì $\overline{A}, A + B, AB$ thuộc $\mathcal{A}$.

Họ $\mathcal{A}$ thỏa các tiên đề i) và ii) thì $\mathcal{A}$ được gọi là đại số.

iii) Nếu $A_1, A_2, \ldots, A_n, \ldots$ là các phần tử của $\mathcal{A}$ thì tổng và tích vô hạn $A_1 + A_2 + \ldots + A_n$ và $A_1 A_2 \ldots A_n \ldots$ cũng thuộc $\mathcal{A}$.

Nếu $\mathcal{A}$ thỏa các điều kiện i), ii), iii) thì $\mathcal{A}$ được gọi là σ đại số.

□ **Định nghĩa 8** *Ta gọi xác suất trên $(\Omega, \mathcal{A})$ là một hàm P số xác định trên $\mathcal{A}$ có giá trị trong [0,1] và thỏa mãn 3 tiên đề sau:*

i) $P(\Omega) = 1$.

ii) $P(A + B) = P(A) + P(B)$ (với A, B xung khắc).

iii) Nếu dãy $\{A_n\}$ có tính chất $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ và $A_1 A_2 \ldots A_n \ldots = \emptyset$ thì $\lim\limits_{n\to\infty} P(A_n) = 0$.

3.5 Các tính chất của xác suất

i) $0 \leq P(A) \leq 1$ với mọi biến cố A
ii) $P(\Omega) = 1$
iii) $P(\emptyset) = 0$
iv) Nếu $A \subset B$ thì $P(A) \leq P(B)$.
v) $P(A) + P(\overline{A}) = 1$.
vi) $P(A) = P(AB) + P(A\overline{B})$.

4. MỘT SỐ CÔNG THỨC TÍNH XÁC SUẤT

4.1 Công thức cộng xác suất

⊙ Công thức 1

Giả sử A và B là hai biến cố xung khắc ($AB = \emptyset$). Ta có

$$\boxed{P(A + B) = P(A) + P(B)}$$

Chứng minh

Giả sử phép thử có n biến cố đồng khả năng có thể xảy ra, trong đó có m_A biến cố thuận lợi cho biến cố A và m_B biến cố thuận lợi cho biến cố B. Khi đó số biến cố thuận lợi cho biến cố $A + B$ là $m = m_A + m_B$.

Do đó
$$P(A + B) = \frac{m_A + m_B}{n} = \frac{m_A}{n} + \frac{m_B}{n} = P(A) + P(B)$$

□ Định nghĩa 9

i) Các biến cố $A_1, A_2, \ldots, A_n$ được gọi là nhóm các biến cố đầy đủ xung khắc từng đôi nếu chúng xung khắc từng đôi và tổng của chúng là biến cố chắc chắn. Ta có

$$A_1 + A_2 + \ldots + A_n = \Omega, \quad A_iA_j = \emptyset$$

ii) Hai biến cố A và B được gọi là hai biến cố độc lập nếu sự tồn tại hay không tồn tại của biến cố này không ảnh hưởng đến sự tồn tại hay không tồn tại của biến cố kia.

iii) Các biến cố $A_1, A_2, \ldots, A_n$ được gọi độc lập toàn phần nếu mỗi biến cố độc lập với tích của một tổ hợp bất kỳ trong các biến cố còn lại.

△ Hệ quả 1

i) Nếu $A_1, A_2, \ldots, A_n$ là biến cố xung khắc từng đôi thì

$$P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$$

ii) Nếu $A_1, A_2, \ldots, A_n$ là nhóm các biến cố đầy đủ xung khắc từng đôi thì

$$\sum_{i=1}^{n} P(A_i) = 1$$

iii) $P(A) = 1 - P(\overline{A})$.

⊙ Công thức 2

$$\boxed{P(A+B) = P(A) + P(B) - P(AB)}$$

Chứng minh

Giả sử phép thử có n biến cố đồng khả năng có thể xảy ra, trong đó có m_A biến cố thuận lợi cho biến cố A, m_B biến cố thuận lợi cho biến cố B và k biến cố thuận lợi cho biến cố AB. Khi đó số biến cố thuận lợi cho biến cố $A+B$ là $m_A + m_B - k$.

Do đó

$$P(A+B) = \frac{m_A + m_B - k}{n} = \frac{m_A}{n} + \frac{m_B}{n} - \frac{k}{n} = P(A) + P(B) - P(AB).$$

△ Hệ quả 2

i) $P(A_1 + A_2 + \ldots, +A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_iA_j) + \sum_{i<j<k} P(A_iA_jA_k) + \ldots + (-1)^{n-1} P(A_1A_2\ldots A_n)$.

ii) Nếu $A_1, A_2, \ldots, A_n$ là các biến cố độc lập toàn phần thì

$$P(A_1 + A_2 + \ldots + A_n) = 1 - P(\overline{A_1}).P(\overline{A_2})\ldots P(\overline{A_n}).$$

• **Ví dụ 23** *Một lô hàng gồm 10 sản phẩm, trong đó có 2 phế phẩm. Lấy ngẫu nhiên không hoàn lại từ lô hàng ra 6 sản phẩm. Tìm xác suất để có không quá 1 phế phẩm trong 6 sản phẩm được lấy ra.*

Giải

Gọi

A là biến cố không có phế phẩm trong 6 sản phẩm lấy ra.
B là biến cố có đúng 1 phế phẩm.
C là biến cố có không quá một phế phẩm

thì A và B là hai biến cố xung khắc và $C = A + B$.

Ta có

$$P(A) = \frac{C_8^6}{C_{10}^6} = \frac{28}{210} = \frac{2}{15}$$

$$P(B) = \frac{C_2^1.C_8^5}{C_{10}^6} = \frac{112}{210} = \frac{8}{15}$$

Do đó

$$P(C) = P(A) + P(B) = \frac{2}{15} + \frac{8}{15} = \frac{2}{3}$$

• **Ví dụ 24** *Một lớp có 100 sinh viên, trong đó có 40 sinh viên giỏi ngoại ngữ, 30 sinh viên giỏi tin học, 20 sinh viên giỏi cả ngoại ngữ lẫn tin học. Sinh viên nào giỏi ít nhất một trong hai môn sẽ được thêm điểm trong kết quả học tập của học kỳ. Chọn ngẫu nhiên một sinh viên trong lớp. Tìm xác suất để sinh viên đó được tăng điểm.*

Giải

Gọi
A là biến cố gọi được sinh viên được tăng điểm.
N là biến cố gọi được sinh viên giỏi ngoại ngữ.
T là biến cố gọi được sinh viên giỏi tin học

thì $A = T + N$.

Ta có

$$P(A) = P(T) + P(N) - P(TN) = \frac{30}{100} + \frac{40}{100} - \frac{20}{100} = \frac{50}{100} = 0,5$$

4.2 Xác suất có điều kiện và công thức nhân xác suất

a) Xác suất có điều kiện

□ **Định nghĩa 10** *Xác suất của biến cố A với điều kiện biến cố B xảy ra được gọi là xác có điều kiện của biến cố A. Kí hiệu $P(A/B)$.*

• **Ví dụ 25** *Trong hộp có 5 viên bi trắng, 3 viên bi đen. Lấy lần lượt ra 2 viên bi (không hoàn lại). Tìm xác suất để lần thứ hai lấy được viên bi trắng biết lần thứ nhất đã lấy được viên bi trắng.*

Giải

Gọi A là biến cố lần thứ hai lấy được viên bi trắng
B là biến cố lần thứ nhất lấy được viên bi trắng.

Ta tìm $P(A/B)$.

Ta thấy lần thứ nhất lấy được viên bi trắng (B đã xảy ra) nên trong hợp còn 7 viên bi trong đ ó có 4 viên bi trắng. Do đó

$$P(A/B) = \frac{C_4^1}{C_7^1} = \frac{4}{7}$$

⊙ Công thức

$$\boxed{P(A/B) = \frac{P(AB)}{P(B)}}$$

Chứng minh

Giả sử phép thử có n biến cố đồng khả năng có thể xảy ra trong đó có m_A biến có thuận lợi cho biến cố A, m_B biến cố thuận lợi cho biến cố B và k biến cố thuận lợi cho biến cố AB.

Theo định nghĩa xác suất theo lối cổ điển ta có

$$P(AB) = \frac{k}{n}, \quad P(B) = \frac{m_B}{n}$$

Ta tìm $P(A/B)$. Vì biến cố B đã xảy ra nên biến cố đồng khả năng của A là m_B, biến cố thuận lợi cho A là k. Do đó

$$P(A/B) = \frac{k}{m_B} = \frac{\frac{k}{n}}{\frac{m_B}{n}} = \frac{P(AB)}{P(B)}.$$

• **Ví dụ 26** *Một bộ bài có 52 lá. Rút ngẫu nhiên 1 lá bài. Tìm xác suất để rút được con "át" biết rằng lá bài rút ra là lá bài màu đen.*

Giải

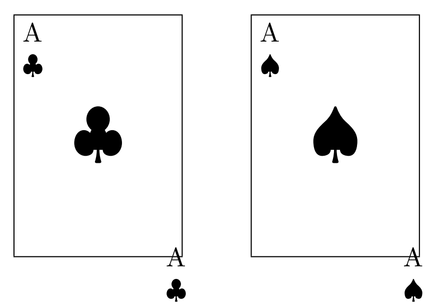
Gọi A là biến cố rút được con "át"
B là biến cố rút được lá bài màu đen.

Ta thấy trong bộ bài có

26 lá bài đen nên $P(B) = \frac{26}{52}$

2 con "át" đen nên $P(AB) = \frac{2}{52}$.

Do đó $P(A/B) = \dfrac{P(AB)}{P(B)} = \dfrac{2/52}{26/52} = \dfrac{1}{13}$



b) Công thức nhân xác suất

Từ công thức xác suất có điều kiện ta có

i) $P(AB) = P(A).P(B/A) = P(B).P(A/B)$.

ii) Nếu A, B là hai biến cố độc lập thì $P(AB) = P(A).P(B)$.

iii) $P(ABC) = P(A).P(B/A).P(C/AB)$
$P(A_1A_2\ldots A_n) = P(A_1)P(A_2/A_1)\ldots P(A_n/A_1A_2\ldots A_{n-1})$.

• **Ví dụ 27** *Hộp thứ nhất có 2 bi trắng và 10 bi đen. Hộp thứ hai có 8 bi trắng và 4 bi đen. Từ mỗi hộp lấy ra 1 viên bi. Tìm xác suất để*

a) Cả 2 viên bi đều trắng,
b) 1 bi trắng, 1 bi đen.

Giải

Gọi T là biến cố lấy ra được cả 2 bi trắng
T_1 là biến cố lấy được bi trắng từ hộp thứ nhất
T_2 là biến cố lấy được bi trắng từ hộp thứ hai

thì T_1, T_2 là 2 biến cố độc lập và $T = T_1T_2$. Ta có

$$P(T_1) = \frac{1}{6}, \quad P(T_2) = \frac{2}{3}$$

Do đó $P(T) = P(T_1T_2) = P(T_1).P(T_2) = \frac{1}{6}.\frac{2}{3} = \frac{1}{9}$.

b) Gọi T_1, T_2 là biến cố lấy được bi trắng ở hộp thứ nhất, thứ hai
D_1, D_2 là biến cố lấy được bi đen ở hộp thứ nhất, thứ hai
T_1D_2 là biến cố lấy được bi trắng ở hộp thứ nhất và bi đen ở hộp thứ hai
T_2D_1 là biến cố lấy được bi trắng ở hộp thứ hai và bi đe n ở hộp thứ nhất

thì $A = T_1D_2 + T_2D_1$.

Ta có

$$P(T_1) = \frac{1}{6}, \quad P(T_2) = \frac{2}{3}$$

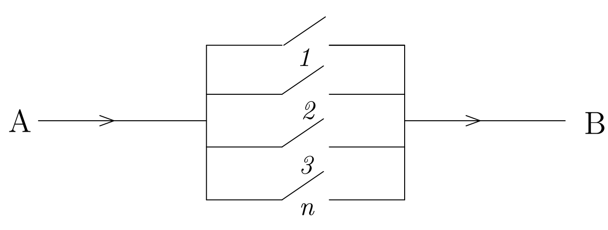
$$P(D_1) = 1 - P(T_1) = \frac{5}{6} \quad P(D_2) = 1 - P(T_2) = \frac{1}{3}$$

Suy ra

$$P(A) = P(T_1D_2) + P(T_2D_1) = P(T_1).P(D_2) + P(T_2).P(T_1)$$

$$= \frac{1}{6}.\frac{1}{3} + \frac{2}{3}.\frac{5}{6} = \frac{11}{8}$$

• **Ví dụ 28** *Một hệ thống được cấu thành bởi n thành phần riêng lẻ được gọi là một hệ thống song song nếu nó hoạt động khi ít nhất một thành phần hoạt động. Thành phần thứ i (độc lập với các thành phần khác) hoạt động với xác suất p_i. Tìm xác suất để hệ thống song song hoạt động.*



Giải

Gọi

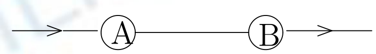
A là biến cố hệ thống hoạt động.

A_i là biến cố thành phần thứ i hoạt động.

Ta có

$$\begin{aligned}
\text{P(A)} &= 1 - P(\overline{A}) \\
&= 1 - P(\overline{A_1}.\overline{A_2}\ldots\overline{A_n}) \\
&= 1 - \prod_{i=1}^{n}\overline{P(A_i)} \\
&= 1 - \prod_{i=1}^{n}(1 - p_i)
\end{aligned}$$

• **Ví dụ 29** (`Hệ xích`) *Xét một hệ thống gồm hai thành phần. Hệ thống hoạt động khi và chỉ khi cả hai thành phần hoạt động (các thành phần được nối theo xích).*



Độ tin cậy R(t) của một thành phần của hệ thống là xác suất mà thành phần có thể hoạt động ít nhất khoảng thời gian t.

Nếu kí hiệu biến cố "thành phần hoạt động ít nhất t đơn vị thời gian" bởi $T > t$ thì

$$R(t) = P(T > t)$$

Gọi P_A và P_B là độ tin cậy của thành phần A và B, nghĩa là

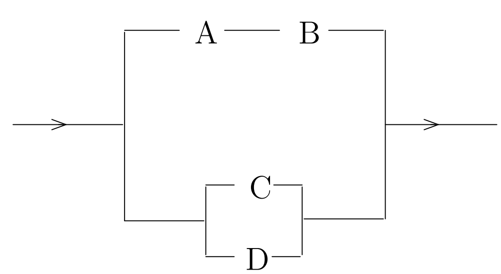
$P_A = P(A$ hoạt động ít nhất t đơn vị thời gian),

$P_B = P(B$ hoạt động ít nhất t đơn vị thời gian).

Nếu các thành phần hoạt động độc lập thì độ tin cậy của hệ thống là $R = p_A.p_B$.

• **Ví dụ 30**

Xét độ tin cậy của hệ thống cho bởi hình bên. Thành phần nối A và B trên đỉnh có thể thay bởi thành phần đơn với độ tin cậy $p_A.p_B$. Thành phần song song của ngắt C và D có thể thay bởi ngắt đơn với độ tin cậy $1-(1-p_C).(1-p_D)$.



Độ tin cậy của hệ thống song song này là

$$1 - (1 - p_A.p_B)[1 - (1 - (1 - p_C).(1 - p_D))]$$

4.3 Công thức xác suất đầy đủ và công thức Bayes

a) Công thức xác suất đầy đủ

⊙ Công thức

Giả sử $A_1, A_2, \ldots, A_n$ là nhóm các biến cố đầy đủ xung khắc từng đôi và B là biến cố bất kỳ có thể xảy ra trong phép thử. Khi đó ta có

$$\boxed{P(B) = \sum_{i=1}^{n} P(A_i).P(B/A_i)}$$

Chứng minh

Vì $A_1 + A_2 + \ldots + A_n = \Omega$ nên

$$B = B(A_1 + A_2 + \ldots + A_n) = BA_1 + B_2 + \ldots + BA_n$$

Do các biến cố $A_1, A_2, \ldots, A_n$ xung khắc từng đôi nên các biến cố tích $BA_1, BA_2, \ldots,$ BA_n cũng xung khắc từng đôi.

Theo định lý cộng xác suất ta có $P(B) = \sum_{i=1}^{n} P(BA_i)$.

Mặt khác theo công thức nhân xác suất thì $P(BA_i) = P(A_i).P(B/A_i)$.

Do đó $P(B) = \sum_{i=1}^{n} P(A_i).P(B/A_i)$.

⊙ Chú ý Công thức trên còn đúng nếu ta thay điều kiện $A_1 + A_2 + \ldots + A_n = \Omega$ bởi $B \subset A_1 + A_2 + \ldots + A_n$.

• **Ví dụ 31** *Xét một lô sản phẩm trong đó số sản phẩm do nhà máy I sản xuất chiếm 20%, nhà máy II sản xuất chiếm 30%, nhà máy III sản xuất chiếm 50%. Xác suất phế phẩm của nhà máy I là 0,001; nhà máy II là 0,005; nhà máy III là 0,006. Tìm xác suất để lấy ngẫu nhiên được đúng 1 phế phẩm.*

Giải

Gọi B là biến cố sản phẩm lấy ra là phế phẩm
A_1, A_2, A_3 là biến cố lấy được sản phẩm của nhà máy I, II, III
thì A_1, A_2, A_3 là nhóm các biến cố xung khắc từng đôi. Ta có

$$P(A_1) = 0,2; \quad P(A_2) = 0,3; \quad P(A_3) = 0,5$$

$$P(B/A_1) = 0,001; \quad P(B/A_2) = 0,005; \quad P(B/A_3) = 0,006$$

Do đó

$$\begin{aligned} P(B) &= P(A_1).P(B/A_1) + P(A_2).P(B/A_2) + P(A_3).P(B/A_3) \\ &= 0,2.0,001 + 0,3.0,005 + 0,5.0,006 \\ &= 0,0065 \end{aligned}$$

• **Ví dụ 32** *Một hộp chứa 4 bi trắng, 3 bi vàng và 1 bi xanh. Lấy lần lượt (không hoàn lại) từ hộp ra 2 bi. Tìm xác suất để lấy được 1 bi trắng và 1 bi vàng.*

Giải

Gọi T là biến cố lấy được bi trắng, V là biến cố lấy được bi vàng.

Ta có

$$P(T) = \frac{4}{8} = \frac{1}{2}; \qquad P(V) = \frac{3}{8};$$
$$P(V/T) = \frac{3}{7}; \qquad P(T/V) = \frac{4}{7}$$

Xác xuất để lấy được 1 bi trắng và 1 bi vàng là

$$P(TV) = P(T).P(V/T) + P(V).P(T/V) = \frac{1}{2}.\frac{3}{7} + \frac{3}{8}.\frac{4}{7} = \frac{3}{7}.$$

□ Cây xác suất

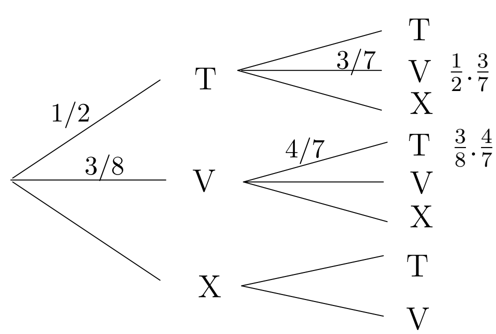
Trong thực tế có nhiều phép thử chứa một dãy nhiều biến cố. *Cây xác suất* cung cấp cho ta một công cụ thuận lợi cho việc xác định cấu trúc các quan hệ bên trong các phép thử khi tính xác suất.

Cấu trúc của cây xác suất được xác định như sau:

i) Vẽ biểu đồ cây xác suất tương ứng với các kết quả của dãy phép thử.

ii) Gán mỗi xác suất với mỗi nhánh.

Cây xác suất sau minh họa cho ví dụ 32.



b) Công thức Bayes

⊙ Công thức

Giả sử $A_1, A_2, \ldots, A_n$ là nhóm các biến cố đầy đủ xung khắc từng đôi và B là biến cố bất kỳ có thể xảy ra trong phép thử. Khi đó ta có

$$\boxed{P(A_i/B) = \frac{P(A_i).P(B/A_i)}{\sum_{i=1}^{n} P(A_i).P(B/A_i)}} \quad i = 1, 2, \ldots, n$$