TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN TOÁN ỨNG DỤNG VÀ TIN HỌC

BÀI GIẢNG

XÁC SUẤT THỐNG KÊ

NGUYỄN THỊ THU THỦY
BỘ MÔN TOÁN ỨNG DỤNG

HÀ NỘI - 01/2020

MỤC LỤC

Lời nói đầu

Lý thuyết xác suất và thống kê toán học là một ngành khoa học đang giữ vị trí quan trọng trong các lĩnh vực ứng dụng rộng rãi và phong phú của đời sống con người. Cùng với sự phát triển mạnh mẽ của khoa học và công nghệ, nhu cầu hiểu biết và sử dụng các công cụ ngẫu nhiên trong phân tích và xử lý thông tin ngày càng trở nên đặc biệt cần thiết. Các kiến thức và phương pháp của xác suất và thống kê đã hỗ trợ hữu hiệu các nhà nghiên cứu trong nhiều lĩnh vực khoa học khác nhau như vật lý, hóa học, sinh học, nông học, kinh tế học, xã hội học, ngôn ngữ học... Do đó "Xác suất thống kê" là học phần rất cần thiết cho sinh viên bậc đại học.

Bài giảng học phần "Xác suất thống kê", mã học phần MI2020 được biên soạn theo Đề cương chi tiết với khối lượng 30 tiết lý thuyết, 30 tiết bài tập dành cho sinh viên hệ đại học chính quy (không phải chuyên ngành Toán Tin) của Trường Đại học Bách khoa Hà Nội.

Mục tiêu học phần: Cung cấp cho sinh viên những kiến thức cơ bản về xác suất là các khái niệm và quy tắc suy diễn xác suất cũng như về biến ngẫu nhiên và các phân phối xác suất thông dụng (một và hai chiều); các khái niệm cơ bản của thống kê toán học nhằm giúp sinh viên biết cách xử lý các bài toán thống kê về ước lượng, kiểm định giả thuyết.... Trên cơ sở đó sinh viên có được một phương pháp tiếp cận với mô hình thực tế và có kiến thức cần thiết để đưa ra lời giải đúng cho các bài toán đó.

Nội dung vắn tắt học phần: Sự kiện ngẫu nhiên và phép tính xác suất, đại lượng ngẫu nhiên, phân phối xác suất, véc tơ ngẫu nhiên, lý thuyết ước lượng thống kê, lý thuyết quyết định thống kê.

Chương 1

Sự kiện ngẫu nhiên và phép tính xác suất

Các hiện tượng trong tự nhiên hay xã hội xảy ra một cách ngẫu nhiên (không biết trước kết quả) hoặc tất định (biết trước kết quả sẽ xảy ra). Chẳng hạn một vật nặng được thả từ trên cao chắc chắn sẽ rơi xuống đất, trong điều kiện bình thường nước sôi ở 100° C... Đó là những hiện tượng diễn ra có tính quy luật, tất nhiên. Trái lại, khi tung đồng xu ta không biết sẽ xuất hiện mặt sấp hay mặt ngửa; ta không thể biết trước có bao nhiêu cuộc gọi đến tổng đài; có bao nhiêu khách hàng đến điểm phục vụ trong khoảng thời gian nào đó; ta không thể xác định trước chỉ số chứng khoán trên thị trường chứng khoán... Đó là những hiện tượng ngẫu nhiên. Tuy nhiên, nếu tiến hành quan sát nhiều lần một hiện tượng ngẫu nhiên trong những hoàn cảnh như nhau, thì trong nhiều trường hợp ta có thể rút ra những kết luận có tính quy luật về những hiện tượng này. Lý thuyết xác suất nghiên cứu các quy luật của các hiện tượng ngẫu nhiên. Việc nắm bắt các quy luật này sẽ cho phép dự báo các hiện tượng ngẫu nhiên đó sẽ xảy ra như thế nào. Chính vì vậy các phương pháp của lý thuyết xác suất được ứng dụng rộng rãi trong việc giải quyết các bài toán thuộc nhiều lĩnh vực khác nhau của khoa học tự nhiên, kỹ thuật và kinh tế–xã hội.

1.1 Sự kiện. Quan hệ giữa các sự kiện

1.1.1 Phép thử. Sự kiện

Định nghĩa 1.1 (Phép thử. Sự kiện)**.** (a) Việc thực hiện một nhóm các điều kiện cơ bản để quan sát một hiện tượng nào đó được gọi là một phép thử (experiment).

(b) Hiện tượng, kết quả xét trong phép thử gọi là sự kiện hay biến cố (event).

(c) Sự kiện sơ cấp hay kết cục của phép thử là một kết quả mà ta không chia nhỏ hơn được, ký hiệu là ω.

(d) Sự kiện phức hợp là sự kiện có thể phân tích thành các sự kiện nhỏ hơn.

(e) Tập hợp tất cả các kết cục của một phép thử tạo thành không gian các sự kiện sơ cấp, ký hiệu là

$$\Omega = \{\omega_i,\ i \in I\}, \quad I \text{ là tập chỉ số.}$$

Ví dụ 1.1. (a) Gieo một con xúc xắc (cân đối, đồng chất, trên mặt phẳng cứng) là một phép thử. Xúc xắc xuất hiện mặt 1, 2, 3, 4, 5, 6 chấm là các sự kiện.

(b) Gieo một đồng xu (cân đối, đồng chất, trên mặt phẳng cứng) là một phép thử. Đồng xu xuất hiện mặt sấp, mặt ngửa là các sự kiện.

Ví dụ 1.2. Gieo một con xúc xắc, khi đó

(a) Sự kiện A_i "xuất hiện mặt i chấm", $i = 1, \ldots, 6$ là sự kiện sơ cấp.

(b) Sự kiện A "xuất hiện mặt chấm chẵn" là sự kiện phức hợp vì có thể phân tích nó thành các sự kiện "xuất hiện mặt 2, 4, 6 chấm".

Ví dụ 1.3. (a) Phép thử gieo một đồng xu (cân đối, đồng chất, trên mặt phẳng cứng) có không gian các sự kiện sơ cấp là $\Omega = \{S, N\}$.

(b) Phép thử gieo đồng thời hai đồng xu (cân đối, đồng chất, trên mặt phẳng cứng) có không gian các sự kiện sơ cấp là $\Omega = \{SS, SN, NS, NN\}$.

Chú ý 1.1. (a) Chú ý rằng bản chất của các sự kiện sơ cấp không có vai trò đặc biệt gì trong lý thuyết xác suất. Chẳng hạn có thể mã hóa các kết quả và xem không gian các sự kiện sơ cấp của phép thử tung đồng xu là $\Omega = \{0, 1\}$, trong đó 0 là sự kiện sơ cấp chỉ mặt sấp xuất hiện và 1 để chỉ mặt ngửa xuất hiện.

(b) Mỗi kết cục ω của phép thử $\mathcal{C}$ được gọi là kết cục thuận lợi cho sự kiện A nếu A xảy ra khi kết cục của phép thử $\mathcal{C}$ là ω.

Ví dụ 1.4. Nếu gọi sự kiện A "xuất hiện mặt chấm chẵn" trong phép thử gieo con xúc xắc thì A có các kết cục thuận lợi là 2, 4, 6.

1.1.2 Phân loại sự kiện

Có 3 loại sự kiện.

(a) **Sự kiện chắc chắn** là sự kiện nhất định sẽ xảy ra khi thực hiện một phép thử. Ký hiệu là U hoặc Ω hoặc S.

(b) **Sự kiện không thể có** là sự kiện nhất định không xảy ra khi thực hiện một phép thử. Ký hiệu là V hoặc $\emptyset$.

(c) **Sự kiện ngẫu nhiên** là sự kiện có thể xảy ra, cũng có thể không xảy ra khi thực hiện một phép thử. Ký hiệu là A, B, C, A_1, A_2 ...

Ví dụ 1.5. Gieo một con xúc xắc, khi đó

(a) Sự kiện S "xuất hiện mặt có số chấm ≤ 6 và ≥ 1" là sự kiện chắc chắn.

(b) Sự kiện $\emptyset$ "xuất hiện mặt 7 chấm" là sự kiện không thể.

(c) Sự kiện A "xuất hiện mặt chấm chẵn" là sự kiện ngẫu nhiên.

1.1.3 Quan hệ giữa các sự kiện

Một cách tương ứng với các phép toán của tập hợp, trong lý thuyết xác suất người ta xét các quan hệ sau đây cho các sự kiện trong cùng một phép thử.

(a) **Quan hệ kéo theo:** Sự kiện A kéo theo sự kiện B, ký hiệu $A \subset B$, nếu khi A xảy ra thì B xảy ra.

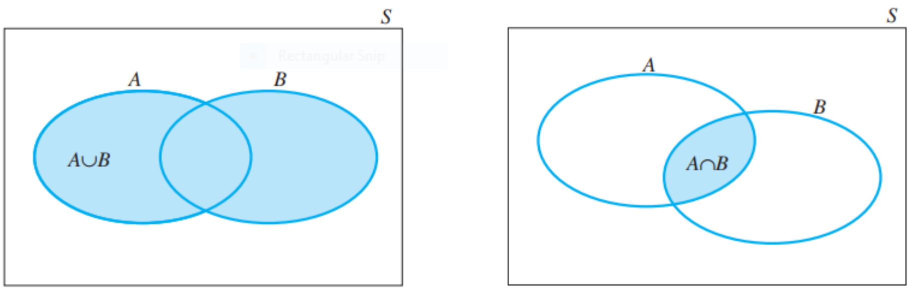
Nếu $A \subset B$ và $B \subset A$ thì ta nói hai sự kiện A và B trùng nhau, viết là $A = B$.

(b) **Tổng các sự kiện:** Sự kiện A được gọi là tổng của các sự kiện A_1, A_2,..., A_n nếu A xảy ra khi và chỉ khi ít nhất một trong các sự kiện A_i xảy ra, $i = 1, 2, \ldots, n$. Viết là:

$$\boxed{A = A_1 + A_2 + \cdots + A_n}$$

hoặc

$$\boxed{A = A_1 \cup A_2 \cup \cdots \cup A_n}$$



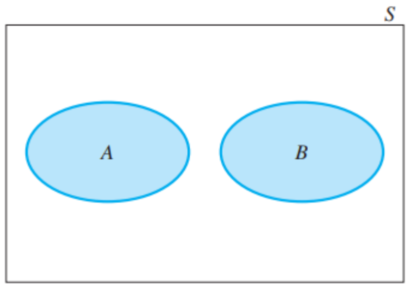
Hình 1.1: Sơ đồ Venn của $A \cup B$ và $A \cap B$

(c) **Tích các sự kiện:** Sự kiện B được gọi là tích của các sự kiện A_1, A_2,..., A_n nếu B xảy ra khi và chỉ khi tất cả các sự kiện A_i xảy ra, $i = 1, 2, \ldots, n$. Viết là:

$$\boxed{B = A_1 A_2 \ldots A_n}$$

hoặc

$$\boxed{B = A_1 \cap A_2 \cap \cdots \cap A_n}$$



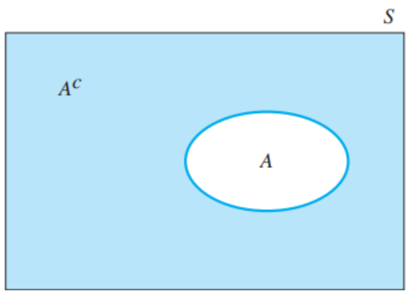
Hình 1.2: Hai sự kiện xung khắc

(d) **Sự kiện xung khắc:** Hai sự kiện A và B được gọi xung khắc với nhau nếu chúng không đồng thời xảy ra trong cùng một phép thử.

Như vậy, nếu A và B xung khắc thì $A \cap B = \emptyset$.

(e) **Sự kiện đối lập:** Sự kiện không xảy ra sự kiện A được gọi là sự kiện đối lập của A, ký hiệu là $\overline{A}$ hoặc A^c.

Như vậy A và $\overline{A}$ thỏa mãn tính chất: $A \cup \overline{A} = S$ và $A \cap \overline{A} = \emptyset$.



Hình 1.3: Sự kiện đối lập

(f) **Hiệu hai sự kiện:** Hiệu của 2 sự kiện A và B, ký hiệu là $A - B$, là sự kiện xảy ra khi và chỉ khi A xảy ra nhưng B không xảy ra.

Trường hợp hay sử dụng sự kiện hiệu: $\overline{A} = S - A$, $A = S - \overline{A}$.

Trường hợp tổng quát, ta biến đổi thành sự kiện tích như sau: $A - B = A \cap \overline{B}$.

(g) **Hệ (nhóm) đầy đủ các sự kiện:** Hệ (nhóm) n sự kiện A_1, A_2,..., A_n được gọi là hệ (nhóm) đầy đủ các sự kiện nếu nhất định phải xảy ra một và chỉ một trong các sự kiện ấy sau phép thử. Như vậy hệ $\{A_1, A_2, \ldots, A_n\}$ là hệ đầy đủ nếu

$$\begin{cases} A_i \cap A_j = \emptyset, & i \neq j, \\ A_1 \cup A_2 \cup \cdots \cup A_n = S. \end{cases}$$

Nhận xét 1.1. Các sự kiện trong cùng một phép thử với phép toán tổng, tích và lấy sự kiện đối tạo thành đại số Boole, do đó các phép toán này có các tính chất như các phép toán hợp, giao, lấy phần bù đối với các tập con của không gian các sự kiện sơ cấp. Chẳng hạn

1. $A \cap \emptyset = \emptyset$.
2. $A \cup \emptyset = A$.
3. $A \cap \overline{A} = \emptyset$.
4. $A \cup \overline{A} = S$.
5. $\overline{S} = \emptyset$.
6. $\overline{\emptyset} = S$.
7. $\overline{(\overline{A})} = A$.
8. $\overline{(A \cap B)} = \overline{A} \cup \overline{B}$.
9. $\overline{(A \cup B)} = \overline{A} \cap \overline{B}$.
10. $\overline{A \cup B} = \overline{A} \cap \overline{B}$; $\overline{A \cap B} = \overline{A} \cup \overline{B}$.
11. $A = A \cap (B \cup \overline{B}) = (A \cap B) \cup (A \cap \overline{B})$.

Chú ý 1.2. (a) Mọi sự kiện ngẫu nhiên đều có thể biểu diễn dưới dạng tổng của một số sự kiện sơ cấp nào đó.

Sự kiện chắc chắn S là tổng của mọi sự kiện sơ cấp có thể. Do đó S còn được gọi là không gian các sự kiện sơ cấp Ω.

(b) Đối với một sự kiện A thì ta có hệ đầy đủ $\left\{A, \overline{A}\right\}$.

Đối với hai sự kiện A và B, một hệ đầy đủ là $\left\{A \cap B, A \cap \overline{B}, \overline{A} \cap B, \overline{A} \cap \overline{B}\right\}$.

Tính chất 1.1. (a) $A \cup B = B \cup A$, $A \cap B = B \cap A$ (giao hoán).

(b) $A \cup B \cup C = (A \cup B) \cup C = A \cup (B \cup C)$, $A \cap B \cap C = (A \cap B) \cap C = A \cap (B \cap C)$ (kết hợp).

(c) $A \cap (B \cup C) = A \cap B \cup A \cap C$ (phân phối của phép cộng và phép nhân).

Đặc biệt $A + A = A$; $AA = A$; $A + S = S$; $AS = A$; $A + \emptyset = A$; $A\emptyset = \emptyset$.

Ví dụ 1.6. (a) Một mạng điện gồm ba bóng đèn mắc nối tiếp. Gọi A_i là sự kiện "bóng đèn thứ i bị cháy", $i = 1, 2, 3$. Gọi A là sự kiện "mạng mất điện". Ta thấy rằng mạng bị mất điện khi ít nhất một trong ba bóng bị cháy. Vậy $A = A_1 + A_2 + A_3$.

(b) Một mạng điện gồm ba bóng đèn mắc song song. Gọi B_i là sự kiện "bóng đèn thứ i bị cháy", $i = 1, 2, 3$. Gọi B là sự kiện "mạng mất điện". Ta thấy rằng mạng bị mất điện khi cả ba bóng bị cháy. Vậy $B = B_1B_2B_3$.

(c) Một nhà máy có ba phân xưởng sản xuất ra cùng một loại sản phẩm. Giả sử rằng mỗi sản phẩm của nhà máy chỉ do một trong ba phân xưởng này sản xuất. Chọn ngẫu nhiên một sản phẩm, gọi C_i là sự kiện "sản phẩm được chọn do phân xưởng thứ i sản xuất", $i = 1, 2, 3$. Khi đó hệ ba sự kiện $\{C_1, C_2, C_3\}$ là hệ đầy đủ.

Ví dụ 1.7. Ba xạ thủ A, B, C mỗi người bắn một viên đạn vào mục tiêu. Gọi A_1, A_2, A_3 lần lượt là sự kiện "A, B, C bắn trúng mục tiêu".

(a) Hãy mô tả các sự kiện $A_1A_2A_3$, $\overline{A}_1\overline{A}_2\overline{A}_3$, $A_1 + A_2 + A_3$.

(b) Biểu diễn các sự kiện sau theo A_1, A_2, A_3:

A: Có ít nhất hai xạ thủ bắn trúng;

B: Có nhiều nhất một xạ thủ bắn trúng;

C: Chỉ có xạ thủ A bắn trúng;

D: Chỉ có một xạ thủ bắn trúng.

(c) Các sự kiện A_1, A_2, A_3 có xung khắc không?

Lời giải:

(a) $A_1A_2A_3$: "cả ba xạ thủ đều bắn trúng";

$\overline{A}_1\overline{A}_2\overline{A}_3$: "cả ba xạ thủ đều bắn trượt";

$A_1 + A_2 + A_3$: "có ít nhất một xạ thủ bắn trúng".

(b) $A = A_1A_2 + A_1A_3 + A_2A_3$;

$B = \overline{A}_1\overline{A}_2 + \overline{A}_1\overline{A}_3 + \overline{A}_2\overline{A}_3$;

$C = A_1\overline{A}_2\overline{A}_3$;

$D = A_1\overline{A}_2\overline{A}_3 + \overline{A}_1A_2\overline{A}_3 + \overline{A}_1\overline{A}_2A_3$.

(c) Các sự kiện A_1, A_2, A_3 không xung khắc vì có thể cả ba xạ thủ đều bắn trúng mục tiêu.

1.2 Giải tích kết hợp

1.2.1 Quy tắc cộng. Quy tắc nhân

1.2.1a Quy tắc cộng

Định nghĩa 1.2 (Quy tắc cộng)**.** Nếu một công việc được chia ra thành k trường hợp để thực hiện, trường hợp một có n_1 cách thực hiện xong công việc, trường hợp hai có n_2 cách thực hiện xong công việc,..., trường hợp k có n_k cách thực hiện xong công việc và không có một cách thực hiện nào ở trường hợp này lại trùng với một cách thực hiện ở trường hợp khác. Khi đó ta có $n = n_1 + n_2 + \cdots + n_k$ cách thực hiện công việc.

1.2.1b Quy tắc nhân

Định nghĩa 1.3 (Quy 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 $n = n_1n_2 \ldots n_k$ cách thực hiện công việc.

Ví dụ 1.8. Giả sử để đi từ A đến C có thể đi qua B, trong đó có 2 đường khác nhau đi trực tiếp từ A đến C, có 3 đường khác nhau để đi từ A đến B và có 2 đường khác nhau để đi từ B đến C. Hỏi có bao nhiêu cách đi từ A đến C?

Lời giải: Đi từ A đến C có 2 lựa chọn: Đi trực tiếp từ A đến C: có $n_1 = 2$ cách; đi gián tiếp từ A đến C thông qua B: có $n_2 = 3 \times 2 = 6$ (cách).

Tổng số cách đi từ A đến C là $n = n_1 + n_2 = 2 + 6 = 8$ (cách).

1.2.2 Chỉnh hợp

Định nghĩa 1.4 (Chỉnh hợp)**.** Chỉnh hợp chập k của n phần tử là một nhóm có thứ tự gồm k phần tử khác nhau lấy từ n phần tử đã cho ($k \le n$).

Ký hiệu và công thức tính:

$$\boxed{A_n^k = \frac{n!}{(n-k)!} = n(n-1)\dots(n-k+1)} \tag{1.1}$$

Ví dụ 1.9. Từ 5 chữ số $1, 2, 3, 4, 5$ lập được bao nhiêu số có 3 chữ số khác nhau?

Lời giải: Số các số được lập là $A_5^3 = 5 \times 4 \times 3 = 60$ (số).

1.2.3 Chỉnh hợp lặp

Định nghĩa 1.5 (Chỉnh hợp lặp)**.** Chỉnh hợp lặp chập k của n phần tử là một nhóm có thứ tự gồm k phần tử có thể giống nhau lấy từ n phần tử đã cho.

Ký hiệu và công thức tính:

$$\boxed{\overline{A}_n^k = n^k} \tag{1.2}$$

Ví dụ 1.10. Từ 5 chữ số $1, 2, 3, 4, 5$ lập được bao nhiêu số có 3 chữ số?

Lời giải: Chọn 3 chữ số từ 5 chữ số có thứ tự và có thể lặp lại. Số các số được lập là $\overline{A}_5^3 = 5^3 = 125$ (số).

1.2.4 Hoán vị

Định nghĩa 1.6 (Hoán vị)**.** Hoán vị của n phần tử là một nhóm có thứ tự gồm đủ mặt n phần tử đã cho. Nói cách khác, hoán vị là một chỉnh hợp chập n của n phần tử.

Ký hiệu và công thức tính:

$$\boxed{P_n = A_n^n = n!} \tag{1.3}$$

Ví dụ 1.11. Có 6 người khách cần xếp vào 6 ghế trên một bàn tròn 6 chỗ.

(a) Nếu có quan tâm đến khung cảnh xung quanh thì có bao nhiêu cách sắp xếp?

(b) Nếu chỉ quan tâm đến người ngồi xung quanh là ai thì sẽ có bao nhiêu cách?

Lời giải: (a) $P_6 = 6! = 720$ (cách); (b) $P_5 = 5! = 120$ (cách).

1.2.5 Tổ hợp

Định nghĩa 1.7 (Tổ hợp)**.** Tổ hợp chập k của n phần tử là một nhóm không phân biệt thứ tự gồm k phần tử khác nhau lấy từ n phần tử đã cho ($k \le n$).

Ký hiệu và công thức tính:

$$\boxed{C_n^k = \frac{n!}{k!(n-k)!}} \tag{1.4}$$

Ví dụ 1.12. Mỗi đề thi gồm 3 câu hỏi lấy trong 25 câu hỏi cho trước. Hỏi có thể lập được bao nhiêu đề thi có nội dung khác nhau?

Lời giải: Số đề thi có thể lập nên là $C_{25}^3 = \dfrac{25 \times 24 \times 23}{3!} = 2300$ (đề).

Chú ý 1.3. (a) Qui ước $0! = 1$.

(b) $C_n^k = C_n^{n-k}$.

(c) $C_n^k = C_{n-1}^{k-1} + C_{n-1}^k$.

(d) Khai triển nhị thức Niu–tơn ($a, b \in \mathbb{R}$, $n \in \mathbb{N}^*$)

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^{n-k} b^k = C_n^0 a^n + C_n^1 a^{n-1} b + \cdots + C_n^{n-1} a b^{n-1} + C_n^n b^n.$$

1.3 Khái niệm và các định nghĩa xác suất

1.3.1 Khái niệm xác suất

Mọi sự kiện ngẫu nhiên đều giống nhau ở chỗ chúng không chắc chắn, nhưng khả năng xảy ra của từng sự kiện lại có thể khác nhau. Để đặc trưng cho khả năng xảy ra (xuất hiện) của các sự kiện người ta dùng các con số, sự kiện nào có khả năng xảy ra nhiều hơn được đặc trưng bởi số lớn hơn. Con số đặc trưng cho khả năng xuất hiện của một sự kiện gọi là xác suất của sự kiện đó.

Định nghĩa 1.8 (Xác suất)**.** Xác suất (probability) của một sự kiện A là một số nằm giữa 0 và 1, số này đo lường khả năng xuất hiện của sự kiện đó khi phép thử được thực hiện.

Ký hiệu là $P(A)$.

1.3.2 Định nghĩa cổ điển về xác suất

Định nghĩa 1.9 (Định nghĩa cổ điển về xác suất)**.** Giả sử trong một phép thử có n kết cục đồng khả năng có thể xảy ra, trong đó có m kết cục thuận lợi cho sự kiện A. Khi đó,

$$\boxed{P(A) = \frac{m}{n} = \frac{\text{số kết cục thuận lợi cho } A}{\text{tổng số kết cục có thể}}} \tag{1.5}$$

Từ định nghĩa này ta suy ra các tính chất sau đây của xác suất.

Tính chất 1.2. (a) $0 \le P(A) \le 1$, A là sự kiện bất kỳ.

(b) $P(S) = 1$.

(c) $P(\emptyset) = 0$.

(d) Nếu $A \subset B$ thì $P(A) \le P(B)$.

Ví dụ 1.13. Một người khi gọi điện thoại quên mất 2 số cuối cùng của số điện thoại cần gọi mà chỉ nhớ được rằng chúng khác nhau. Tìm xác suất để người đó chọn ngẫu nhiên một số để gọi thì được đúng số cần gọi.

Lời giải: Gọi A là sự kiện "chọn ngẫu nhiên một số để gọi thì được đúng số cần gọi".

Số kết cục đồng khả năng là $n = A_{10}^2$.

Số kết cục thuận lợi cho A là $m = 1$.

Vậy $P(A) = \dfrac{m}{n} = \dfrac{1}{90}$.

Ví dụ 1.14. Từ bộ bài tú-lơ-khơ 52 cây đã trộn kỹ rút ngẫu nhiên ra 2 cây. Tính xác suất xảy ra các sự kiện sau:

(a) Hai cây rút ra đều là Át.

(b) Hai cây rút ra có 1 cây Át, 1 cây K.

Lời giải: Số kết cục đồng khả năng $n = C_{52}^2 = 1326$.

(a) Số kết cục thuận lợi cho sự kiện A "hai cây rút ra đều là Át" là $m_A = C_4^2$. Vậy

$$P(A) = \frac{m_A}{n} = \frac{C_4^2}{1326} = \frac{1}{221}.$$

(b) Số kết cục thuận lợi cho sự kiện B "hai cây rút ra có 1 cây Át, 1 cây K" là $m_B = C_4^1 \times C_4^1$, suy ra $P(B) = \dfrac{m_B}{n} = \dfrac{C_4^1 \times C_4^1}{1326} = \dfrac{8}{663}$.

Ví dụ 1.15. Một đoàn tàu có 4 toa được đánh số I, II, III, IV đỗ ở sân ga. Có 6 hành khách từ sân ga lên tàu. Mỗi người độc lập với nhau chọn ngẫu nhiên một toa. Tính xác suất để:

(a) toa I có 3 người, toa II có 2 người và toa III có 1 người;

(b) một toa có 3 người, một toa 2 người, một toa có 1 người;

(c) mỗi toa có ít nhất 1 người.

Lời giải: Số trường hợp đồng khả năng có thể có là $n = 4^6 = 4096$.

(a) Số trường hợp thuận lợi cho sự kiện A "toa I có 3 người, toa II có 2 người và toa III có 1 người" là $C_6^3 \times C_3^2 \times C_1^1 = 60$, suy ra $P(A) = \dfrac{60}{4096} \simeq 0,0146$.

(b) Số trường hợp thuận lợi cho sự kiện B "một toa có 3 người, một toa 2 người, một toa có 1 người" là $C_6^3 \times 4 \times C_3^2 \times 3 \times C_1^1 \times 2 = 1440$, suy ra $P(B) = \dfrac{1440}{4096} \simeq 0,3516$.

(b) Gọi C là sự kiện "mỗi toa có ít nhất 1 người". Số trường hợp thuận lợi cho sự kiện C là $C_6^3 \times 4 \times 3! + C_6^2 \times C_4^2 \times C_4^2 \times 2! = 480 + 1080 = 1560$. Do đó, $P(C) = \dfrac{1560}{4096} \simeq 0,3809$.

Ví dụ 1.16. Ba nữ nhân viên phục vụ A, B và C thay nhau rửa đĩa chén và giả sử ba người này đều "khéo léo" như nhau. Trong một tháng có 4 chén bị vỡ. Tìm xác suất để:

(a) chị A đánh vỡ 3 chén và chị B đánh vỡ 1 chén;

(b) một trong ba người đánh vỡ 3 chén;

(c) một trong ba người đánh vỡ cả 4 chén.

Lời giải: Số kết cục đồng khả năng có thể có là $n = 3^4$.

(a) Số kết cục thuận lợi cho sự kiện D "chị A đánh vỡ 3 chén và chị B đánh vỡ 1 chén" là $m_D = C_4^3 \times 1 = 4$, suy ra $P(D) = \dfrac{4}{81} \simeq 0,0494$.

(b) Số kết cục thuận lợi cho sự kiện E "một trong 3 người đánh vỡ 3 chén" là $m_E = C_3^1 \times C_4^3 \times 2 = 24$, nên $P(E) = \dfrac{24}{81} \simeq 0,2963$.

(c) Số kết cục thuận lợi cho sự kiện F "một trong 3 người đánh vỡ 4 chén" là $m_F = C_3^1 \times C_4^4 = 3$. Vậy $P(F) = \dfrac{3}{81} \simeq 0,037$.

Nhận xét 1.2. Định nghĩa cổ điển về xác suất có ưu điểm là dễ vận dụng tuy nhiên định nghĩa này chỉ áp dụng được với các phép thử có hữu hạn kết cục đồng khả năng xảy ra. Trong trường hợp có vô hạn kết cục đồng khả năng ta sử dụng định nghĩa xác suất theo quan điểm hình học.

1.3.3 Định nghĩa xác suất theo quan điểm hình học

Định nghĩa 1.10 (Định nghĩa xác suất theo quan điểm hình học)**.** Giả sử tập hợp vô hạn các kết cục đồng khả năng của một phép thử có thể biểu thị bởi một miền hình học G (đo được, hữu hạn, khác 0), còn các kết cục thuận lợi cho A bởi miền con H của G. Khi đó

$$\boxed{P(A) = \frac{\text{độ đo } H}{\text{độ đo } G}} \tag{1.6}$$

Chú ý 1.4. Tùy theo G là đoạn thẳng, miền phẳng hay khối không gian mà độ đo được hiểu là độ dài, diện tích hay thể tích.

Ví dụ 1.17. Hai người bạn hẹn gặp nhau ở một địa điểm trong khoảng thời gian từ 7h00 đến 8h00. Mỗi người có thể đến điểm hẹn một cách ngẫu nhiên tại một thời điểm trong khoảng thời gian nói trên và họ quy ước rằng ai đến trước thì chỉ đợi người kia trong vòng 10 phút. Tính xác suất để hai người gặp nhau.

Lời giải: Gọi x, y lần lượt là thời điểm đến điểm hẹn của hai người, $0 \le x, y \le 60$. Vậy mỗi cặp thời điểm đến (x, y) của hai người là một điểm của miền

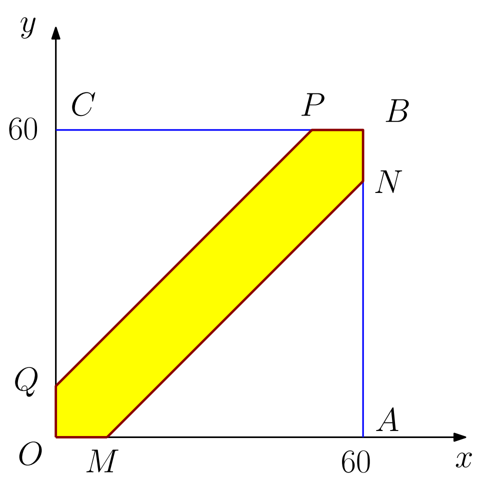
$$G = \{(x,y) \in \mathbb{R}^2 : 0 \le x \le 60;\ 0 \le y \le 60\} \quad \text{(hình vuông } OABC\text{)}.$$

Gọi E là sự kiện "hai người gặp nhau", khi đó E được biểu diễn bởi

$$H = \{(x,y) \in G : |x-y| \le 10\} \quad \text{(đa giác } OMNBPQ\text{)}.$$

Sử dụng định nghĩa xác suất theo quan điểm hình học,

$$P(E) = \frac{\text{diện tích } H}{\text{diện tích } G} = \frac{\text{diện tích } (OMNBPQ)}{\text{diện tích } (OABC)} = \frac{60^2 - 50^2}{60^2} = \frac{11}{36} \simeq 0,3056.$$



Hình 1.4: Minh họa cho Ví dụ 1.17

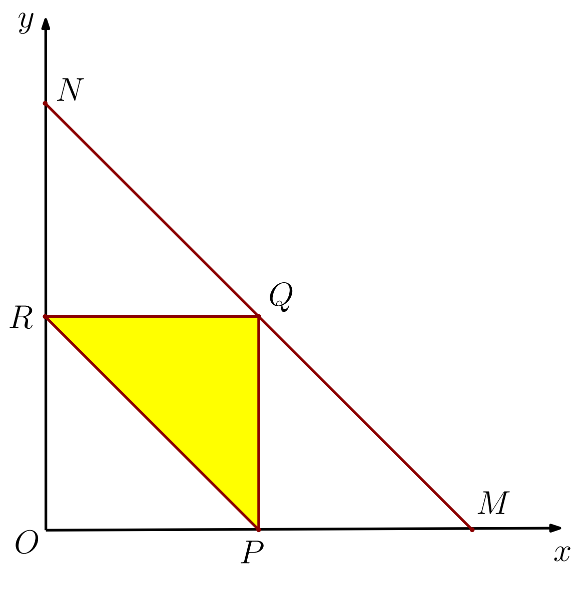
Ví dụ 1.18. Cho đoạn thẳng AB có độ dài 10cm. Lấy hai điểm C, D bất kỳ trên đoạn AB (C nằm giữa A và D). Tính xác suất độ dài AC, CD, DB tạo thành 3 cạnh của một tam giác.

Lời giải: Gọi x là độ dài đoạn AC, y là độ dài đoạn CD thì độ dài đoạn DB là $10 - x - y$. Khi đó ta có điều kiện $0 \le x \le 10$, $0 \le y \le 10$ và $0 \le x + y \le 10$. Tập hợp các giá trị (x, y) thỏa mãn điều kiện này tương ứng với miền

$$G = \{(x,y) \in \mathbb{R}^2 : 0 \le x \le 10,\ 0 \le y \le 10,\ 0 \le x + y \le 10\} \quad \text{(tam giác } OMN).$$

Độ dài các đoạn AC, CD, DB tạo thành 3 cạnh một tam giác phải thỏa mãn tính chất "tổng hai cạnh lớn hơn một cạnh", tức là $x + y > 10 - x - y$, $x + (10 - x - y) > y$, $y + (10 - x - y) > x$ hay $x + y > 5$, $x < 5$ và $y < 5$. Tập các giá trị (x, y) thỏa mãn điều kiện này tương ứng với miền

$$H = \{(x,y) \in G : x + y > 5,\ x < 5,\ y < 5\} \quad \text{(tam giác } PQR).$$



Hình 1.5: Minh họa cho Ví dụ 1.18

Theo định nghĩa hình học, xác suất cần tìm là $p = \dfrac{\text{diện tích tam giác } (PQR)}{\text{diện tích tam giác } (OMN)} = \dfrac{1}{4} = 0,25.$

Ví dụ 1.19. Trên mặt phẳng đã kẻ sẵn các đường thẳng song song cách đều nhau một khoảng có độ dài $2a$, người ta gieo ngẫu nhiên một chiếc kim dài $2b$ ($b < a$). Tính xác suất sao cho kim cắt một đường thẳng trong số những đường thẳng đó.

Lời giải: Gọi x là khoảng cách từ trung điểm của kim đến đường thẳng song song gần nhất và φ là góc mà kim tạo với các đường này. Ta có $0 \le x \le a$, $0 \le \varphi \le \pi$. Như vậy có thể biểu diễn miền đồng khả năng bởi hình chữ nhật $G = [a, \pi] \times [a, \pi]$. Miền thuận lợi cho sự kiện kim cắt đường thẳng song song là $H = \{(x, \varphi) \in G : 0 \le x \le b \sin\varphi;\ 0 \le \varphi \le \pi\}$. Do đó,

$$p = \frac{\text{diện tích } H}{\text{diện tích } G} = \frac{\displaystyle\int_0^{\pi} b \sin\varphi d\varphi}{a \times \pi} = \frac{2b}{a\pi}.$$

Nhận xét 1.3. Định nghĩa cổ điển về xác suất và định nghĩa xác suất theo quan điểm hình học chỉ áp dụng được với các phép thử có kết cục đồng khả năng xảy ra. Trong nhiều bài toán thực tế, việc tính hết các kết cục của một phép thử không dễ dàng, bên cạnh đó điều kiện các kết cục đồng khả năng trên thực tế thường khó thỏa mãn.

1.3.4 Định nghĩa thống kê về xác suất

Định nghĩa 1.11 (Tần suất)**.** Giả sử trong một điều kiện nào đó ta có thể lặp lại n lần một phép thử và thấy có m lần xuất hiện sự kiện A. Khi đó, tỷ số $\dfrac{m}{n}$ gọi là tần suất xuất hiện A, ký hiệu là $f(A)$.

Như vậy

$$\boxed{f(A) = \frac{m}{n}} \tag{1.7}$$

Ví dụ 1.20. Để xác định tần suất xuất hiện mặt sấp khi tung một đồng xu nhiều lần, người ta ghi lại kết quả sau:

Người thí nghiệm	Số lần tung n	Số lần xuất hiện mặt sấp m	Tần suất f
Buýp-phông	4040	2048	0,5080
Piêc-sơn	12000	6019	0,5016
Piêc-sơn	24000	12012	0,5005

Nhận xét 1.4. Tần suất của sự kiện A có tính chất ổn định, nghĩa là nó dao động rất ít xung quanh một số xác định p nào đó khi số phép thử khá lớn. Số ấy gọi là xác suất của sự kiện A theo quan điểm thống kê.

Định nghĩa 1.12 (Định nghĩa thống kê về xác suất)**.** Nếu tần suất xuất hiện sự kiện A luôn luôn dao động xung quanh một số xác định p nào đó và khi số phép thử tăng lên khá lớn mà tần suất xuất hiện sự kiện A càng gần tới p thì số p được gọi là xác suất của sự kiện A theo quan điểm thống kê.

Chú ý 1.5. Bằng định nghĩa thống kê về xác suất, người ta đã tìm được xác suất để sinh con trai trong mỗi lần sinh là $p = 0,518$, con số này hầu như không thay đổi theo thời gian, địa phương và chủng tộc.

(a) Nhà toán học Láp-la–xơ trong 10 năm liền theo dõi ở thành phố Pê–tec–bua, Luân–đôn và Béc–lin thấy tỷ số đó là $22/43$. Ông cũng đã theo dõi 40 năm liền ở Pa–ri thấy tỷ số đó là $25/49$.

(b) Nhà toán học Crame theo dõi ở Thụy–điển năm 1935 cũng thấy tỷ số đó là 0,518.

Nhận xét 1.5. (a) Định nghĩa thống kê của xác suất khắc phục được một nhược điểm của định nghĩa cổ điển là không dùng đến khái niệm đồng khả năng.

(b) Định nghĩa này không giúp ta tính được chính xác xác suất của một sự kiện mà chỉ tìm được giá trị gần đúng; đồng thời số phép thử phải đủ lớn và chỉ dùng được cho các phép thử ngẫu nhiên có thể lặp lại nhiều lần một cách độc lập trong các điều kiện giống nhau.

Các định nghĩa trên về xác suất giúp ta một cách tích cực trong việc tính xác suất, nhưng mỗi định nghĩa đều có nhược điểm của nó. Để khắc phục các nhược điểm đó, năm 1933 nhà toán học Xô–viết Can–mơ–gơ–rôp đã đưa ra xác suất theo phương pháp tiên đề. Tuy nhiên ta không đề cập đến trong chương trình này.

1.3.5 Nguyên lý xác suất nhỏ, nguyên lý xác suất lớn

1.3.5a Nguyên lý xác suất nhỏ

Sự kiện không thể có có xác suất bằng 0, một sự kiện có xác suất gần bằng 0 vẫn có thể xảy ra khi thực hiện một số lớn các phép thử. Tuy nhiên qua thực nghiệm và quan sát thực tế, người ta thấy rằng các sự kiện có xác suất nhỏ sẽ không xảy ra khi ta chỉ thực hiện một phép thử hay một vài phép thử. Từ đó ta thừa nhận nguyên lý sau đây, gọi là "Nguyên lý xác suất nhỏ": *Nếu một sự kiện có xác suất rất nhỏ thì thực tế có thể cho rằng trong một phép thử sự kiện đó sẽ không xảy ra".*

Nhận xét 1.6. (a) Mức xác suất được coi là nhỏ tùy thuộc vào từng bài toán cụ thể và gọi là mức ý nghĩa, ký hiệu là α.

(b) Nguyên lý xác suất nhỏ là cơ sở của phương pháp kiểm định (sẽ được đề cập ở Chương 5).

1.3.5b Nguyên lý xác suất lớn

Tương tự như trên, ta có thể đưa ra nguyên lý xác suất lớn: *Nếu sự kiện A có xác suất gần bằng 1 thì trên thực tế có thể cho rằng trong một phép thử sự kiện đó sẽ xảy ra".*

Nhận xét 1.7. (a) Mức xác suất đủ lớn gọi là độ tin cậy, ký hiệu là $\gamma = 1 - \alpha$. Việc quy định một mức xác suất thế nào là lớn sẽ tùy thuộc vào từng bài toán cụ thể.

(b) Nguyên lý xác suất lớn là cơ sở của phương pháp ước lượng bằng khoảng tin cậy (sẽ được đề cập ở Chương 4).

1.4 Công thức cộng và nhân xác suất

1.4.1 Xác suất có điều kiện

Định nghĩa 1.13 (Xác suất có điều kiện)**.** Giả sử trong một phép thử ta có $P(B) > 0$. Khi đó xác suất có điều kiện của sự kiện A nào đó, biết rằng đã có B, là một số không âm ký hiệu là

$$\boxed{P(A|B) = \frac{P(AB)}{P(B)}} \tag{1.8}$$

Tương tự

$$\boxed{P(B|A) = \frac{P(AB)}{P(A)}, \quad P(A) > 0} \tag{1.9}$$

Nhận xét 1.8. (a) Xác suất điều kiện có mọi tính chất của một xác suất bình thường, chẳng hạn $P(A|B) \geq 0$, $P(A|A) = 1$.

(b) Ta có thể tính xác suất điều kiện bằng cách áp dụng các công thức (1.8) hoặc (1.9) hoặc tính trực tiếp.

Ví dụ 1.21. Từ một bộ bài tú-lơ-khơ 52 cây đã trộn kỹ rút ngẫu nhiên ra một cây bài. Biết đó là cây đen, tính xác suất đó là cây át.

Lời giải: Gọi A là sự kiện "rút được cây át" và B là sự kiện "rút được cây đen". Xác suất cần tính là $P(A|B) = \dfrac{2}{26}$.

Ví dụ 1.22. Trong một thùng kín có N quả cầu giống nhau trong đó có M cầu trắng ($M < N$). Lấy ngẫu nhiên lần lượt không hoàn lại 2 quả cầu. Tính xác suất để cầu thứ hai lấy ra là trắng, biết rằng cầu thứ nhất lấy ra đã là trắng.

Lời giải: Gọi A_i là sự kiện "cầu thứ i lấy ra là trắng", $i = 1, 2$. Sử dụng công thức (1.8) ta được

$$P(A_2|A_1) = \frac{P(A_1A_2)}{P(A_1)} = \frac{\frac{M\times(M-1)}{N\times(N-1)}}{\frac{M}{N}} = \frac{M-1}{N-1}.$$

1.4.2 Công thức nhân xác suất

1.4.2a Tính độc lập của các sự kiện

Định nghĩa 1.14 (Sự kiện độc lập)**.** (a) Hai sự kiện A và B được gọi là độc lập với nhau nếu sự kiện này xảy ra hay không xảy ra không làm ảnh hưởng tới khả năng xảy ra của sự kiện kia, nghĩa là $P(A|B) = P(A|\overline{B}) = P(A)$, $P(B|A) = P(B|\overline{A}) = P(B)$.

(b) Các sự kiện $A_1, A_2, \ldots, A_n$ được gọi là độc lập từng đôi với nhau nếu mỗi cặp 2 trong n sự kiện đó độc lập với nhau.

(c) Các sự kiện $A_1, A_2, \dots, A_n$ được gọi là độc lập trong tổng thể nếu mỗi sự kiện trong chúng độc lập với tích của một số bất kỳ sự kiện trong các sự kiện còn lại.

Chú ý 1.6. (a) Nếu A và B độc lập thì các cặp A và $\overline{B}$; $\overline{A}$ và B; $\overline{A}$ và $\overline{B}$ cũng độc lập.

(b) Thông thường tính độc lập của các sự kiện được suy ra từ ý nghĩa thực tế.

1.4.2b Công thức nhân xác suất

(a) Nếu A và B là hai sự kiện bất kỳ thì

$$\boxed{P(AB) = P(A)P(B|A) = P(B)P(A|B)} \tag{1.10}$$

Nếu A và B là hai sự kiện độc lập thì

$$\boxed{P(AB) = P(A)P(B)} \tag{1.11}$$

(b) Mở rộng cho tích n sự kiện bất kỳ $A_1, A_2, \dots, A_n$:

$$\boxed{P(A_1A_2\dots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\dots P(A_n|A_1A_2\dots A_{n-1})} \tag{1.12}$$

Nếu $A_1, A_2, \dots, A_n$ độc lập trong tổng thể, thì:

$$\boxed{P(A_1A_2\dots A_n) = P(A_1)P(A_2)\dots P(A_n)} \tag{1.13}$$

Nhận xét 1.9. Công thức nhân (1.11) cung cấp cho ta một phương pháp dễ thực hành để kiểm tra tính độc lập của hai sự kiện ngẫu nhiên.

Ví dụ 1.23. Có 4 que thăm, trong đó có 3 que thăm dài bằng nhau và 1 que thăm ngắn hơn. Bốn người lần lượt lên rút ngẫu nhiên một que thăm. Tính xác suất người thứ i rút được thăm ngắn ($i = 1, 2, 3, 4$).

Lời giải: Gọi A_i là sự kiện "người thứ i rút được thăm ngắn", $i = 1, 2, 3, 4$. Khi đó,

$$P(A_1) = \frac{1}{4}.$$
$$P(A_2) = P(\overline{A}_1A_2) = P(\overline{A}_1)P(A_2|\overline{A}_1) = \frac{3}{4} \times \frac{1}{3} = \frac{.1}{4}$$
$$P(A_3) = P(\overline{A}_1\overline{A}_2A_3) = P(\overline{A}_1)P(\overline{A}_2|\overline{A}_1)P(A_3|\overline{A}_1\overline{A}_2) = \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2} = \frac{1}{4}.$$
$$P(A_4) = P(\overline{A}_1\overline{A}_2\overline{A}_3A_4) = P(\overline{A}_1)P(\overline{A}_2|\overline{A}_1)P(\overline{A}_3|\overline{A}_1\overline{A}_2)P(A_4|\overline{A}_1\overline{A}_2\overline{A}_3) = \frac{1}{4}.$$

Vậy khả năng rút được thăm ngắn của 4 người là như nhau và bằng $\frac{1}{4}$.

Ví dụ 1.24. Một tổ có 15 sinh viên trong đó có 5 sinh viên học giỏi môn Xác suất thống kê. Chia tổ này thành 5 nhóm, mỗi nhóm 3 người. Tính xác suất để nhóm nào cũng có một sinh viên học giỏi môn Xác suất thống kê.

Lời giải: Gọi A là sự kiện "nhóm nào cũng có một sinh viên học giỏi môn Xác suất thống kê"; A_i là sự kiện "nhóm i có một sinh viên học giỏi môn Xác suất thống kê", $i = 1, \ldots, 5$. Khi đó $A = A_1A_2A_3A_4A_5$. Sử dụng công thức nhân (1.12)

$$P(A) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)P(A_4|A_1A_2A_3)P(A_5|A_1A_2A_3A_4),$$

trong đó

$$P(A_1) = \frac{C_5^1 \times C_{10}^2}{C_{15}^3} = \frac{45}{91}, \qquad P(A_2|A_1) = \frac{C_4^1 \times C_8^2}{C_{12}^3} = \frac{28}{55},$$
$$P(A_3|A_1A_2) = \frac{C_3^1 \times C_6^2}{C_9^3} = \frac{15}{28}, \quad P(A_4|A_1A_2A_3) = \frac{C_2^1 \times C_4^2}{C_6^3} = \frac{3}{5},$$
$$P(A_5|A_1A_2A_3A_4) = \frac{C_1^1 \times C_2^2}{C_3^3} = 1.$$

Vậy $P(A) = \dfrac{81}{1001} \simeq 0,0809$.

Ví dụ 1.25 (Đề thi MI2020 kỳ 20151)**.** Ra khỏi phòng khách, 6 người cùng xỏ ngẫu nhiên vào một đôi giày trong bóng tối. Mỗi người chỉ có thể phân biệt chiếc giày phải với chiếc giày trái, còn không thể phân biệt được giày của mình với giày của người khác. Tính xác suất để

(a) Mỗi người khách xỏ vào đúng đôi giày của mình.

(b) Mỗi người khách xỏ vào đúng hai chiếc giày của cùng một đôi nào đó.

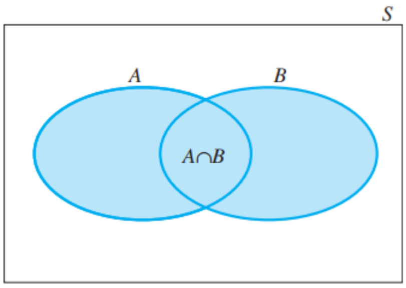
Lời giải:

(a) Gọi A là sự kiện "cả 6 người khách đều xỏ đúng đôi giày của mình"; A_i là sự kiện "người thứ i xỏ đúng đôi giày của mình", $i = 1, 2, \ldots, 6$. Khi đó $A = A_1A_2A_3A_4A_5A_6$ và

$$P(A) = P(A_1)P(A_2|A_1)\ldots P(A_6|A_1A_2A_3A_4A_5) = \frac{1}{6^2} \times \frac{1}{5^2} \times \cdots \times \frac{1}{1^2} = \frac{1}{(6!)^2}.$$

(b) Gọi B là sự kiện "mỗi người khách đều xỏ đúng 2 chiếc giày của cùng một đôi"; B_i là sự kiện "người thứ i xỏ đúng 2 chiếc giày của cùng một đôi", $i = 1, 2, \ldots, 6$. Khi đó $B = B_1B_2B_3B_4B_5B_6$ và

$$P(B) = P(B_1)P(B_2|B_1)\ldots P(B_6|B_1B_2B_3B_4B_5) = \frac{1}{6} \times \frac{1}{5} \times \cdots \times \frac{1}{1} = \frac{1}{6!}.$$



Hình 1.6: Minh họa công thức cộng

1.4.3 Công thức cộng xác suất

(a) Nếu A và B là hai sự kiện bất kỳ thì

$$\boxed{P(A+B) = P(A) + P(B) - P(AB)} \tag{1.14}$$

Nếu A và B là hai sự kiện xung khắc thì

$$\boxed{P(A+B) = P(A) + P(B)} \tag{1.15}$$

(b) Nếu A, B và C là ba sự kiện bất kỳ thì

$$\boxed{P(A+B+C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)} \tag{1.16}$$

(c) Mở rộng cho tổng n sự kiện bất kỳ $A_1, A_2, \ldots, A_n$:

$$P\Big(\sum_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_iA_j) + \sum_{i<j<k} P(A_iA_jA_k) - \ldots + (-1)^{n-1}P(A_1A_2\ldots A_n). \tag{1.17}$$

Nếu $A_1, A_2, \ldots, A_n$ xung khắc từng đôi thì

$$\boxed{P\Big(\sum_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} P(A_i)} \tag{1.18}$$

Đặc biệt:

(a) Nếu $A_1, A_2, \ldots, A_n$ là hệ đầy đủ các sự kiện thì $\sum_{i=1}^{n} P(A_i) = 1$.

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

(c) $P(A) = P(AB) + P(A\overline{B})$.

Ví dụ 1.26. Một lớp có 100 sinh viên, trong đó có 40 sinh viên giỏi ngoại ngữ, 30 sinh viên giỏi toán xác suất, 20 sinh viên giỏi cả ngoại ngữ lẫn toán xác suất. Chọn ngẫu nhiên một sinh viên trong lớp. Tìm xác suất để sinh viên đó giỏi ít nhất 1 trong 2 môn trên.

Lời giải: Gọi A là sự kiện "sinh viên đó giỏi ít nhất 1 trong 2 môn ngoại ngữ, toán xác suất"; N là sự kiện "sinh viên đó giỏi ngoại ngữ"; T là sự kiện "sinh viên đó giỏi toán xác suất". Khi đó, $A = T + N$. Suy ra

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

Ví dụ 1.27. Ba xạ thủ độc lập với nhau cùng bắn súng vào bia. Xác suất bắn trúng bia của xạ thủ thứ nhất, thứ hai, thứ ba tương ứng là 0,7, 0,8 và 0,9. Tính xác suất để:

(a) có duy nhất một xạ thủ bắn trúng bia;

(b) có đúng hai xạ thủ bắn trúng bia;

(c) có ít nhất một xạ thủ bắn trúng bia;

(d) xạ thủ thứ nhất bắn trúng bia biết rằng có hai xạ thủ bắn trúng bia.

Lời giải: Gọi A_i là các sự kiện "xạ thủ A, B, C bắn trúng bia" tương ứng, $i = 1, 2, 3$.

(a) Gọi A là sự kiện "có duy nhất một xạ thủ bắn trúng bia". Khi đó,

$$A = A_1\overline{A}_2\overline{A}_3 + \overline{A}_1A_2\overline{A}_3 + \overline{A}_1\overline{A}_2A_3.$$

Sử dụng công thức cộng (1.18) và công thức nhân (1.13) trong trường hợp các sự kiện xung khắc và độc lập suy ra

$$\begin{aligned} P(A) &= P(A_1)P(\overline{A}_2)P(\overline{A}_3) + P(\overline{A}_1)P(A_2)P(\overline{A}_3) + P(\overline{A}_1)P(\overline{A}_2)P(A_3) \\ &= 0,7 \times 0,2 \times 0,1 + 0,3 \times 0,8 \times 0,1 + 0,3 \times 0,2 \times 0,9 = 0,092. \end{aligned}$$

(b) Gọi B là sự kiện "có đúng hai xạ thủ bắn trúng bia". Khi đó,

$$B = A_1A_2\overline{A}_3 + A_1\overline{A}_2A_3 + \overline{A}_1A_2A_3.$$

Làm tương tự ý (a), $P(B) = 0,398$.

(c) Gọi C là sự kiện "ít nhất một xạ thủ bắn trúng bia", khi đó

Hoặc $C = A_1 + A_2 + A_3$,

$$\begin{aligned} P(C) &= P(A_1) + P(A_2) + P(A_3) - P(A_1A_2) - P(A_1A_3) - P(A_2A_3) + P(A_1A_2A_3) \\ &= 0,994. \end{aligned}$$

Hoặc $\overline{C} = \overline{A}_1\overline{A}_2\overline{A}_3$,

$$P(\overline{C}) = P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3), \quad P(C) = 1 - P(\overline{C}) = 1 - 0,006 = 0,994.$$

Hoặc $C = A + B + A_1A_2A_3$, $P(C) = P(A) + P(B) + P(A_1A_2A_3) = 0,994$.

(d) $P(A_1|B) = \dfrac{P(A_1B)}{P(B)} = \dfrac{P(A_1A_2\overline{A}_3 + A_1\overline{A}_2A_3)}{P(B)} = \dfrac{0,182}{0,398} = 0,46$.

Ví dụ 1.28 (Đề thi MI2020 kỳ 20182)**.** Cho ba sự kiện A, B, C độc lập từng đôi thỏa mãn $P(A) = P(B) = P(C) = p$ và $P(ABC) = 0$.

(a) Tính $P(AB\overline{C})$; $P(A\overline{B}\,\overline{C})$; $P(\overline{A}\,\overline{B}\,\overline{C})$.

(b) Tìm giá trị p lớn nhất có thể có.

Lời giải:

(a) Vì $AB\overline{C} + ABC = AB$; $AB\overline{C}$ và ABC xung khắc; A và B độc lập, nên

$$P(AB\overline{C}) = P(AB) - P(ABC) = P(A)P(B) - 0 = p^2.$$

Vì $A = A\overline{B}\,\overline{C} + ABC + A\overline{B}C + AB\overline{C}$, sử dụng tính xung khắc của các sự kiện,

$$P(A\overline{B}\,\overline{C}) = P(A) - P(ABC) - P(A\overline{B}C) - P(AB\overline{C}) = p - 2p^2.$$

Vì $\overline{A}\,\overline{B}\,\overline{C} + A\overline{B}\,\overline{C} = \overline{B}\,\overline{C}$ nên $P(\overline{A}\,\overline{B}\,\overline{C}) = P(\overline{B}\,\overline{C}) - P(A\overline{B}\,\overline{C}) = 1 - 3p + 3p^2$.

(b) Từ ý (a) và đầu bài ta có $P(ABC) = 0$, $P(AB\overline{C}) = P(A\overline{B}C) = P(\overline{A}BC) = p^2$, $P(A\overline{B}\,\overline{C}) = P(\overline{A}B\overline{C}) = P(\overline{A}\,\overline{B}C) = p - 2p^2$, $P(\overline{A}\,\overline{B}\,\overline{C}) = 1 - 3p + 3p^2$. Khi đó p thỏa mãn hệ

$$\begin{cases} 0 \le p^2 \le 1, \\ 0 \le p - 2p^2 \le 1, \\ 0 \le 1 - 3p + 3p^2 \le 1. \end{cases}$$

Hệ này tương đương với $0 \le p \le 0,5$. Vậy giá trị p lớn nhất là $0,5$.

Ví dụ 1.29 (Đề thi MI2020 kỳ 20183)**.** Có một nhóm 4 sinh viên, mỗi người có một chiếc mũ giống hệt nhau để trên giá. Khi ra khỏi phòng, mỗi người lấy ngẫu nhiên một chiếc mũ để đội. Tính xác suất để:

(a) Sinh viên thứ nhất và sinh viên thứ hai lấy đúng mũ của mình.

(b) Có ít nhất một sinh viên lấy đúng mũ của mình.

Lời giải: Gọi A là sự kiện "có ít nhất một sinh viên lấy đúng mũ của mình"; A_i là sự kiện "sinh viên thứ i lấy đúng mũ của mình", $i = 1,2,3,4$.

(a) $P(A_1A_2) = P(A_1)P(A_2|A_1) = \dfrac{3!}{4!} \times \dfrac{2!}{3!} = \dfrac{1}{12} \simeq 0,0833$.

(b)

$$\begin{aligned}
P(A) &= P(A_1 + A_2 + A_3 + A_4) \\
&= \sum_{i=1}^{4} P(A_i) - P(A_1A_2) - P(A_1A_3) - P(A_1A_4) - P(A_2A_3) \\
&\quad - P(A_2A_4) - P(A_3A_4) + P(A_1A_2A_3) + P(A_1A_2A_4) \\
&\quad + P(A_1A_3A_4) + P(A_2A_3A_4) - P(A_1A_2A_3A_4) \\
&= 4 \times \frac{1}{4} - 6 \times \frac{1}{12} + 4 \times \frac{1}{24} - \frac{1}{24} = 0,625.
\end{aligned}$$

Ví dụ 1.30 (Đề thi MI2020 kỳ 20191)**.** Lớp MI2020 có 80 sinh viên trong đó có 20 sinh viên thuộc tổ I, 25 sinh viên thuộc tổ II và 35 sinh viên thuộc tổ III. Chọn ngẫu nhiên 10 sinh viên trong lớp tham dự trại hè. Tính xác suất để mỗi tổ có ít nhất 1 sinh viên được chọn.

Lời giải: Gọi A là sự kiện "Mỗi tổ có ít nhất 1 sinh viên được chọn", A_i: "tổ i có ít nhất 1 sinh viên được chọn", $i = 1, 2, 3$. Khi đó, $A = A_1A_2A_3$ và

$$P(A) = P(A_1A_2A_3) = 1 - P(\overline{A}_1 + \overline{A}_2 + \overline{A}_3).$$

Sử dụng công thức (1.16),

$$\begin{aligned}
P(\overline{A}_1 + \overline{A}_2 + \overline{A}_3) &= P(\overline{A}_1) + P(\overline{A}_2) + P(\overline{A}_3) - P(\overline{A}_1\overline{A}_2) - P(\overline{A}_1\overline{A}_3) - P(\overline{A}_2\overline{A}_2) + P(\overline{A}_1\overline{A}_2\overline{A}_3) \\
&= \frac{1}{C_{80}^{10}}\Big[(C_{60}^{10} + C_{55}^{10} + C_{45}^{10}) - (C_{35}^{10} + C_{25}^{10} + C_{20}^{10}) + 0\Big] \\
&\simeq 1 - 0,06538 = 0,93462.
\end{aligned}$$

Ví dụ 1.31 (Đề thi MI2020 kỳ 20161)**.** Biết từ vị trí A đến B có hai đường đi với xác suất bị ngập của mỗi con đường là p; từ B đến C cũng có hai đường đi với xác suất bị ngập của mỗi con đường cũng là p. Biết đường đi từ A đến C bị ngập, tính xác suất để đường đi từ A đến B không bị ngập.

Lời giải: Gọi E_{AB} là sự kiện "đường đi từ A đến B không ngập", khi đó $\overline{E}_{AB}$ là sự kiện "đường đi từ A đến B bị ngập". Xác suất cần tìm là

$$P(E_{AB}|\overline{E}_{AC}) = \frac{P[(E_{AB})(\overline{E}_{AC})]}{P(\overline{E}_{AC})} = \frac{P[(E_{AB})(\overline{E}_{BC})]}{P(\overline{E}_{AC})} = \frac{P(E_{AB})P(\overline{E}_{BC})}{P(\overline{E}_{AC})}.$$

Đường đi từ B đến C bị ngập nếu cả hai đường đi đều bị ngập, do đó xác suất để đường đi từ B đến C bị ngập là $P(\overline{E}_{BC}) = p^2$ và xác suất để đường đi từ A đến B không ngập là $P(E_{AB}) = 1 - p^2$.

Đường đi từ A đến C không ngập nếu đường đi từ A đến B không ngập và đường đi từ B đến C cũng không ngập, nên xác suất để đường đi từ A đến C bị ngập là $P(\overline{E}_{AC}) = 1 - (1 - p^2)^2$.

Vậy

$$P(E_{AB}|\overline{E}_{AC}) = \frac{(1 - p^2)p^2}{1 - (1 - p^2)^2}.$$

1.5 Công thức Béc–nu–li

1.5.1 Dãy phép thử độc lập

Định nghĩa 1.15 (Dãy phép thử độc lập)**.** Các phép thử được gọi là độc lập với nhau nếu xác suất để xảy ra một sự kiện nào đó trong từng phép thử sẽ không phụ thuộc vào việc sự kiện đó có xảy ra ở các phép thử khác hay không.

Ví dụ 1.32. Tung nhiều lần một đồng xu sẽ tạo nên các phép thử độc lập. Lấy nhiều lần sản phẩm từ một lô sản phẩm theo phương thức có hoàn lại cũng tạo nên các phép thử độc lập.

1.5.2 Lược đồ Béc–nu–li

1. Giả sử ta tiến hành n phép thử độc lập.
2. Trong mỗi phép thử chỉ có hai trường hợp: hoặc sự kiện A xảy ra, hoặc sự kiện A không xảy ra (tức là xảy ra $\overline{A}$).
3. Xác suất xảy ra A trong mỗi phép thử đều bằng p (tức là $P(A) = p$) và xác suất không xảy ra A trong mỗi phép thử đều bằng $q = 1 - p$ (tức là $P(\overline{A}) = 1 - p$).

Những bài toán thỏa mãn cả ba điều giả thiết trên được gọi là tuân theo lược đồ Béc-nu-li (hay dãy phép thử Béc-nu-li).

1.5.3 Công thức Béc–nu–li

Định lý 1.1. Trong lược đồ Béc–nu–li (hay dãy phép thử Béc–nu–li)

(a) Xác suất để sự kiện A xuất hiện đúng k lần, ký hiệu là $P_n(k)$, được xác định bởi

$$\boxed{P_n(k) = C_n^k p^k q^{n-k}, \quad k = 0, 1, \ldots, n} \tag{1.19}$$

(b) Xác suất để sự kiện A xuất hiện từ k_1 đến k_2 lần, ký hiệu là $P_n(k_1, k_2)$:

$$\boxed{P_n(k_1, k_2) = \sum_{k=k_1}^{k_2} P_n(k) = \sum_{k=k_1}^{k_2} C_n^k p^k q^{n-k}} \tag{1.20}$$

Chứng minh. (a) Gọi B là sự kiện "trong dãy n phép thử Béc–nu–li, sự kiện A xuất hiện đúng k lần". Ta thấy B có thể xảy ra nhiều phương án khác nhau miễn sao trong đó sự kiện A xuất hiện đúng k lần. Khi đó, có C_n^k phương án như vậy. Còn xác suất xảy ra một phương án sẽ là $p^k q^{n-k}$ do các phép thử là độc lập, sự kiện A xuất hiện k lần, sự kiện $\overline{A}$ xuất hiện $n - k$ lần. Từ đó ta có công thức cần chứng minh.

(b) Suy trực tiếp từ ý (a).

Nhận xét 1.10. Nếu một bài toán thỏa mãn lược đồ Béc–nu–li thì việc sử dụng công thức (1.19) hay (1.20) sẽ đơn giản hơn rất nhiều so với việc dùng các công thức nhân xác suất và cộng xác suất. Do đó chúng có ý nghĩa thực tiễn rất lớn.

Ví dụ 1.33. Trong phân xưởng có 5 máy hoạt động độc lập. Xác suất để trong mỗi ca mỗi máy bị hỏng đều bằng 0,1.

(a) Tìm xác suất để trong ca đó có đúng 2 máy hỏng.

(b) (Đề thi giữa kỳ 20182) Biết rằng trong một ca có đúng 2 máy hỏng. Tính xác suất để máy thứ nhất không hỏng.

Lời giải:

(a) Coi sự hoạt động của mỗi máy là một phép thử. Ta có 5 phép thử độc lập; trong mỗi phép thử chỉ có 2 trường hợp: hoặc máy hỏng, hoặc máy không hỏng; xác suất hỏng của mỗi máy đều bằng 0,1. Như vậy, bài toán thỏa mãn lược đồ Béc–nu–li với $n = 5$, $p = 0,1$ và $k = 2$. Áp dụng công thức Béc–nu–li (1.19) ta có xác suất cần tìm là:

$$P_5(2) = C_5^2 \times (0,1)^2 \times (0,9)^3 = 0,0729.$$

Nếu sử dụng công thức cộng và nhân xác suất với A là sự kiện "trong ca đó có đúng 2 máy hỏng", A_i là sự kiện "máy i bị hỏng trong ca", $i = 1, 2, \ldots, 5$, ta sẽ tính xác suất của A trên cơ sở phân tích:

$$\begin{aligned} A = {} & A_1A_2\overline{A}_3\overline{A}_4\overline{A}_5 + A_1\overline{A}_2A_3\overline{A}_4\overline{A}_5 + A_1\overline{A}_2\overline{A}_3A_4\overline{A}_5 + A_1\overline{A}_2\overline{A}_3\overline{A}_4A_5 + \\ & + \overline{A}_1A_2A_3\overline{A}_4\overline{A}_5 + \overline{A}_1A_2\overline{A}_3A_4\overline{A}_5 + \overline{A}_1A_2\overline{A}_3\overline{A}_4A_5 + \\ & + \overline{A}_1\overline{A}_2A_3A_4\overline{A}_5 + \overline{A}_1\overline{A}_2A_3\overline{A}_4A_5 + \overline{A}_1\overline{A}_2\overline{A}_3A_4A_5 \end{aligned}$$

và sử dụng tính xung khắc, tính độc lập của các sự kiện. Rõ ràng việc sử dụng công thức (1.19) cho ví dụ này đơn giản hơn rất nhiều.

(b) $P(\overline{A}_1|A) = \dfrac{P(\overline{A}_1)P(A|\overline{A}_1)}{P(A)} = \dfrac{0,9 \times C_4^2 \times (0,1)^2 \times (0,9)^2}{0,0729} = \dfrac{0,04374}{0,0729} = 0,6.$

Ví dụ 1.34. Hai vận động viên bóng bàn A và B đấu một trận gồm tối đa 5 ván (không có kết quả hòa sau mỗi ván và trận đấu sẽ dừng nếu một người nào đó thắng trước 3 ván). Xác suất để A thắng được ở một ván là 0,7.

(a) Tính các xác suất để A thắng sau x ván ($x = 3, 4, 5$).

(b) Tính xác suất để trận đấu kết thúc sau 5 ván.

Lời giải: (a) Việc A thắng sau x ván ($x = 3, 4, 5$) tương đương với sự kiện "ván thứ x người A thắng và trong $x-1$ ván đầu người A thắng 2 ván". Khi đó, xác suất cần tìm là

$$p_x = 0,7 \times P_{x-1}(2) = 0,7 \times C_{x-1}^2 \times (0,7)^2 \times (0,3)^{x-3},$$

cụ thể:

$$p_3 = 0,7 \times P_2(2) = 0,7 \times C_2^2 \times (0,7)^2 \times (0,3)^0 = 0,343,$$
$$p_4 = 0,7 \times P_3(2) = 0,7 \times C_3^2 \times (0,7)^2 \times (0,3)^1 = 0,3087,$$
$$p_5 = 0,7 \times P_4(2) = 0,7 \times C_4^2 \times (0,7)^2 \times (0,3)^2 = 0,18522.$$

(b) Sự kiện "trận đấu kết thúc sau 5 ván" tương đương với sự kiện "trong 4 ván đầu mỗi người thắng 2 ván". Khi đó xác suất cần tìm là:

$$P_4(2) = C_4^2 \times (0,7)^2 \times (0,3)^2 = 0,2646.$$

Ví dụ 1.35. Tỷ lệ phế phẩm của một lô hàng là 1%. Hỏi cỡ mẫu cần chọn ra là bao nhiêu (có hoàn lại) sao cho trong mẫu có ít nhất 1 phế phẩm với xác suất lớn hơn 0,95?

Lời giải: Giả sử mẫu chọn ra có kích cỡ là n và việc chọn ra một sản phẩm có hoàn lại là một phép thử Béc–nu–li với $p = 0,01$. Gọi A là sự kiện "trong mẫu có ít nhất một phế phẩm" thì $\overline{A}$ sẽ là sự kiện "trong mẫu không có phế phẩm nào". Khi đó $\overline{A} = A_1 A_2 \ldots A_2$, với A_i là sự kiện "sản phẩm thứ i lấy ra không là phế phẩm", $i = 1, 2, \ldots, n$. Suy ra

$$P(A) = 1 - P(\overline{A}) = 1 - (0,99)^n.$$

Theo yêu cầu của đầu bài, $P(A) > 0,95$ tức là $1 - (0,99)^n > 0,95$ hay $0,05 > (0,99)^n$. Từ đây suy ra

$$n > \frac{\log 0,05}{\log 0,99} \simeq 298.$$

1.5.4 Số có khả năng nhất trong lược đồ Béc–nu–li

Trong lược đồ Béc–nu–li, số x_0 mà tại đó xác suất đạt giá trị lớn nhất gọi là số có khả năng nhất (hay số lần xuất hiện chắc chắn nhất).

- Nếu $np - q \in \mathbb{Z}$ thì có hai số có khả năng nhất $x_0 = np - q$ và $x_0 = np - q + 1$.
- Nếu $np = q \notin \mathbb{Z}$ thì $x_0 = [np - q] + 1$, ở đây $[np - q]$ là phần nguyên của $np - q$.

Ví dụ 1.36. Tỷ lệ mắc một loại bệnh A ở một vùng là 10%. Trong đợt khám bệnh cho vùng đó người ta đã khám 100 người. Tìm số người bị bệnh A có khả năng nhất? Tính xác suất tương ứng.

Lời giải: Bài toán thỏa mãn lược đồ Béc–nu–li với $n = 100$, $p = 0,1$. Theo bài ra ta có $np - q = 100 \times 0,1 - 0,9 = 9,1 \notin \mathbb{Z}$. Vậy số người bị bệnh A có khả năng nhất khi khám 100 người là $[9,1] + 1 = 10$ người và xác suất tương ứng là $P_{100}(10) = C_{100}^{10} \times (0,1)^{10} \times (0,9)^{90} \simeq 0,1319$.

1.5.5 Công thức xấp xỉ

Khi n và k khá lớn thì việc tính toán xác suất theo (1.19) và (1.20) rất cồng kềnh và khó khăn, vì vậy người ta tìm cách tính gần đúng các xác suất đó.

(a) **Xấp xỉ Poa-xông:** Nếu n rất lớn, trong khi p rất nhỏ, xác suất theo công thức (1.19) có thể xấp xỉ bằng

$$\boxed{P_n(k) \simeq \frac{(\lambda)^k}{k!}e^{-\lambda}} \tag{1.21}$$

Xác suất này được tính sẵn trong bảng giá trị hàm khối lượng xác suất Poa–xông (Phụ lục 5) với $\lambda = np$.

(b) **Xấp xỉ chuẩn** (định lý giới hạn địa phương Moa–vrơ–Láp–la–xơ): Nếu n lớn nhưng p không quá bé và quá lớn ta có xấp xỉ

$$\boxed{P_n(k) \simeq \frac{\varphi(x_k)}{\sqrt{npq}}, \quad x_k = \frac{k-np}{\sqrt{npq}}} \tag{1.22}$$

trong đó $\varphi(x) = \dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ là hàm Gao–xơ với các giá trị được tính trong bảng giá trị hàm Gao–xơ (Phụ lục 1) đối với các giá trị x dương. Hàm $\varphi(x)$ là hàm chẵn, tức là $\varphi(-x) = \varphi(x)$. Khi $x > 4$ ta có thể lấy $\varphi(x) \simeq 0$.

(c) **Xấp xỉ cho công thức** (1.20) (định lý giới hạn tích phân Moa–vrơ–Láp–la–xơ): Nếu n lớn nhưng p không quá bé và quá lớn thì xác suất trong (1.20) có thể xấp xỉ bằng

$$\boxed{P_n(k_1;k_2) \simeq \phi(x_2) - \phi(x_1), \quad x_i = \frac{k_i - np}{\sqrt{npq}},\ i = 1,2} \tag{1.23}$$

trong đó

$$\phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt \tag{1.24}$$

là hàm Láp–la–xơ với các giá trị được tính trong bảng giá trị hàm Láp–la–xơ (Phụ lục 2) đối với các giá trị x dương. Hàm $\phi(x)$ là hàm lẻ, tức là $\phi(-x) = -\phi(x)$. Khi $x > 5$ ta có thể lấy $\phi(x) \simeq 0,5$.

Ví dụ 1.37. Xác suất để sản phẩm sau khi sản xuất không được kiểm tra chất lượng bằng 0,2. Tìm xác suất để trong 400 sản phẩm sản xuất ra có:

(a) 80 sản phẩm không được kiểm tra chất lượng;

(b) từ 70 đến 100 sản phẩm không được kiểm tra chất lượng.

Lời giải: Bài toán thỏa mãn lược đồ Béc–nu–li với $n = 400$, $p = 0,2$.

(a) Ta phải tính $P_{400}(80)$ theo công thức Béc–nu–li (1.19):

$$P_{400}(80) = C_{400}^{80} \times (0,2)^{80} \times (0,8)^{320}.$$

Việc tính xác suất theo công thức này khá phức tạp vì $n = 400$ khá lớn, $p = 0,2$ không quá bé hoặc quá lớn. Do đó, ta sẽ tính xấp xỉ theo (1.22):

$$P_{400}(80) \simeq \frac{\varphi(0)}{8} \simeq 0,04986$$

ở đây $\varphi(0) = 0,3989$ được tra từ bảng giá trị hàm Gau-xơ (Phụ lục 1).

(b) Tương tự, thay việc dùng công thức (1.20) ta sử dụng xấp xỉ (1.23):

$$P_{400}(70; 100) \simeq \phi(2,5) - \phi(-1,25) \simeq 0,49379 + 0,39435 = 0,88814,$$

ở đây $\phi(-1,25) = -0,39435$, $\phi(2,5) = 0,49379$ tra từ bảng giá trị hàm Láp–la–xơ (Phụ lục 2).

Ví dụ 1.38. Vận chuyển 4000 chai rượu đến một cửa hàng. Xác suất để mỗi chai rượu bị vỡ trong quá trình vận chuyển là 0,001. Tính xác suất để có 7 chai rượu bị vỡ trong quá trình vận chuyển.

Lời giải: Bài toán thỏa mãn lược đồ Béc–nu–li với $n = 4000$, $p = 0,001$. Ta phải tính $P_{4000}(7)$ theo công thức Béc–nu–li (1.19):

$$P_{4000}(7) = C_{4000}^{7} \times (0,001)^{7} \times (0,999)^{3993}.$$

Vì $n = 4000$ khá lớn, $p = 0,001$ khá bé, nên ta sẽ tính xấp xỉ theo (1.21):

$$P_{4000}(7) \simeq \frac{4^7}{7!}(2,71828)^{-4} \simeq 0,05954.$$

Ta có thể tính trực tiếp hoặc tra bảng giá trị hàm khối lượng Poa-xông (Phụ lục 5).

1.6 Công thức xác suất đầy đủ. Công thức Bay–ét

1.6.1 Công thức xác suất đầy đủ

Định lý 1.2. Giả sử các sự kiện $A_1, A_2, \ldots, A_n$ lập thành một hệ đầy đủ và H là một sự kiện nào đó. Khi đó,

$$\boxed{P(H) = \sum_{i=1}^{n} P(A_i)P(H|A_i)} \tag{1.25}$$

Công thức (1.25) được gọi là công thức xác suất đầy đủ (hay công thức xác suất toàn phần). Công thức này cho phép ta tính xác suất $P(H)$ nếu biết các xác suất $P(A_i)$ và $P(H|A_i)$, $i = 1, 2, \ldots, n$.

Chứng minh. Từ giả thiết ta có $H = HS = H(A_1 + A_2 + \cdots + A_2)$. Sử dụng tính xung khắc của các sự kiện và công thức nhân suy ra

$$\begin{aligned} P(H) &= P(HA_1) + P(HA_2) + \cdots + P)HA_n \\ &= P(A_1)P(H|A_1) + P(A_2)P(H|A_2) + \cdots + P(A_n)P(H|A_n). \end{aligned}$$

1.6.2 Công thức Bay–ét

Định lý 1.3. Giả sử ta có một hệ đầy đủ $A_1, A_2, \ldots, A_n$, sau đó có thêm sự kiện H nào đó. Khi đó xác suất $P(A_k|H)$, $k = 1, 2, \ldots, n$, được xác định bởi:

$$\boxed{P(A_k|H) = \frac{P(A_k)P(H|A_k)}{\sum_{i=1}^{n} P(A_i)P(H|A_i)}, \quad k = 1, 2, \ldots, n} \tag{1.26}$$

Công thức (1.26) được gọi là công thức Bay-ét.

Chứng minh. Sử dụng công thức nhân (1.10) $P(A_kH) = P(A_k)P(H|A_k) = P(H)P(A_k|H)$. Suy ra

$$P(A_k|H) = \frac{P(A_k)P(H|A_k)}{P(H)}. \tag{1.27}$$

Từ đây sử dụng công thức xác suất đầy đủ (1.25) suy ra công thức (1.26).

Nhận xét 1.11. (a) Các xác suất $P(A_i)$, $i = 1, 2, \ldots, n$ đã được xác định từ trước, thường được gọi là xác suất tiên nghiệm.

(b) Các xác suất $P(A_i|H)$, $i = 1, 2, \ldots, n$ được xác định sau khi đã có kết quả thí nghiệm nào đó thể hiện qua sự xuất hiện của H, thường gọi là xác suất hậu nghiệm. Như vậy, công thức Bay–ét cho phép đánh giá lại xác suất xảy ra các sự kiện A_i sau khi đã có thêm thông tin về H.

Chú ý 1.7. (a) Muốn dùng công thức xác suất đầy đủ (1.25) hoặc công thức Bay–ét (1.26) nhất định phải có hệ đầy đủ.

(b) Nếu (1.25) cho ta xác suất không có điều kiện thì (1.26) cho phép tính xác suất có điều kiện, trong đó sự kiện A_i cần tính xác suất phải là một thành viên của nhóm đầy đủ đang xét. Từ đó thấy rằng việc dùng công thức Bay–ét để tính xác suất có điều kiện đã gợi ý cho ta cách chọn nhóm đầy đủ sao cho sự kiện quan tâm phải là thành viên.

(c) Trong trường hợp không có, hoặc rất khó xác định nhóm đầy đủ ta nên dùng công thức (1.26), trong trường hợp này tính $P(H)$ sẽ khó hơn là dùng công thức (1.25).

Ví dụ 1.39. Một nhà máy có ba phân xưởng sản xuất ra cùng một loại sản phẩm. Xác suất để phân xưởng 1, phân xưởng 2 và phân xưởng 3 sản xuất được sản phẩm loại một lần lượt là 0,7, 0,8 và 0,6. Từ một lô hàng gồm 20% sản phẩm của phân xưởng 1, 50% sản phẩm của phân xưởng 2 và 30% sản phẩm của phân xưởng 3 người ta lấy ra một sản phẩm để kiểm tra.

(a) Tính xác suất để sản phẩm được kiểm tra là loại một.

(b) Biết sản phẩm được kiểm tra là loại một. Tính xác suất để sản phẩm này do phân xưởng 2 sản xuất.

Lời giải: Gọi H là sự kiện "sản phẩm được kiểm tra là loại một"; A_i là sự kiện "sản phẩm được kiểm tra do phân xưởng i sản xuất", $i = 1, 2, 3$. Ta thấy A_1, A_2, A_3 tạo thành một hệ đầy đủ với $P(A_1) = 0,2$, $P(A_2) = 0,5$ và $P(A_3) = 0,3$.

(a) Áp dụng công thức xác suất đầy đủ (1.25) với $P(H|A_1) = 0,7$; $P(H|A_2) = 0,8$ và $P(H|A_3) = 0,6$ ta nhận được

$$\begin{aligned} P(H) &= P(A_1)P(H|A_1) + P(A_2)P(H|A_2) + P(A_3)P(H|A_3) \\ &= 0,2 \times 0,7 + 0,5 \times 0,8 + 0,3 \times 0,6 = 0,72 = 72\%. \end{aligned}$$

Ý nghĩa của xác suất này là tỷ lệ sản phẩm loại một của nhà máy.

(b) Áp dụng công thức Bay–ét (1.26) ta tính

$$P(A_2|H) = \frac{P(A_2)P(H|A_2)}{\sum_{i=1}^{3} P(A_i)P(H|A_i)} = \frac{0,5 \times 0,8}{0,72} = \frac{5}{9}.$$

Ví dụ 1.40. Có hai lô sản phẩm: lô I có 7 chính phẩm 3 phế phẩm; lô II có 6 chính phẩm 2 phế phẩm. Lấy ngẫu nhiên 2 sản phẩm từ lô I bỏ sang lô II, sau đó từ lô II lấy ngẫu nhiên ra 2 sản phẩm.

(a) Tính xác suất để 2 sản phẩm lấy ra sau cùng là chính phẩm.

(b) Giả sử 2 sản phẩm lấy ra sau cùng là chính phẩm. Hãy tính xác suất để 2 chính phẩm này là của lô I (ban đầu).

Lời giải:

(a) Gọi H là sự kiện "hai sản phẩm lấy ra sau cùng là chính phẩm"; A_i là sự kiện "trong 2 sản phẩm lấy từ lô I bỏ sang lô II có i chính phẩm", $i = 0, 1, 2$. Khi đó A_0, A_1, A_2 tạo thành một hệ đầy đủ với

$$P(A_0) = \frac{C_3^2}{C_{10}^2} = \frac{1}{15};\ P(A_1) = \frac{C_7^1 \times C_3^1}{C_{10}^2} = \frac{7}{15};\ P(A_2) = \frac{C_7^2}{C_{10}^2} = \frac{7}{15};$$

và
$$P(H|A_0) = \frac{C_6^2}{C_{10}^2} = \frac{15}{45};\ P(H|A_1) = \frac{C_7^2}{C_{10}^2} = \frac{21}{45};\ P(H|A_2) = \frac{C_8^2}{C_{10}^2} = \frac{28}{45}.$$

Áp dụng công thức xác suất đầy đủ (1.25)

$$\begin{aligned} P(H) &= P(A_0)P(H|A_0) + P(A_1)P(H|A_1) + P(A_2)P(H|A_2) \\ &= \frac{1}{15} \times \frac{15}{45} + \frac{7}{15} \times \frac{21}{45} + \frac{7}{15} \times \frac{28}{45} = \frac{358}{675} \simeq 0,5304. \end{aligned}$$

(b) Ta không thể chọn nhóm đầy đủ như trong ý (a), vì sự kiện cần tính xác suất không là thành viên của nhóm này. Việc chọn nhóm đầy đủ thích hợp xem như là bài tập.

Ví dụ 1.41. Một người có ba chỗ ưa thích như nhau để câu cá. Xác suất để câu được cá ở mỗi chỗ tương ứng là 0,6; 0,7 và 0,8. Biết rằng đến một chỗ người đó thả câu 3 lần và chỉ câu được một con cá. Tính xác suất để cá câu được ở chỗ thứ nhất.

Lời giải: Gọi A_i là sự kiện "người đó chọn chỗ thứ i", $i = 1, 2, 3$, A là sự kiện "câu được cá". Khi đó,
$$P(A) = P(A_1)P(A|A_1) + P(A_2)P(A|A_2) + P(A_3)P(A|A_3) = 0,191,$$
trong đó

$$\begin{aligned} &P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}, \\ &P(A|A_1) = P_3(1) = C_3^1 \times (0,6)^1 \times (0,4)^2 = 0,288, \\ &P(A|A_2) = P_3(1) = C_3^1 \times (0,7)^1 \times (0,3)^2 = 0,189, \\ &P(A|A_3) = P_3(1) = C_3^1 \times (0,8)^1 \times (0,2)^2 = 0,096. \end{aligned}$$

Từ đây suy ra
$$P(A_1|A) = \frac{P(A_1)P(A|A_1)}{P(A)} = 0,5026.$$

Ví dụ 1.42. Người ta dùng một thiết bị để kiểm tra một loại sản phẩm nhằm xác định sản phẩm có đạt yêu cầu không. Biết rằng sản phẩm có tỉ lệ phế phẩm là 0,01. Thiết bị có khả năng phát hiện đúng sản phẩm là phế phẩm với xác suất 0,85 và phát hiện đúng sản phẩm đạt chất lượng với xác suất 0,9. Kiểm tra ngẫu nhiên một sản phẩm, tìm xác suất sao cho sản phẩm này:

(a) Được kết luận là phế phẩm.

(b) Được kết luận là đạt chất lượng thì lại là phế phẩm.

(c) Được kết luận đúng với thực chất của nó.

Lời giải: Gọi A là sự kiện "sản phẩm được chọn là phế phẩm", $P(A) = 0,01$, $P(\overline{A}) = 0,99$.

(a) Gọi H là sự kiện "sản phẩm được kết luận là phế phẩm", khi đó $\overline{H}$ là sự kiện "sản phẩm được kết luận là đạt chất lượng". Theo đầu bài, $P(H|A) = 0,85$, $P(\overline{H}|\overline{A}) = 0,9$. Suy ra

$$P(H) = P(A)P(H|A) + P(\overline{A})P(H|\overline{A}) = 0,01 \times 0,85 + 0,99 \times 0,1 = 0,1075.$$

(b) $P(\overline{H}) = 1 - 0,1075 = 0,8925$. Suy ra

$$P(A|\overline{H}) = \frac{P(A\overline{H})}{P(\overline{H})} = \frac{P(A)P(\overline{H}|A)}{P(\overline{H})} = \frac{0,01 \times 0,15}{0,8925} = 0,0017.$$

(c) $P(AH) + P(\overline{A}\,\overline{H}) = P(A)P(H|A) + P(\overline{A})P(\overline{H}|\overline{A}) = 0,01 \times 0,85 + 0,99 \times 0,9 = 0,8995.$

Ví dụ 1.43. Một hãng hàng không cho biết rằng 5% số khách đặt trước vé cho các chuyến đã định sẽ hoãn không đi chuyến bay đó. Do đó hãng đã đưa ra một chính sách là sẽ bán 52 ghế cho một chuyến bay mà trong đó mỗi chuyến chỉ trở được 50 khách hàng. Tìm xác suất để tất cả các khách đặt chỗ trước và không hoãn chuyến bay đều có ghế. Biết rằng xác suất bán được 51 vé hoặc 52 vé là như nhau và bằng 10%.

Lời giải: Gọi A là sự kiện "bán được 52 vé", B là sự kiện "bán được 51 vé", C là sự kiện "bán được ≤ 50 vé". Khi đó A, B, C tạo thành một nhóm đầy đủ, $P(A) = P(B) = 0,1$ và $P(C) = 0,8$.

Gọi H là sự kiện "tất cả các khách hàng đặt chỗ trước và không hoãn chuyến bay đều đủ chỗ", suy ra $\overline{H}$ là sự kiện "khách hàng không đủ chỗ". Khi đó

$$P(\overline{H}) = P(A)P(\overline{H}|A) + P(B)P(\overline{H}|B) + P(C)P(\overline{H}|C),$$

trong đó

$$\begin{aligned} P(\overline{H}|A) &= P_{52}(0) + P_{52}(1) = (0,95)^{52} + 52 \times (0,95)^{51} \times (0,05)^1, \\ P(\overline{H}|B) &= P_{51}(0) = (0,95)^{51}, \\ P(\overline{H}|C) &= 0. \end{aligned}$$

Từ đó $P(\overline{H}) = 0,0333$, suy ra $P(H) = 0,6667$.

Ví dụ 1.44. Ba người thợ cùng may một loại áo với xác suất may được sản phẩm chất lượng cao tương ứng là 0,9; 0,9 và 0,8. Biết một người khi may 8 áo thì có 6 sản phẩm chất lượng cao. Tìm xác suất để người đó may 8 áo nữa thì có 6 áo chất lượng cao.

Lời giải: Gọi A là sự kiện "trong 8 áo đầu tiên có 6 áo chất lượng cao"; A_i là sự kiện "8 áo đầu tiên do người thợ thứ i may", $i = 1,2,3$ với $P(A_i) = \dfrac{1}{3}$, $i = 1,2,3$. Theo công thức xác suất đầy đủ

$$\begin{aligned} P(A) &= P(A_1)P(A|A_1) + P(A_2)P(A|A_2) + P(A_3)P(A|A_3) \\ &= \frac{1}{3} \times \left[C_8^6 \times (0,9)^6 \times (0,1)^2 + C_8^6 \times (0,9)^6 \times (0,1)^2 + C_8^6 \times (0,8)^6 \times (0,2)^2\right] \simeq 0,2. \end{aligned}$$

Gọi B là sự kiện "trong 8 áo sau có 6 áo chất lượng cao".

$$P(B) = \sum_{i=1}^{3} P(A_i|A)P(B|A_iA) = 0,225,$$

trong đó các xác suất $P(A_1|A)$, $P(A_2|A)$, $P(A_3|A)$ được xác định theo công thức Bay-et.