Một sự kiện có thể coi là có xác suất nhỏ tuỳ thuộc vào bài toán cụ thể. Chẳng hạn mỗi chiếc máy bay đều có một xác suất rất nhỏ bị xảy ra tai nạn (như ví dụ trên là 0,000133%). Nhưng trên thực tế ta vẫn không từ chối đi máy bay vì tin rằng trong chuyến bay ta đi sự kiện máy bay rơi không xảy ra.

Hiển nhiên việc quy định một mức xác suất thế nào được gọi là nhỏ sẽ phụ thuộc vào từng bài toán cụ thể. Chẳng hạn nếu xác suất để máy bay rơi là 0,01% thì xác suất đó chưa thể được coi là nhỏ. Song nếu xác suất một chuyến tàu khởi hành chậm là 0,01% thì có thể coi rằng xác suất này là nhỏ.

Mức xác suất nhỏ này được gọi là mức ý nghĩa. Nếu α là mức ý nghĩa thì số $\beta = 1 - \alpha$ gọi là độ tin cậy. Khi dựa trên nguyên lí xác suất nhỏ ta nói rằng: "Sự kiện A có xác suất nhỏ (tức là $P(A) \leq \alpha$) sẽ không xảy ra trên thực tế với độ tin cậy là β". Tính đúng đắn của kết luận sẽ xảy ra trong $100.\beta\%$ trường hợp.

Tương tự như vậy ta có thể đưa ra *Nguyên lí xác suất lớn*: "Nếu sự kiện A có xác suất lớn (gần bằng 1) thì trên thực tế có thể hiểu sự kiện đó sẽ xảy ra trong một (hoặc một vài) phép thử tiếp theo".

1.3. XÁC SUẤT CÓ ĐIỀU KIỆN

Như chúng ta đã biết, xác suất của một sự kiện có thể phụ thuộc vào nhiều yếu tố, điều kiện khác nhau. Để chỉ ra một cách cụ thể hơn về việc xác suất của một sự kiện A nào đó phụ thuộc vào một điều kiện B cho trước ra sao, người ta đưa ra khái niệm xác suất có điều kiện. Điều kiện B cũng có thể hiểu là một sự kiện, tức là sự kiện "có B".

Ví dụ 1.19. Một kênh liên lạc kỹ thuật số có tỷ lệ lỗi là 1 bit cho mỗi 10^6 lượt truyền tin. Lỗi này có xác suất nhỏ, nhưng khi chúng xảy ra, chúng có xu hướng ảnh hưởng đến nhiều bit liên tiếp. Như vậy một bit được truyền đi thì xác suất lỗi là $1/10^6$; tuy nhiên, nếu bit trước đó bị lỗi thì chắc chắn xác suất để lỗi ở bit tiếp theo sẽ lớn hơn $1/10^6$.

1.3.1. Định nghĩa và tính chất

Định nghĩa 1.7. Giả sử điều kiện B có $P(B) > 0$, khi đó xác suất của sự kiện A biết rằng điều kiện B đã xảy ra, kí hiệu là $P(A|B)$, được định nghĩa là

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{1.6}$$

Một hệ quả trực tiếp của định nghĩa xác suất có điều kiện là công thức tích:

$$P(A \cap B) = P(A|B).P(B). \tag{1.7}$$

Tất nhiên, ta cũng có thể coi B là sự kiện, A là điều kiện, và khi đó ta có

$$P(A \cap B) = P(B|A).P(A). \tag{1.8}$$

Để mô tả xác suất có điều kiện bằng tần suất, ta kí hiệu n_A, n_B và n_{AB} lần lượt là số lần xảy ra sự kiện A, B và $A \cap B$ trong loạt n phép thử với n đủ lớn. Theo định nghĩa xác suất cổ điển, ta có

$$P(A \cap B) = \frac{n_{AB}}{n} \quad \text{và} \quad P(B) = \frac{n_B}{n}.$$

Vì điều kiện B đã xảy ra nên số kết quả có thể cho sự kiện A là n_B, trong đó số kết quả thuận lợi cho sự kiện A là n_{AB}. Do đó

$$P(A|B) = \frac{n_{AB}}{n_B} = \frac{n_{AB}/n}{n_B/n} = \frac{P(A \cap B)}{P(B)}.$$

Ví dụ 1.20. Một tổ điều tra dân số vào thăm một gia đình có 2 con.

a) Tính xác suất gia đình có 2 con trai.

b) Gia đình đang nói chuyện thì có 1 cậu con trai ra chào khách. Tính xác suất để gia đình này có 2 con trai.

Lời giải. Với một gia đình có 2 con thì ta có 4 khả năng xảy ra: $\{TT, TG, GT, GG\}$.

a) Gọi A là sự kiện "gia đình này có 2 con trai" thì $P(A) = 1/4$.

b) Gọi B là sự kiện "gia đình này có con trai" thì $P(B) = 3/4$. Sự kiện "cậu con trai ra chào khách" tức là điều kiện B đã xảy ra. Xác suất để gia đình có 2 con trai là

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

Như vậy, khi điều kiện B chưa xảy ra thì xác suất của sự kiện A là $1/4$. Tuy nhiên khi điều kiện B xảy ra thì khả năng xảy ra sự kiện A đã tăng lên là $P(A|B) = 1/3$.

Tính chất:

(i) $0 \leq P(A|B) \leq 1$,

(ii) $P(B|B) = 1$,

(iii) $P(\overline{A}|B) = 1 - P(A|B)$,

(iv) $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B) - P(A_1 \cap A_2|B)$.

Ví dụ 1.21. Một học viên thi 2 môn, xác suất đậu môn thứ nhất là 0,6. Nếu môn thứ nhất đậu thì khả năng học viên đó đậu môn thứ hai là 0,8. Nếu môn thứ nhất không đậu thì khả năng học viên đó đậu môn thứ hai chỉ là 0,6. Tính xác suất trong các trường hợp sau:

a) Học viên đó đậu chỉ một môn.

b) Học viên đó đậu 2 môn.

Lời giải.

a) Gọi A là sự kiện "học viên đó đậu chỉ một môn",

A_i là sự kiện "học viên đó đậu môn thứ $i\,(i = 1, 2)$".

Ta có

$$\begin{aligned} P(A) &= P(A_1\overline{A_2} \cup \overline{A_1}A_2) = P(A_1\overline{A_2}) + P(\overline{A_1}A_2) \\ &= P(A_1).P(\overline{A_2}|A_1) + P(\overline{A_1}).P(A_2|\overline{A_1}) = 0,6.0,2 + 0,4.0,6 = 0,36. \end{aligned}$$

b) Gọi B là sự kiện "học viên đó đậu 2 môn" thì $B = A_1 \cap A_2$. Suy ra

$$P(B) = P(A_1 \cap A_2) = P(A_1).P(A_2|A_1) = 0,6.0,8 = 0,48.$$

Xác suất có điều kiện cho phép chúng ta sử dụng thông tin về khả năng xảy ra của một sự kiện để dự báo xác suất xảy ra một sự kiện khác.

Từ *(1.7)* có thể mở rộng ra công thức xác suất của n sự kiện diễn ra đồng thời

$$P(A_1 \cap \cdots \cap A_n) = P(A_1).P(A_2|A_1).P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap \cdots \cap A_{n-1}). \quad (1.9)$$

Ví dụ 1.22. Một lô hàng có 100 USB trong đó có 90 USB tốt và 10 USB lỗi. Kiểm tra ngẫu nhiên liên tiếp không hoàn lại 5 USB. Nếu có ít nhất 1 USB lỗi trong 5 USB được kiểm tra thì không nhận lô hàng. Tìm xác suất để nhận lô hàng.

Lời giải.

Gọi A_i là sự kiện "USB được kiểm tra thứ i là tốt", $i = \overline{1,5}$,

A là sự kiện "nhận lô hàng".

Khi đó $A = A_1 \cap \cdots \cap A_5$. Theo *(1.9)* ta có

$$\begin{aligned} P(A) &= P(A_1).P(A_2|A_1).P(A_3|A_1 \cap A_2).P(A_4|A_1 \cap A_2 \cap A_3).P(A_5|A_1 \cap \cdots \cap A_4) \\ &= \frac{90}{100} \cdot \frac{89}{99} \cdot \frac{88}{98} \cdot \frac{87}{97} \cdot \frac{86}{96}. \end{aligned}$$

1.3.2. Sự độc lập và phụ thuộc của các sự kiện

Hai sự kiện A và B được gọi là độc lập với nhau khi việc có hay không xảy ra sự kiện B không ảnh hưởng gì đến việc có hay không xảy ra sự kiện A. Nói cách khác, xác suất của A với điều kiện B bằng với xác suất của A khi không tính đến điều kiện B.

Định nghĩa 1.8. Ta nói rằng A và B là 2 sự kiện độc lập nếu

$$P(A|B) = P(A) \quad \text{hoặc} \quad P(B|A) = P(B) \qquad (1.10)$$

hay viết cách khác

$$P(A \cap B) = P(A).P(B).$$

Việc kiểm tra biểu thức *(1.10)* trong thực tiễn rất khó khăn và trong nhiều trường hợp là không thể. Vì vậy dựa vào thực tế và trực giác mà ta thừa nhận các sự kiện độc lập trong các bài tập sau này.

Một cách tổng quát, giả sử một họ $\mathcal{E}$ (hữu hạn hoặc vô hạn) các sự kiện. Khi đó:

Định nghĩa 1.9. Họ $\mathcal{E}$ được gọi là một họ các sự kiện độc lập, nếu như với bất kì k sự kiện $A_1, A_2 \cdots, A_k$ khác nhau nào trong họ $\mathcal{E}$ ta cũng có

$$P\Big(\bigcap_{i=1}^{k} A_i\Big) = \prod_{i=1}^{k} P(A_i). \qquad (1.11)$$

Nếu $P(A \cap B) = P(A).P(B)$ với bất kì hai sự kiện khác nhau nào trong họ $\mathcal{E}$ thì họ $\mathcal{E}$ được gọi là họ các sự kiện độc lập từng đôi một.

Nếu như hai sự kiện không độc lập với nhau, thì người ta nói là chúng phụ thuộc nhau. Do tính chất đối xứng, nếu sự kiện A phụ thuộc vào sự kiện B thì B

cũng phụ thuộc vào A. Nếu như $P(A|B) > P(A)$ thì ta có thể nói là điều kiện B thuận lợi cho sự kiện A, và ngược lại nếu $P(A|B) < P(A)$ thì điều kiện B không thuận lợi cho sự kiện A.

Công thức $P(A|B).P(B) = P(B|A).P(A)$ tương đương với $\dfrac{P(A|B)}{P(A)} = \dfrac{P(B|A)}{P(B)}$ có thể được hiểu như sau: B thuận lợi cho A (tức là $P(A|B) > P(A)$) thì A cũng thuận lợi cho B và ngược lại.

1.4. CÔNG THỨC BAYES

1.4.1. Công thức xác suất đầy đủ

Định nghĩa 1.10. Nhóm các sự kiện $A_1, A_2, \cdots, A_n\,(n \geq 2)$ của một phép thử được gọi là một nhóm đầy đủ nếu

(i) $A_i \cap A_j = \emptyset$, $\forall i \neq j$ (xung khắc từng đôi),

(ii) $A_1 \cup A_2 \cup \cdots \cup A_n = \mathbf{\Omega}$.

Theo định nghĩa này ở phép thử đang xét chỉ có thể xuất hiện một sự kiện trong số n sự kiện $A_1, \cdots, A_n$ (và phải có một sự kiện). Đặc biệt với mọi sự kiện A thì hệ $\{A, \overline{A}\}$ là đầy đủ.

Ví dụ 1.23. Một tiểu đoàn có 3 đại đội cùng trồng một loại bí xanh. Chọn ngẫu nhiên một quả bí xanh và gọi A_1, A_2, A_3 lần lượt là sự kiện quả bí xanh được chọn do đại đội 1, đại đội 2 và đại đội 3 trồng. Khi đó hệ $\{A_1, A_2, A_3\}$ là đầy đủ.

Nếu như ta chưa biết xác suất $P(A)$ của một sự kiện A nào đó, nhưng biết các xát suất $P(A_i)$ của một hệ đầy đủ $\{A_1, A_2, \cdots, A_n\}$ của không gian mẫu $\mathbf{\Omega}$, và biết các xác suất có điều kiện $P(A|A_i)$, thì ta có thể dùng công thức sau, gọi là *công thức xác suất đầy đủ*, để tính xác suất của sự kiện A

$$P(A) = \sum_{i=1}^{n} P(A \cap A_i) = \sum_{i=1}^{n} P(A_i).P(A|A_i). \tag{1.12}$$

Khi A xảy ra thì có một và chỉ một biến cố A_i cùng xảy ra với A. Trường hợp riêng của *(1.12)* là khi ta có hai sự kiện A, B, có thể sử dụng hệ đầy đủ $\{B, \overline{B}\}$ của $\mathbf{\Omega}$ để tính xác suất của A:

$$P(A) = P(A \cap B) + P(A \cap \overline{B}) = P(B).P(A|B) + P(\overline{B}).P(A|\overline{B}). \tag{1.13}$$

Chú ý: Vận dụng công thức xác suất đầy đủ để giải một bài toán, vấn đề quan trọng là phải chỉ ra được nhóm sự kiện đầy đủ và xung khắc từng đôi. Trong thực tế việc này thường gặp ở 2 hình thức sau:

(i) Công việc tiến hành trải qua 2 phép thử. Thực hiện phép thử thứ nhất ta có một trong n khả năng xảy ra là các sự kiện: $A_1, A_2, \cdots, A_n$. Sau khi thực hiện phép thử thứ nhất ta thực hiện phép thử thứ hai. Trong phép thử thứ hai ta quan tâm đến sự kiện A. Khi đó sự kiện A sẽ được tính theo công thức xác suất đầy đủ với hệ đầy đủ $\{A_1, A_2, \cdots, A_n\}$.

(ii) Một tập hợp chứa n nhóm phần tử. Mỗi nhóm phần tử có một tỉ lệ phần tử có tính chất **P** nào đó. Lấy ngẫu nhiên từ tập hợp ra 1 phần tử. Gọi A là sự kiện chọn được phần tử thuộc nhóm thứ i. Khi đó xác suất của sự kiện chọn được

phần tử có tính chất **P** trong phép thử sẽ được tính theo công thức xác suất đầy đủ với hệ đầy đủ $\{A_1, A_2, \cdots, A_n\}$.

Ví dụ 1.24. Xét một lô giày chiến sĩ được sản xuất bởi 3 nhà máy với tỉ lệ lần lượt là 20%, 30% và 50%. Xác suất giày hỏng của các nhà máy lần lượt là 0,001; 0,005 và 0,006. Lấy ngẫu nhiên một chiếc giày từ lô hàng. Tìm xác suất để chiếc giày lấy ra bị hỏng.

Lời giải. Gọi A là sự kiện "lấy được giày hỏng",
A_i là sự kiện "lấy được giày của nhà máy i" ($i = 1, 2, 3$).
Ta có $\{A_1, A_2, A_3\}$ là hệ đầy đủ. Theo công thức *(1.12)*, ta có

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

1.4.2. Công thức Bayes

Công thức Bayes, mang tên của linh mục và nhà toán học người Anh Thomas Bayes (1702-1761), là công thức ngược, cho phép tính xác suất có điều kiện $P(B|A)$ khi biết xác suất có điều kiện $P(A|B)$ và một số thông tin khác. Dạng đơn giản nhất của công thức này là: Nếu A, B là hai sự kiện bất kì với xác suất khác 0 thì từ *(1.7)* và *(1.8)*, ta có:

$$P(B|A) = \frac{P(A|B).P(B)}{P(A)}. \tag{1.14}$$

Kết hợp *(1.14)* với công thức xác suất đầy đủ *(1.12)* cho $P(A)$, ta được

Định lý 1.1. *Giả sử $\{A_1, A_2, \cdots, A_n\}$ là hệ đầy đủ và A là một sự kiện bất kì có thể xảy ra trong phép thử. Khi đó ta có công thức Bayes:*

$$P(A_k|A) = \frac{P(A_k).P(A|A_k)}{P(A)} = \frac{P(A_k).P(A|A_k)}{\sum_{i=1}^{n} P(A_i).P(A|A_i)} \tag{1.15}$$

với mọi $k = 1, 2, \cdots, n$.

Ví dụ 1.25. Mạng Bayes được sử dụng trên các trang Web của các nhà sản xuất công nghệ cao nhằm giúp khách hàng nhanh chóng chẩn đoán các vấn đề với sản phẩm. Chẳng hạn, một doanh nghiệp sản xuất máy in thu được các kết quả từ cơ sở phân tích dữ liệu khảo sát của khách hàng. Trong đó, các lỗi xuất hiện ở máy in được chia làm ba nhóm: phần cứng, phần mềm và các loại khác (chẳng hạn như trình kết nối) với xác suất tương ứng là 0,1; 0,6 và 0,3. Xác suất lỗi do phần cứng gây ra là 0,9; do phần mềm là 0,2 và do các vấn đề khác là 0,5. Nếu khách hàng truy cập trang web của nhà sản xuất để chẩn đoán lỗi máy in, nguyên nhân nào có khả năng cao nhất?

Lời giải.

Gọi A là sự kiện "máy in bị lỗi", B là sự kiện "lỗi thuộc phần cứng", C là sự kiện "lỗi thuộc phần mềm" và D là sự kiện "lỗi thuộc các vấn đề khác".

Ta có $\{B, C, D\}$ là hệ đầy đủ. Áp dụng *(1.15)* ta có:

Xác suất để lỗi của máy in thuộc nhóm phần cứng là

$$P(B|A) = \frac{P(B).P(A|B)}{P(B).P(A|B) + P(C).P(A|C) + P(D).P(A|D)}$$
$$= \frac{0,1.0,9}{0,1.0,9 + 0,6.0,2 + 0,3.0,5} = \frac{1}{4}.$$

Xác suất để lỗi của máy in thuộc nhóm phần mềm là

$$P(C|A) = \frac{P(C).P(A|C)}{P(B).P(A|B) + P(C).P(A|C) + P(D).P(A|D)}$$
$$= \frac{0,6.0,2}{0,1.0,9 + 0,6.0,2 + 0,3.0,5} = \frac{1}{3}.$$

Xác suất để lỗi của máy in thuộc nhóm các loại khác là

$$P(D|A) = \frac{P(D).P(A|D)}{P(B).P(A|B) + P(C).P(A|C) + P(D).P(A|D)}$$
$$= \frac{0,3.0,5}{0,1.0,9 + 0,6.0,2 + 0,3.0,5} = \frac{5}{12}.$$

Vậy lỗi máy có nguyên nhân từ nhóm các loại khác có khả năng cao nhất.

Ví dụ 1.26. Một trạm chỉ phát hai tín hiệu A và B với xác suất tương ứng 0,85 và 0,15. Do có nhiễu trên đường truyền nên 1/7 tín hiệu A bị méo và thu được như tín hiệu B còn 1/8 tín hiệu B bị méo và thu được như tín hiệu A.

a) Tìm xác suất để máy thu được tín hiệu A.

b) Giả sử máy đã thu được tín hiệu A. Hỏi xác suất thu được đúng tín hiệu phát đi?

Lời giải.

Gọi A là sự kiện "phát tín hiệu A", B là sự kiện "phát tín hiệu B" thì $\{A, B\}$ là hệ đầy đủ.

Gọi T_A là sự kiện "thu được tín hiệu A", T_B là sự kiện "thu được tín hiệu B".

Theo đề bài ta có:

$$P(A) = 0,85; \quad P(B) = 0,15; \quad P(T_B|A) = 1/7; \quad P(T_A|B) = 1/8.$$

a) Áp dụng công thức xác suất đầy đủ, máy thu được tín hiệu A với xác suất

$$P(T_A) = P(A).P(T_A|A) + P(B).P(T_A|B) = 0,85.\frac{6}{7} + 0,15.\frac{1}{8} \approx 0,7473.$$

b) Áp dụng công thức Bayes ta có:

$$P(A|T_A) = \frac{P(A).P(T_A|A)}{P(T_A)} = \frac{0,85.\frac{6}{7}}{0,7473} \approx 0,975.$$

Nhận xét: Công thức Bayes cho phép đánh giá lại xác suất xảy ra các sự kiện A_k, sau khi đã có thêm thông tin về sự kiện A. Cần phải nhấn mạnh rằng nếu muốn dùng các công thức *(1.12)* hoặc *(1.15)*, nhất thiết phải có hệ đầy đủ. Ngoài ra nếu *(1.12)* cho ta xác suất không có điều kiện, thì *(1.15)* cho phép tính xác suất có điều kiện, trong đó sự kiện A_k cần tính xác suất phải là một thành viên của hệ đầy đủ đang xét. Từ đó thấy rằng việc dùng công thức Bayes để tính xác suất có điều kiện đã gợi ý cho ta cách chọn hệ đầy đủ sao cho sự kiện quan tâm phải là thành viên. Trong trường hợp không có (hoặc rất khó xác định) hệ đầy đủ, ta nên dùng công thức *(1.14)*, khi đó việc tính $P(A)$ sẽ khó hơn là dùng công thức *(1.12)*.

Ví dụ 1.27. Một mạch điện gồm 2 bộ phận mắc nối tiếp, với xác suất làm việc tốt trong một khoảng thời gian nào đó của mỗi bộ phận là 0,95 và 0,98. Ở một thời điểm trong khoảng thời gian trên, người ta thấy mạch điện ngừng làm việc (do bộ phận nào đó hỏng), tìm xác suất để chỉ bộ phận thứ hai hỏng.

Lời giải.
Do 2 bộ phận mắc nối tiếp nên chỉ 1 bộ phận hỏng là mạch ngừng làm việc.
Gọi A_i là sự kiện "bộ phận thứ i tốt" ($i = 1, 2$), khi đó có 4 khả năng khác nhau:
- B_0 là sự kiện "cả 2 bộ phận đều tốt";
- B_1 là sự kiện "bộ phận I tốt, bộ phận II hỏng";
- B_2 là sự kiện "bộ phận I hỏng, bộ phận II tốt";
- B_3 là sự kiện "cả 2 bộ phận đều hỏng".
Ta có hệ $\{B_0, B_1, B_2, B_3\}$ là hệ đầy đủ. Do tính độc lập của các thiết bị nên

$$\begin{aligned}
P(B_0) &= P(A_1 \cap A_2) = P(A_1).P(A_2) = 0,95.0,98 = 0,931,\\
P(B_1) &= P(A_1 \cap \overline{A_2}) = P(A_1).P(\overline{A_2}) = 0,95.0,02 = 0,019,\\
P(B_2) &= P(\overline{A_1} \cap A_2) = P(\overline{A_1}).P(A_2) = 0,05.0,98 = 0,049,\\
P(B_3) &= P(\overline{A_1} \cap \overline{A_2}) = P(\overline{A_1}).P(\overline{A_2}) = 0,05.0,02 = 0,001.
\end{aligned}$$

Gọi A là sự kiện "mạch không làm việc" thì

$$P(A|B_0) = 0, \quad P(A|B_1) = P(A|B_2) = P(A|B_3) = 1.$$

Áp dụng công thức *(1.15)* ta có xác suất cần tìm là:

$$P(B_1|A) = \frac{P(B_1).P(A|B_1)}{\sum_{i=0}^{3} P(B_i).P(A|B_i)} = \frac{0,019}{0,019 + 0,049 + 0,001} = \frac{19}{69}.$$

Hoặc ta có thể dùng *(1.14)* để tính $P(B_1|A)$. Ta có

$$A = (A_1 \cap \overline{A_2}) \cup (\overline{A_1} \cap A_2) \cup (\overline{A_1} \cap \overline{A_2}).$$

Do tính xung khắc và độc lập của các sự kiện tương ứng nên

$$P(A) = P(A_1).P(\overline{A_2}) + P(\overline{A_1}).P(A_2) + P(\overline{A_1}).P(\overline{A_2}) = 0,069.$$

Mặt khác $B_1 \cap A = A_1 \cap \overline{A_2}$ nên $P(B_1 \cap A) = 0,019$, từ đó ta suy ra kết quả cần tìm mà không cần đến hệ đầy đủ. Tuy nhiên mọi khó khăn rơi vào việc tính $P(A)$.