



Xác suất có điều kiện

PiMA 2021: The Mathematics of Data Science

Nguyễn Thị Minh Thư

Ngày 27 tháng 7 năm 2021



1. Xác suất có điều kiện
2. Độc lập
3. Xác suất đầy đủ
4. Định lý Bayes

Xác suất có điều kiện



Xác suất có điều kiện cho chúng ta cách lập luận về kết quả của một thử nghiệm, dựa trên một vài thông tin.

Ví dụ: Tung 2 viên xúc xắc cân đối lần lượt, biết rằng số chấm trên viên xúc xắc đầu tiên là 3. Vậy xác suất để tổng số chấm trên 2 viên xúc xắc là 8 là bao nhiêu?



Cho 2 biến cố A và B . Xác suất có điều kiện (Conditional probability) là xác suất của biến cố A , biết rằng B xảy ra. Kí hiệu là $P(A|B)$.



Tung 2 viên xúc xắc cân đối lần lượt, biết rằng số chấm trên viên xúc xắc đầu tiên là 3. Vậy xác suất để tổng số chấm trên 2 viên xúc xắc là 8 là bao nhiêu?



- Không gian mẫu: $\Omega = \{(x, y) \mid x, y = 1, 2, 3, 4, 5, 6\}$
- $A = \{\text{số chấm trên viên xúc xắc đầu tiên là } 3\}$, $B = \{\text{tổng số chấm của 2 viên xúc xắc là } 8\}$
- Nếu A xảy ra, có thể có 6 kết quả xảy ra cho viên xúc xắc còn lại $(3, y)$, $y = 1, 2, \dots, 6$
- Xác suất để B xảy ra dưới điều kiện A là

$$P(B \mid A) = P((3, 5) \mid A) = 1/6$$



Chọn ngẫu nhiên một người trong trại hè.

	Trại sinh	Ban tổ chức	Tổng cộng
Nam	20	12	28
Nữ	12	10	26
Tổng cộng	32	22	

Tính:

- Xác suất chọn ra một bạn trại sinh nam trong những người tham gia trại.
- Xác suất chọn ra một bạn trại sinh.
- Xác suất để chọn ra một bạn nam trong số các bạn trại sinh.



	Trại sinh	Ban tổ chức	Tổng cộng
Nam	20	12	28
Nữ	12	10	26
Tổng cộng	32	22	

Đặt M: bạn nam được chọn, S: bạn được chọn là trại sinh

Xác suất chọn ra một bạn trại sinh nam trong những người tham gia trại là

$$P(M \cap S) = \frac{20}{54} = \frac{10}{27}$$

Xác suất chọn ra một bạn trại sinh là

$$P(S) = \frac{32}{54} = \frac{16}{27}$$



	Trại sinh	Ban tổ chức	Tổng cộng
Nam	20	12	28
Nữ	12	10	26
Tổng cộng	32	22	

Xác suất để chọn được một bạn nam trong tổng số các trại sinh là

$$P(M | S) = \frac{20}{32} = 62.5\%$$

Nhận xét là

$$P(M | S) = \frac{20}{32} = \frac{10/27}{16/27} = \frac{P(M \cap S)}{P(S)}$$



Xác suất của A trong trường hợp xảy ra B được kí hiệu là $P(A|B)$

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



1. Phần bù

$$P(B^c | A) = 1 - P(B | A)$$

2. Kết hợp

$$P(B \cup C | A) = P(B | A) + P(C | A) - P(B \cap C | A)$$



	Trại sinh	Ban tổ chức	Tổng cộng
Nam	20	12	28
Nữ	12	10	26
Tổng cộng	32	22	

Xác suất để chọn ra một bạn nữ trong số các bạn trại sinh là

$$P(M^c | S) = 1 - P(M | S) = 1 - \frac{20}{32} = \frac{12}{32} = 37.5\%$$



	Trại sinh	Ban tổ chức	Tổng cộng
Nam	20	12	28
Nữ	12	10	26
Tổng cộng	32	22	

Xác suất để chọn ra một bạn nam hoặc nữ trong số các bạn trại sinh là

$$\begin{aligned}
 P(M \cup M^c | S) &= P(M | S) + P(M^c | S) - P(M \cap M^c | S) \\
 &= 67.5\% + 37.5\% - 0 = 1
 \end{aligned}$$

Độc lập



Hai biến cố A và B được gọi là độc lập nếu B xảy ra không liên quan đến A .

A và B độc lập khi và chỉ khi một trong hai đẳng thức sau xảy ra:

$$P(A|B) = P(A)$$

hoặc

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



Nếu A độc lập với B , thì A cũng độc lập với B^c .

Chứng minh tính chất trên:

Cho A và B là 2 biến cố độc lập.

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

$$\begin{aligned} P(A \cap B^c) &= P(A) - P(A \cap B) \\ &= P(A)[1 - P(B)] = P(A)P(B^c) \end{aligned}$$



Cho A_1, A_2, \dots, A_k là các biến cố độc lập

$$P(A_1 A_2 \dots A_k) = P(A_1) P(A_2) \dots P(A_k)$$



Tung một viên xúc xắc cân đối 2 lần. Tính xác suất để số chấm ở lần tung đầu tiên và thứ hai lần lượt là 1 và 2.

$$P((1,2)) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$



Làm sao để tính xác suất của $A \cap B$ khi A và B phụ thuộc?

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

Ví dụ: Tổng số trại sinh của PiMA là 35 người, trong đó có 10 bạn nữ. Bốc thăm ngẫu nhiên lần lượt tên của 2 bạn. Xác suất để hai bạn được chọn đều là nữ là bao nhiêu?



Xác suất bạn thứ nhất được bốc trúng là nữ là:

$$\frac{10}{35} = \frac{2}{7}$$

Xác suất bạn thứ hai được bốc trúng là nữ biết bạn thứ nhất đã là nữ là:

$$\frac{10-1}{35-1} = \frac{9}{34}$$

Vậy xác suất để cả hai bạn được bốc thăm trúng đều là nữ là:

$$\frac{2}{7} \times \frac{9}{34} = \frac{9}{119}$$



Cho một chuỗi các biến cố A_1, A_2, \dots, A_k

$$P(A_1 A_2 \dots A_k) = P(A_1) \times P(A_2 | A_1) \times P(A_3 | A_1 A_2) \times \dots \\ \times P(A_k | A_1 A_2 \dots A_{k-1})$$

Ví dụ: Cho một hộp gồm 8 quả bóng đỏ và 4 quả bóng xanh. Chọn ngẫu nhiên 3 quả bóng. Xác suất để 3 quả bóng đều màu đỏ là bao nhiêu?



Đặt A_i là quả bóng thứ i có màu đỏ, $i = 1, 2, 3$

- $P(A_1) = \frac{8}{12}$
- $P(A_2 | A_1) = \frac{7}{11}$
- $P(A_3 | A_1 A_2) = \frac{6}{10}$

Vậy xác suất để 3 quả bóng được chọn có màu đỏ là

$$P(A_1 A_2 A_3) = P(A_1) \times P(A_2 | A_1) \times P(A_3 | A_1 A_2) = \frac{14}{55}$$

Xác suất đầy đủ



A_1, A_2, \dots, A_n được gọi là một hệ biến cố đầy đủ của Ω nếu:

- Các biến cố này là biến cố xung khắc: $A_i A_j = \emptyset$ với $i \neq j$
- $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$



- Cho không gian mẫu Ω , và A_1, A_2, \dots, A_n, B là các biến cố
- A_1, A_2, \dots, A_n là một hệ biến cố đầy đủ.
- Biết giá trị $P(B | A_i)$ với $i = \overline{1, n}$

Xác suất của B hay $P(B)$ có thể được tính bằng:

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



Một bạn trại sinh PiMA rất thích chơi cờ, lúc tham gia trại bạn rủ các bạn trong trại cùng chơi một ván cờ với mình. Biết khi chơi với một nửa các bạn (nhóm 1 - N_1) trong trại xác suất để bạn trại sinh đó thắng là 0.3, khi chơi với 1/4 các bạn trong trại (N_2) xác suất để bạn đó chiến thắng là 0.4, và khi chơi với các bạn còn lại (N_3) xác suất để thắng là 0.5. Biết các đối thủ trong một ván cờ được chọn ngẫu nhiên. Tính xác suất để bạn đó thắng một ván cờ.



Đặt $P(W)$ là xác suất để bạn trải sinh thắng một ván cờ.

Ta có:

$$P(W | N_1) = \frac{P(WN_1)}{P(N_1)} = 0.3$$

$$\Rightarrow P(WN_1) = P(W | N_1)P(N_1) = 0.5 \times 0.3 = 0.15$$

$$P(W | N_2) = \frac{P(WN_2)}{P(N_2)} = 0.4$$

$$\Rightarrow P(WN_2) = P(W | N_2)P(N_2) = 0.25 \times 0.4 = 0.1$$

$$P(W | N_3) = \frac{P(WN_3)}{P(N_3)} = 0.5$$

$$\Rightarrow P(WN_3) = P(W | N_3)P(N_3) = 0.25 \times 0.5 = 0.125$$



Vậy xác suất để bạn trại sinh đó thắng một ván cờ là

$$\begin{aligned}P(W) &= P(WN_1) + P(WN_2) + P(WN_3) \\ &= 0.15 + 0.1 + 0.125 = 0.375\end{aligned}$$

Định lý Bayes



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



- Biến cố A là tác nhân ảnh hưởng đến biến cố B .
- Ta chỉ có thể quan sát được biến cố B nhưng không quan sát được biến cố A .

Vấn đề: Giả sử quan sát được B xảy ra thì A có xảy ra không? Biết:

- $P(A)$: xác suất tiên nghiệm của A .
- $P(B | A)$: xác suất hậu nghiệm của B xảy ra nếu A xảy ra.
- $P(B | A^c)$: xác suất hậu nghiệm của B xảy ra nếu A^c xảy ra.

$$\Rightarrow P(A | B) = \frac{P(B | A)P(A)}{P(B | A)P(A) + P(B | A^c)P(A^c)}$$



Ví dụ:

- A : có bệnh nhưng khó quan sát A .
- B : kết quả xét nghiệm dương tính và có thể quan sát được.

Biết:

- $P(B | A)$: xác định đúng bệnh
- $P(B | A^c)$: dương tính giả

\Rightarrow Có thể tính được $P(A | B)$, xác suất bị bệnh khi đã có kết quả dương tính.



Tổng quát

- A_1, \dots, A_k là một hệ biến cố đầy đủ
- Xác suất tiên nghiệm $P(A_i)$
- Biết giá trị $P(B | A_i)$
- Nếu biến cố B xảy ra, $P(A_i | B)$ được tính bằng công thức

$$P(A_i | B) = \frac{P(B | A_i)P(A_i)}{\sum_k P(B | A_k)P(A_k)}$$



Xét nghiệm Covid-19 có hai kết quả âm tính và dương tính. Biết 95% người nhiễm Covid-19 cho kết quả dương tính. Nhưng 2% số người không mắc bệnh cũng cho kết quả dương tính. Cho rằng 10% dân số toàn cầu bị nhiễm bệnh. Tính xác suất chọn ngẫu nhiên một người mắc bệnh, biết kết quả xét nghiệm của người đó là dương tính.



Tung một cục xí ngầu cân đối 2 lần. Đặt X và Y lần lượt là số chấm trên xí ngầu của lần tung thứ nhất và thứ hai. Tính $P(A|B)$.
Biết $A = \{\max(X, Y) = 5\}$ và $B = \{\min(X, Y) = 3\}$



Dimitri P. Bertsekas and John N. Tsitsiklis.

Introduction to Probability - 2nd Edition.

Athena Scientific, 2008.



Sheldon Ross.

A first course in probability - 8th Edition.

Pearson, 2009.



Sheldon M. Ross.

Introduction to Probability and Statistics for Engineers and Scientists - 3rd Edition.

Elsevier Academic Press, 2004.