XSTK A

Lê Minh Tuấn, Ph.D. [1]

[1] Department of Mathematics and Applications,
Saigon University.

Xác suất Thống kê

Nội dung chính của XSTK:

Tập hợp

• Tập hợp (gọi tắt là tập) là một khái niệm cơ bản của toán học, không được định nghĩa.

• Cho tập hợp A và phần tử x.
Nếu x có mặt trong tập A ta nói x là một phần tử của tập A hay x thuộc A, kí hiệu $x \in A$ hoặc $A \ni x$.
Nếu x không có mặt trong tập A ta nói x không thuộc A, kí hiệu $x \notin A$ hoặc $A \not\ni x$.

• Ta thường dừng các chữ cái in hoa như $A, B, C, \ldots$ để kí hiệu cho tập hợp và các chữ cái in thường, $a, b, c \ldots$ để kí hiệu các phần tử của tập hợp.

Tập hợp rỗng

Tập hợp rỗng là tập hợp không chứa bất kỳ phần tử nào. Tập hợp rỗng được kí hiệu là $\varnothing$.

Cách xác định tập hợp

• Liệt kê rõ ràng tất cả các phần tử
Ví dụ. $A = \{9, 10, 11, 12\}$

• Chỉ ra tính chất đặc trưng (trưng tính) của tất cả các phần tử trong tập hợp .
Ví dụ. $B = \{n \in \mathbb{N} \mid 9 \leq n \leq 12\}$

Một số tập hợp đặc biệt

• Tập hợp các số tự nhiên $\mathbb{N} := \{1, 2, 3, \cdots\}$,
• Tập hợp các số nguyên $\mathbb{Z} := \{\cdots, -2, -1, 0, 1, 2, \cdots\}$,
• Tập hợp các số hữu tỉ $\mathbb{Q} := \left\{\frac{m}{n} : m, n \in \mathbb{Z}; n \neq 0\right\}$,
Tập hợp các số hữu tỉ dương $\mathbb{Q}^{+} := \left\{\frac{m}{n} \middle| m, n \in \mathbb{N}; n \neq 0\right\}$,
• Tập hợp các số thực $\mathbb{R}$,
• Tập hợp các số phức $\mathbb{C}$.

Ví dụ 1.1.

Liệt kê các phần tử của các tập hợp sau:

1. $A = \{n \in \mathbb{N} \mid n < 5\}$.
2. B: tập hợp các số nguyên lớn hơn 0 và bé hơn 5.
3. $C = \{x \in \mathbb{R} \mid (x-1)(x+2) = 0\}$.

Giải.

1. $A = \{1;2;3;4\}$.
2. $B = \{1;2;3;4\}$.
3. Ta có $(x-1)(x+2) = 0 \Leftrightarrow \left[\begin{array}{lll} x & = & 1 \\ x & = & -2. \end{array}\right.$

   Mà $x \in \mathbb{R}$ nên $C = \{-2;1\}$.

Tập con. Hai tập hợp bằng nhau

• Tập A là một tập con của tập B khi và chỉ khi mỗi phần tử của A cũng là một phần tử của B. Ta dùng ký hiệu $A \subset B$ để chỉ A là một tập con của B.
Với kí hiệu đó, ta có:
$$A \subset B \quad \Leftrightarrow \quad \forall x, (x \in A \Rightarrow x \in B)$$

• Hai tập hợp A và B được gọi là bằng nhau, và được viết là $A = B$, nếu chúng cùng chứa tất cả các phần tử giống nhau, i.e. $\forall x, (x \in A \Leftrightarrow x \in B)$.
Với định nghĩa đó, ta có:
$$A = B \quad \Leftrightarrow \quad (A \subset B \text{ và } B \subset A)$$

Tính chất

1. $\varnothing \subset A$, với mọi tập A.
2. $A \subset A$, với mọi tập A
3. $(A \subset B) \wedge (B \subset C) \implies (A \subset C)$.

Bài tập 1.1.

Cho biết x là một phần tử của tập hợp A, xác định tính đúng sai của các mệnh đề sau:

a) $x \in A$. | b) $\{x\} \in A$. | c) $x \subset A$. | d) $\{x\} \subset A$.

Giải.

Bài tập 1.2.

Cho tập hợp $A = \{1;2;3;4;5;6\}$. Tìm tất cả các tập con có 3 phần tử của tập hợp A sao cho tổng các phần tử này là một số lẻ.

Giải.

Bài tập 1.3.

Xác định tất cả các tập hợp con của mỗi tập hợp:

a) $A = \{1;2;3\}$

b) $B = \{x;y\}$.

Giải.

Hợp của hai tập hợp

Hợp của tập hợp A và tập hợp B là tập hợp chứa tất cả các phần tử nằm trong A hoặc nằm trong B.

$$A \cup B := \left\{ x \middle| (x \in A) \vee (x \in B) \right\}$$

Giao của hai tập hợp

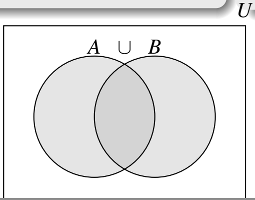
Tập hợp gồm tất cả các phần tử vừa thuộc tập hợp A, vừa thuộc tập hợp B được gọi là giao của A và B.

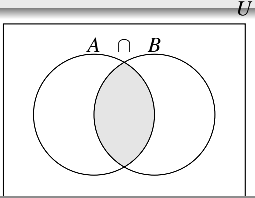
$$A \cap B = \left\{ x \middle| x \in A \text{ và } x \in B \right\}.$$

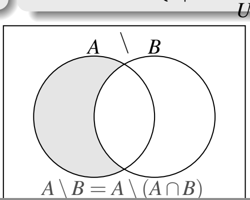
Phần bù của B trong A

Phần bù của tập hợp B trong tập A là tập hợp chứa tất cả các phần tử nằm trong A nhưng không nằm trong B.

$$A \setminus B := \left\{ x \middle| x \in A \text{ và } x \notin B \right\}$$





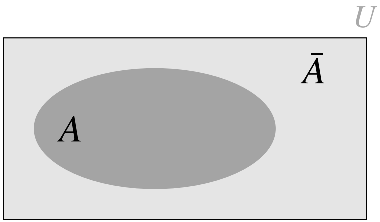


Hai tập rời nhau

Hai tập hợp A và B được gọi là rời nhau nếu chúng không có phần tử chung, nghĩa là $A \cap B = \varnothing$.

Phần bù

Cho $A \subset U$. Phần bù của A trong U là:
$\mathscr{C}_U A = A^c = \bar{A} = U \setminus A$.



Chú ý.

Tập hợp không chứa bất kỳ phần tử nào được gọi là tập rỗng (tập trống), và được ký hiệu là $\varnothing$.

Tập hợp "rất lớn" chứa các tập hợp khác (tùy ngữ cảnh) được gọi là **tập phổ dụng** (tập vũ trụ), và thường được ký hiệu là U.

Phần bù của A (trong tập phổ dụng U) là tập hợp tất cả các phần tử không nằm trong A:
$A^c := \left\{ x \in U \middle| x \notin A \right\}$.

Định Lý 1.2.

Cho A,B,C là các tập hợp và U là tập phổ dụng (i.e. các tập A,B,C đều là các tập con của U). Ta có:

(i) *Luật đồng nhất:*
$A \cap U = A$
$A \cup \varnothing = A$

(ii) *Luật thống trị:* $A \cup U = U$
$A \cap \varnothing = \varnothing$

(iii) *Luật bất biến:* $A \cap A = A$
$A \cup A = A$

(iv) *Phần bù:* $\left(A^c\right)^c = A$

(v) *Luật giao hoán:*
$A \cup B = B \cup A$
$A \cap B = B \cap A$

(vi) *Luật kết hợp:*
$A \cup (B \cup C) = (A \cup B) \cup C$
$A \cap (B \cap C) = (A \cap B) \cap C$

(vii) *Luật phân phối:*
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

(viii) *Luật De Morgan:*
$(A \cup B)^c = A^c \cap B^c$
$(A \cap B)^c = A^c \cup B^c$

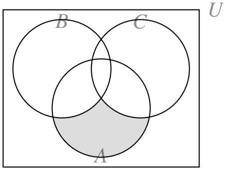
(ix) *Luật hấp thụ:*
$A \cup (A \cap B) = A$
$A \cap (A \cup B) = A$

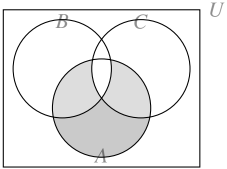
(x) *Luật bù trừ:*
$A \cup A^c = U$
$A \cap A^c = \varnothing$

Hệ quả 1.3 ((De Morgan)).

Cho A,B,C là các tập hợp. Khi đó, ta có:

- $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$
- $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$





- $A \setminus (B \cup C) = A \cap (B \cup C)^c = A \cap (B^c \cap C^c)$
  $= (A \cap B^c) \cap (A \cap C^c) = (A \setminus B) \cap (A \setminus C)$
- $A \setminus (B \cap C) = A \cap (B \cap C)^c = A \cap (B^c \cup C^c)$
  $= (A \cap B^c) \cup (A \cap C^c) = (A \setminus B) \cup (A \setminus C)$

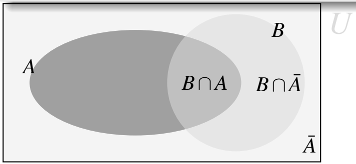
Tập hợp tất cả các tập con

Cho X là một tập hợp, tập hợp tất cả các tập con của X được ký hiệu là $\mathscr{P}(X)$. 2^X

Chú ý. Chúng ta cần cẩn thận khi dùng các ký hiệu $\in$ hay $\subset$. Ta xét một ví dụ đơn giản sau.

Ví dụ 1.4.

Cho $A := \{a,b,c\}$. Khi đó các mệnh đề sau là đúng: $a \in A$, $\{b\} \subset A$, $\{c\} \in \mathscr{P}(A)$,
nhưng các mệnh đề như $\{a\} \in A$ hay $b \subset A$ là sai.



Tính chất

Với mọi tập $A,B \subset U$, ta luôn có:

$$B = B \cap U = B \cap (A \cup \bar{A}) = (B \cap A) \cup (B \cap \bar{A})$$

Mệnh đề 1.5.

Cho $\{A_i\}_{i\in I}$ *là một họ các tập con của* U *và* $B \subset U$*. Khi đó,*

a) $$\bigcup_{i\in I}(B\cap A_i) = B\cap\left(\bigcup_{i\in I}A_i\right)$$
$$\bigcap_{i\in I}(B\cup A_i) = B\cup\left(\bigcap_{i\in I}A_i\right).$$

b) $$\left(\bigcup_{i\in I}A_i\right)^c = \bigcap_{i\in I}A_i^c$$
$$\left(\bigcap_{i\in I}A_i\right)^c = \bigcup_{i\in I}A_i^c$$

Bài tập 1.4.

Hãy xác định tập hợp X biết rằng: $\{1;3;5;7\}\subset X$, $\{3;5;9\}\subset X, X\subset\{1;3;5;7;9\}$.

Giải.

Bài tập 1.5.

Cho hai tập hợp bất kì A,B. Chứng minh rằng $A \cup B = A \cap B \Leftrightarrow A = B$.

Giải.

Bài tập 1.6.

Cho $A = \{1,2,3,4,5\}$ và $B = \{1,3,5,7\}$. Tìm các tập hợp $A \setminus B, B \setminus A$.

Giải.

Bài tập 1.7.

Cho hai tập hợp A,B. Biết $A\setminus B=\{1,2\}, B\setminus A=\{3\}$ và $B=\{3,4,5\}$. Tìm tập hợp A.

Giải.

Bài tập 1.8.

Chứng minh rằng: nếu $A\setminus B=\varnothing$ thì $A\subset B$.

Giải.

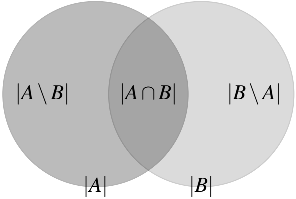
Bài tập 1.9.

Chứng minh rằng: nếu $A\setminus B=B\setminus A$ thì $A=B$

Giải.

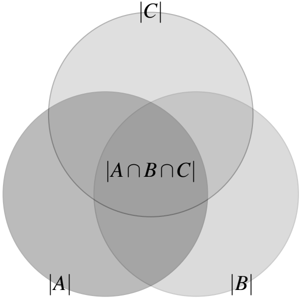
Định nghĩa 1.6.

Cho A là một tập có hữu hạn phần tử. Số các phần tử của A được ký hiệu là $|A|$.



Cho A và B là hai tập hữu hạn. Khi đó, ta có:

$$\begin{aligned}
|A \cup B| &= |A| + |B \setminus A| \\
&= |A \setminus B| + |B| \\
&= |A \setminus B| + |A \cap B| + |B \setminus A| \\
&= |A \setminus B| + |A \cap B| + |B \setminus A| + |A \cap B| - |A \cap B| \\
\hline
|A \cup B| &= |A| + |B| - |A \cap B|
\end{aligned}$$



Cho A, B và C là các tập hữu hạn. Khi đó, ta có:

$$\begin{aligned}
|(A \cup B) \cup C| &= |A \cup B| + |C| - |(A \cup B) \cap C| && \text{(áp dụng công thức cho hai tập)} \\
&= \big[|A| + |B| - |A \cap B|\big] + |C| - |(A \cap C) \cup (B \cap C)| && \text{(phân phối)} \\
&= |A| + |B| + |C| - |A \cap B| - \big[|A \cap C| + |B \cap C| - |(A \cap C) \cap (B \cap C)|\big] \\
&= |A| + |B| + |C| - |A \cap B| - \big[|A \cap C| + |B \cap C| - |A \cap B \cap C|\big] \\
\hline
|A \cup B \cup C| &= |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|
\end{aligned}$$

Bài tập 1.10.

Trong năm vừa qua, trường THPT A có 25 bạn thi học sinh giỏi 2 môn Văn và Toán. Trong đó có 14 bạn thi Toán và 16 bạn thi Văn. Hỏi trường có bao nhiêu bạn thi cả 2 môn Văn và Toán.

Giải.

Bài tập 1.11.

Lớp 10A có 15 bạn thích môn Văn, 20 bạn thích môn Toán. . Trong lớp có 8 bạn thích cả 2 môn và 10 bạn không thích môn nào trong 2 môn Văn và Toán. Hỏi lớp 10A có bao nhiêu bạn?

Giải.

Bài tập 1.12.

Mỗi học sinh của lớp $10A$ đều chơi bóng đá hoặc bóng chuyền. Biết rằng có 25 bạn chơi bóng đá, 20 bạn chơi bóng chuyền và 10 bạn chơi cả 2 môn thể thao. Hỏi lớp $10A$ có bao nhiêu học sinh?

Giải.

Bài tập 1.13.

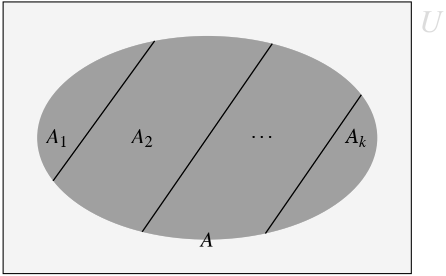
Ở xứ sở thần thoại, ngoài các vị thần thì còn có các sinh vật gồm: 27 con người, 311 yêu quái một mắt, 205 yêu quái tóc rắn và yêu quái vừa một mắt vừa tóc rắn. Tìm số yêu quái vừa một mắt vừa tóc rắn biết có tổng số sinh vật là 500.

Giải.

Định Lý 1.7.

Cho k tập hữu hạn $A_1, A_2, \ldots, A_k$ *đôi một rời nhau, i.e.* $A_i \cap A_j = \varnothing, \forall i, j \in \{1, 2, \ldots, k\} : i \neq j$.

Khi đó, ta có: $$|A_1 \cup A_2 \cup \cdots \cup A_k| = \sum_{i=1}^{k} |A_i| = |A_1| + |A_2| + \cdots + |A_k|.$$



Quy tắc cộng

Để thực hiện một công việc A có thể sử một trong k phương án: $A_1, A_2, \ldots, A_k$. Nếu

- phương án A_1 có n_1 cách thực hiện,
- phương án A_2 có n_2 cách thực hiện,

  $\vdots$
- phương án A_k có n_k cách thực hiện,

thì tổng số cách để thực hiện công việc A là
$$\sum_{i=1}^{k} n_i = n_1 + n_2 + \cdots + n_k.$$

Định Lý 1.8.

Cho k tập hữu hạn $A_1, A_2, \ldots, A_k$*, trong đó:* $|A_1| = n_1$, $|A_2| = n_2$, $\ldots$, $|A_k| = n_k$*. Khi đó, số phần tử của tích Descartes* $A_1 \times A_2 \times \cdots \times A_k = \left\{(x_1, x_2, \ldots, x_k) \mid x_i \in A_i, \forall i = 1, 2, \ldots, k\right\}$ *là:*

$$|A_1 \times A_2 \times \cdots \times A_k| = \prod_{i=1}^{k} |A_i| = \prod_{i=1}^{k} n_i = n_1 \times n_2 \times \cdots \times n_k.$$

Quy tắc nhân

Để thực hiện một công việc A, ta phải thực hiện lần lượt k bước (giai đoạn): $B_1, B_2, \ldots, B_k$. Nếu

- bước thứ nhất (B_1) có n_1 cách thực hiện,
- bước thứ hai (B_2) có n_2 cách thực hiện,

  ⋮
- bước thứ k (B_k) có n_k cách thực hiện,

thì tổng số cách để thực hiện công việc A là: $\displaystyle\prod_{i=1}^{k} n_i = n_1 \times n_2 \times \cdots \times n_k$.

Chỉnh hợp - Hoán vị - Tổ hợp

Cho A là một tập có n phần tử, i.e. $|A| = n$, và k là một số nguyên thỏa $0 \leq k \leq n$.

- Một bộ **có thứ tự** bao gồm k phần tử phân biệt được lấy ra từ tập A được gọi là một **chỉnh hợp (*không lặp*)** chập k từ n phần tử của tập A. **Số chỉnh hợp chập** k **của** n **phần tử** là
$$P(n,k) = {}^{n}P_k = P_k^n = {}_nP_k = A_n^k = \frac{n!}{(n-k)!} = n \times (n-1) \times \cdots \times (n-k+1)$$
- Một bộ **có thứ tự** bao gồm n phần tử phân biệt được lấy ra từ tập A được gọi là một **hoán vị** của tập A. **Số hoán vị của tập** A là
$$P(n,n) = {}^{n}P_n = P_n^n = {}_nP_n = A_n^n = \frac{n!}{(n-n)!} = n! = n \times (n-1) \times \cdots \times 1$$
- Một tập con bao gồm k phần tử phân biệt được lấy ra từ tập A được gọi là một **tổ hợp (*không lặp*)** chập k từ n phần tử của tập A. **Số tổ hợp chập** k **của** n **phần tử** là
$$C(n,k) = {}_nC_k = \binom{n}{k} = \frac{{}_nP_k}{k!} = C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k! \times (n-k)!} = \frac{n \times (n-1) \times \cdots \times (n-k+1)}{1 \times 2 \times \cdots \times k}.$$

Ví dụ 1.9.

Quán Lifetime coffee phục vụ bữa trưa đồng giá 100 nghìn đồng, mỗi suất ăn gồm có:

- **Cơm**: gà kho gừng, giò xông khói, bò xào, cá thu xốt cà.
- **Canh**: canh gà lá giang, canh chua cá lóc, canh cá thác lác với cải bẹ xanh.
- **Bánh**: Bánh quy chocolate, bánh trái cây.
- **Thức uống**: trà, cà phê, Coke, nước cam, nước tinh khiết.

Quán này có thể phục vụ bao nhiêu suất ăn khác nhau?

Giải.

Ví dụ 1.10.

Thầy dạy môn XSTK A có một số bài tập gồm: 3 bài toán khó, 5 bài toán trung bình và 6 bài toán dễ và 4 bài siêu dễ. Thầy muốn lập một đề thi gồm 1 câu hỏi khó, 1 câu hỏi trung bình, 1 câu hỏi dễ và 1 câu siêu dễ. Hỏi có bao nhiêu cách chọn bài cho đề thi?

Giải.

Ví dụ 1.11.

Cho $A = \{a,b,c,d\}$, hãy liệt kê tất cả các chỉnh hợp chập 3 và các tổ hợp chập 3 từ 4 phần tử của tập A.

abc acb abd adb acd adc bac bca bad bda bcd bdc

cab cba cad cda cbd cdb dab dba dac dca dbc dcb

$\{a,b,c\},\{a,c,d\},\{a,b,d\},\{b,c,d\}$

$$4 = C_4^3 = \frac{A_4^3}{3!} = \frac{24}{6}$$

Bài Tú Lơ Khơ

Bộ bài Tú-Lơ-Khơ gồm có 52 lá:

- 13 giá trị (số): $2, 3, 4, \ldots, 9, 10, J, Q, K, A$.
- 4 chất (nước): ♥ , ♦ , ♣ , ♠ .

Bài tập 1.14.

a) Có bao nhiêu bộ năm lá bài mà trong đó chỉ có đúng một đôi (chỉ có đúng 2 lá bài cùng giá trị)? (đôi)

b) Có bao nhiêu bộ năm lá bài mà trong đó: 3 lá bài có cùng giá trị, hai lá còn lại có giá tri khác nhau và khác giá trị chung của 3 lá ban đầu? (xám)

Giải.

Bài tập 1.15.

Một hộp gồm 100 sản phẩm, trong đó có 97 chính phẩm và 3 phế phẩm. Nhân viên kiểm soát chất lượng lấy ngẫu nhiên đồng thời 5 sản phẩm để kiểm tra. Có bao nhiêu cách chọn ra 5 sản phẩm từ hộp trong đó:

a) có đúng 2 phế phẩm.

b) có ít nhất 2 phế phẩm.

c) có không quá 1 phế phẩm.

d) không có quá 4 chính phẩm.

Giải.

Thí nghiệm ngẫu nhiên - Không gian mẫu

- **Thí nghiệm ngẫu nhiên** là thí nghiệm ở đó kết quả ở đầu ra không được xác định duy nhất từ những hiểu biết về đầu vào.
- Kết quả đầu ra của thí nghiệm được quy định là **kết quả đơn**, *không phân tách được*, mỗi lần thử chỉ có một kết quả. Vì thế ta hay gọi chúng là những kết cục (hay biến cố sơ cấp), ký hiệu bởi ζ hay thêm vào chỉ số: $\zeta_1, \zeta_2, \ldots$

  Ví dụ. Khi ta tung một đồng xu, kết quả thu được sẽ là **S**ấp hoặc **N**gửa.

  Tung một cục xúc xắc 6 mặt, kết quả thu được sẽ là một trong 6 mặt: ⚀, ⚁, ⚂, ⚃, ⚄, ⚅.
- Tập tất cả những kết cục có thể có của một thí nghiệm ngẫu nhiên, ký hiệu bởi Ω (nhiều tài liệu viết là S hoặc U), được gọi là không gian mẫu của thí nghiệm đó.

  Không gian mẫu còn được gọi là *không gian các biến cố sơ cấp*.

Ví dụ 1.12.

- Khi ta tung một đồng xu, không gian mẫu $\Omega = \{S, N\}$.
- Khi ta tung một đồng xu hai lần, không gian mẫu $\Omega = \big\{(S,S), (S,N), (N,S), (N,N)\big\}$.
- Khi ta tung một cục xúc xắc 6 mặt, không gian mẫu $\Omega = \Big\{$ ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ $\Big\}$.

Biến cố

- Một biến cố là một tập con của Ω.
  Bản thân tập Ω cũng là một biến cố, được gọi là **biến cố chắc chắn** (luôn luôn xảy ra).
  Biến cố trống (biến cố không) *không chứa* bất cứ kết cục nào, ký hiệu bởi $\varnothing$, được gọi là biến cố bất khả (hay biến cố không thể), tức là không bao giờ xảy ra.
  Biến cố $\{\zeta\}$ chỉ gồm một kết cục ζ được gọi là biến cố sơ cấp
- Các biến cố được ký hiệu bởi chữ cái in hoa hoặc thêm chỉ số: $\mathrm{A}, \mathrm{B}, \ldots, \mathrm{A}_1,\ \mathrm{A}_2, \ldots$
  Chúng ta có thể thể hiện biến cố bằng cách liệt kê các kết cục hoặc nêu các thuộc tính của nó, tất cả được viết trong dấu ngoặc nhọn $\{\cdot\}$ (giống như cách biểu diễn tập hợp).

Ví dụ 1.13.

A: "Tung đồng xu một lần và thu được mặt sấp" $\leftrightarrow$ $A = \{S\}$.

B: "Tung một đồng xu hai lần và được ít nhất một lần sấp" $\leftrightarrow$ $B = \{SS, SN, NS\}$.

C: "Tung một cục xúc xắc và thu được một số chẵn" $\leftrightarrow$ $C = \Big\{$ , ,  $\Big\}$.

- Nếu kết quả của lần thử nào đó là ζ và $\zeta \in A$ thì ta nói biến cố A **xảy ra** ở lần thử này.
  Nếu như $\zeta \notin A$ thì ta nói biến cố A **không xảy ra** ở lần thử này.

Định nghĩa 1.14.

Cho $A, B \subset \Omega$ là hai biến cố.

- Nếu $A \subset B$ thì ta nói "*biến cố A kéo theo biến cố B*" (nghĩa là: nếu biến cố A xảy ra thì biến cố B xảy ra).
- Nếu $A = B$ thì ta nói "*biến cố A tương đương với biến cố B*"
  (i.e. biến cố A xảy ra khi và chỉ khi biến cố B xảy ra).
- *Biến cố tổng của A và B* chính là biến cố $A \cup B$, còn được ký hiệu là $A + B$. $A + B \equiv A \cup B$
  $(A + B)$ xảy ra khi và chỉ khi có ít nhất một trong hai biến cố A, B xảy ra.
- *Biến cố tích của A và B* chính là biến cố $A \cap B$, còn được ký hiệu là $A \cdot B$. $A \cdot B \equiv A \cap B$
  $(A \cdot B)$ xảy ra khi và chỉ khi có cả hai biến cố A, B đồng thời xảy ra.
- A và B được gọi là **hai biến cố xung khắc** nếu A và B không bao giờ đồng thời xảy ra, nghĩa là
  $A \cdot B = A \cap B = \varnothing$.
  n biến cố $A_1, A_2, \ldots A_n$ được gọi là đôi một xung khắc nếu $A_i \cap A_j = \varnothing, \forall i, j = \overline{1,n} : i \neq j$
- A và B được gọi là hai biến cố **đối lập** nếu với mọi phép thử, trong hai biến cố A và B, luôn luôn có đúng một và chỉ một biến cố xảy ra. Khi đó, ta có thể chứng tỏ được rằng: $B = \bar{A}$.
  Với mọi biến cố A, ta luôn có: A đối lập với $\bar{A}$ $\quad A \cap \bar{A} = \varnothing, \quad A \cup \bar{A} = \Omega.$
- *Hiệu của hai biến cố A và B* là biến cố $A \setminus B$, xảy ra khi: A xảy ra và B không xảy ra.
  $A \setminus B = A \cdot \bar{B}, \quad B \setminus A = B \cdot \bar{A}.$

Ví dụ 1.15.

Thực hiện phép thử "tung một cục xúc xắc"

- A: "mặt xuất hiện là mặt chẵn"
- B: "mặt xuất hiện là mặt lẻ"
- C: "mặt xuất hiện là lớn hơn 3"
- D: "mặt xuất hiện là mặt "
- E: "mặt xuất hiện là  hoặc "

Khi đó, ta có:

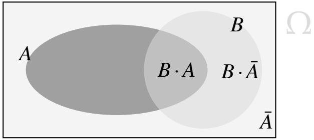
- $D \subset B$ và $E \subset B$
- $D + E = B$
- $B \cdot C = D$
- A và D xung khắc nhau; A và E xung khắc nhau.
- A, D, E đôi một cung khắc.
- A và B đối lập
- $B \setminus D = E$ và $B \setminus C = E$.

Chú ý

- Nếu A và B đều đồng thời xảy ra hoặc đồng thời không xảy ra thì A **và** B **không đối lặp.**
- $\bar{A}$ **xảy ra** khi và chỉ khi A **không xảy ra.**
- Nếu *A và B đối lặp nhau* thì A và B xung khắc.

Biến cố độc lập - Hệ biến cố độc lập toàn phần

- Hai biến cố A và B được gọi là **độc lập** nếu việc xảy ra hay không xảy ra biến cố này không làm ảnh hưởng đến việc xảy ra hay không xảy ra biến cố kia và ngược lại.
- Hệ biến cố $\{A_1, A_2, \ldots, A_n\}$ được gọi là **độc lập toàn phần** nếu mỗi biến cố trong hệ độc lập với tích của một tổ hợp bất kỳ các biến cố còn lại.



Chú ý

Với mọi biến cố $A,B \subset \Omega$, ta luôn có:

$$B = B \cdot \Omega = B \cdot (A + \bar{A}) = B \cdot A + B \cdot \bar{A}$$

Khi đó, A và $\bar{A}$ tạo thành một phân hoạch cho Ω.

Tương tự như các tính chất của tập hợp, ta có các tính chất của biến cố.

Các tính chất của biến cố

Cho $A, B, C \subset \Omega$ là các biến cố. Ta có:

1. (i) Luật đồng nhất: $A \cdot \Omega = A$, $A + \varnothing = A$
2. (ii) Luật thống trị: $A + \Omega = \Omega$, $A \cdot \varnothing = \varnothing$
3. (iii) Luật bất biến: $A \cdot A = A$, $A + A = A$
4. (iv) Phần bù: $\overline{(\bar{A})} = A$
5. (v) Luật giao hoán: $A + B = B + A$, $A \cdot B = B \cdot A$
6. (vi) Luật kết hợp: $A + (B + C) = (A + B) + C$, $A \cdot (B \cdot C) = (A \cdot B) \cdot C$
7. (vii) Luật phân phối: $A + (B \cdot C) = (A + B) \cdot (A + C)$, $A \cdot (B + C) = (A \cdot B) + (A \cdot C)$
8. (viii) Luật De Morgan: $\overline{(A + B)} = \bar{A} \cdot \bar{B}$, $\overline{(A \cdot B)} = \bar{A} + \bar{B}$
9. (ix) Luật hấp thụ: $A + (A \cdot B) = A$, $A \cdot (A + B) = A$. Nếu $A \subset B$ thì $A + B = B$ và $A \cdot B = A$.
10. (x) Luật bù trừ: $A + \bar{A} = \Omega$, $A \cdot \bar{A} = \varnothing$

Biến cố đồng khả năng

Các biến cố A, B được gọi là **đồng khả năng** nếu khả năng chúng xuất hiện trong phép thử T là như nhau.

Ví dụ 1.16.

Tung ngẫu nhiên một đồng xu, gọi S là biến cố "đồng xu xuất hiện mặt sấp", N là biến cố "đồng xu xuất hiện mặt ngửa". Khi đó S, N là hai biến cố đồng khả năng.
Tung ngẫu nhiên một cục xúc xắc, gọi C là biến cố "mặt xuất hiện là mặt chẵn", L là biến cố "mặt xuất hiện là mặt lẻ". Khi đó C, L là hai biến cố đồng khả năng.

Hệ biến cố đầy đủ

Hệ biến cố $\{\mathrm{A}_1,\ \mathrm{A}_2,\ldots,\mathrm{A}_\mathrm{n}\}$ được gọi là một **hệ đầy đủ các biến cố** nếu hệ thỏa mãn hai điều sau:

1. Các biến cố $A_1, A_2, \ldots A_n$ là đôi một xung khắc, i.e. $A_i \cap A_j = \varnothing, \forall i,j = \overline{1,n} : i \neq j$.
2. $\Omega = A_1 \cup A_2 \cup \cdots \cup A_n = \bigcup\limits_{i=1}^{n} A_i$.

Ví dụ 1.17.

a) Trong ví dụ 1.15, $\{A, D, E\}$ là một hệ biến cố đầy đủ.
b) Với mọi biến cố $A \subset \Omega$, ta có $\left\{A, \bar{A}\right\}$ là một hệ biến cố đầy đủ.

Định nghĩa xác suất cổ điển

Giả sử phép thử ngẫu nhiên T có không gian mẫu là Ω, chỉ có ***hữu hạn*** *biến cố sơ cấp đồng khả năng*.

Khi đó, xác suất của biến cố A được định nghĩa bởi công thức sau:

$$P(A) = \frac{\text{Số lượng các biến cố sơ cấp trong } A}{\text{Số lượng các biến cố sơ cấp trong khong gian mẫu } \Omega} = \frac{|A|}{|\Omega|}$$

Ví dụ 1.18.

Từ bộ bài Tú-lơ-khơ 52 lá, ta rút ra một lá bài. Hãy tính xác suất để lá bài được rút ra:

a) là J. | b) là 6 ♣ . | c) là 3 hoặc ♦ . | d) là 3 hoặc 6.

Giải.

Chú ý

1. Xác suất của biến cố chắc chắn, Ω, là $P(\Omega) = |\Omega| \big/ |\Omega| = 1$.
   Tổng xác suất của tất cả các biến cố sơ cấp là 1.
2. Xác suất của biến cố bất khả, $\varnothing$, là $P(\varnothing) = |\varnothing| \big/ |\Omega| = 0 \big/ |\Omega| = 0$.
3. Với mọi biến cố $A \subset \Omega$, ta có: $0 \leq |A| \leq |\Omega|$.
   Do đó, $0 \leq P(A) = |A| \big/ |\Omega| \leq 1$.
4. Nếu biến cố $A \subset B \subset \Omega$ thì $0 \leq |A| \leq |B| \leq |\Omega|$ và
   $$0 \leq P(A) = \frac{|A|}{|\Omega|} \leq P(B) = \frac{|B|}{|\Omega|} \leq 1$$
5. Với mọi biến cố $A \subset \Omega$, do A và $\bar{A}$ là hai biến cố đối lập nên $|\Omega| = |A| + |\bar{A}|$. Khi đó,
   $$P(A) + P(\bar{A}) = \frac{|A|}{|\Omega|} + \frac{|\bar{A}|}{|\Omega|} = \frac{|A| + |\bar{A}|}{|\Omega|} = \frac{|\Omega|}{|\Omega|} = 1$$

Chú ý

- Định nghĩa xác suất theo phương pháp cổ điển có ưu điểm là tính được chính xác giá trị của xác suất mà (có thể) không phải thực hiện phép thử.
- Định nghĩa xác suất theo phương pháp cổ điển có hạn chế là: chỉ xét cho hệ hữu hạn các biến cố sơ cấp đồng khả năng (không phải lúc nào cũng phân tích được thành hữu hạn các biến cố đồng khả năng)

Ví dụ 1.19.

Tung một cục xúc xắc hai lần, tính xác suất để

a) tổng số nút là 6.

b) số nút ở hai lần giống nhau.

c) tổng số nút là 7 hoặc 11.

d) tổng số nút nhỏ hơn 10.

Giải.

Định nghĩa xác suất theo thống kê

Giả sử thực hiện 1 phép thử nào đó n lần độc lập (kết quả của phép thử sau không phụ thuộc vào kết quả của phép thử trước), trong đó biến cố A xảy ra m lần.

Khi đó: m gọi là tần số xuất hiện của biến cố A;

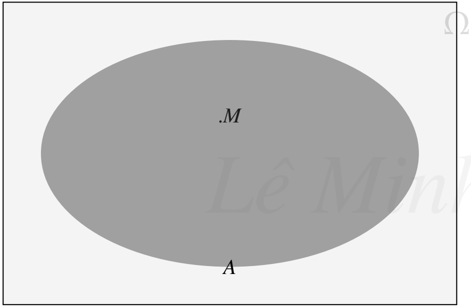
$\mathrm{f} = \frac{m}{n}$ gọi là tần suất của biến cố A.

Khi $\mathrm{n} \to \infty$, tần suất f đạt giá trị ổn định và giá trị đó được xem là xác suất của biến cố A.

Ta có: $P(A) = \lim\limits_{n\to\infty} f = \lim\limits_{n\to\infty} \frac{m}{n}$

Các chuyên gia tung đồng xu

Người thực hiện	Tổng số	Ngửa	tỷ lệ
Count Buffon (TK 18)	4040	2048	50.69%
Karl Pearson (1900)	24000	12012	50.05%
John Kerrich (WW II)	10000	5067	50.67%



Định nghĩa xác suất theo hình học

Xét một phép thử đồng khả năng, có không gian các biến cố sơ cấp là *miền hình học* Ω, có **độ đo** (độ dài, diện tích, thể tích,...) hữu hạn, khác không.
Giả sử một (chất) điểm M rơi ngẫu nhiên vào miền Ω, xét miền con A của Ω.
Khi đó, xác suất để M rơi vào miền A là:

$$P(A) = \frac{\text{Độ đo miền } A}{\text{Độ đo miền } \Omega}$$

Các tính chất của xác suất

1. $P(\varnothing) = 0$.
2. $P(\Omega) = 1$.
3. Với mọi biến cố $A \subset \Omega$, $0 \leq P(A) \leq 1$.
4. Với mọi biến cố $A \subset \Omega$, $P(A) + P(\bar{A}) = 1$.
5. Cho A, B là hai biến cố.
   Nếu $A \subset B$ thì $P(A) \leq P(B)$.

Đại số

Cho $\mathscr{F}$ là một họ khác trống các biến cố của không gian mẫu Ω, i.e. $\mathscr{F} \subset \mathscr{P}(\Omega)$. $\mathscr{F}$ được gọi là một đại số (hay trường) nếu $\mathscr{F}$ thỏa:

(i) $\Omega \in \mathscr{F}$;

(ii) Nếu $A \in \mathscr{F}$ thì $\bar{A} \in \mathscr{F}$;

(iii) Nếu $A,B \in \mathscr{F}$ thì $(A \cup B) \in \mathscr{F}$.

Xác suất

$(\Omega, \mathscr{F})$ là một bộ gồm không gian mẫu Ω và đại số $\mathscr{F}$. Xác suất là một ánh xạ $P : \mathscr{F} \to \mathbb{R}$ thỏa:

(i) $P(A) \geq 0, \forall A \in \mathscr{F}$;

(ii) $P(\Omega) = 1$;

(iii) Nếu $A,B \in \mathscr{F}$ là hai biến cố xung khắc thì $P(A+B) = P(A) + P(B)$.

σ-đại số

Cho $\mathscr{F}$ là một họ khác trống các biến cố của không gian mẫu Ω, i.e. $\mathscr{F} \subset \mathscr{P}(\Omega)$. $\mathscr{F}$ được gọi là một σ-đại số (hay σ-trường) nếu $\mathscr{F}$ thỏa:

(i) $\Omega \in \mathscr{F}$;

(ii) Nếu $A \in \mathscr{F}$ thì $\bar{A} \in \mathscr{F}$;

(iii) Nếu $A_1, A_2, \ldots \in \mathscr{F}$ thì $\bigcup_{i=1}^{\infty} A_i \in \mathscr{F}$.

Xác suất

$(\Omega, \mathscr{F})$ là một bộ gồm không gian mẫu Ω và σ-đại số $\mathscr{F}$. Xác suất là một ánh xạ $P : \mathscr{F} \to \mathbb{R}$ thỏa:

(i) $P(A) \geq 0, \forall A \in \mathscr{F}$;

(ii) $P(\Omega) = 1$;

(iii) Nếu $A_1, A_2, \ldots \in \mathscr{F}$ là dãy các biến cố đôi một xung khắc thì $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.

Không gian xác suất

Bộ ba $(\Omega, \mathscr{F}, P)$ bao gồm không gian mẫu Ω, (σ-)đại số $\mathscr{F}$ và xác suất P là một không gian xác suất. Mỗi một thí nghiệm ngẫu nhiên được mô hình hóa bởi một không gian xác suất nào đó.

Các tính chất của xác suất

1. $P(\varnothing) = 0$ và $P(\Omega) = 1$.
2. Với mọi biến cố $A \subset \Omega$, $0 \le P(A) \le 1$.
3. Với mọi biến cố $A \subset \Omega$, $P(A) + P(\bar{A}) = 1$.
4. Cho A, B là hai biến cố. Nếu $A \subset B$ thì $P(A) \le P(B)$.
5. Nếu A và B xung khắc thì $P(A+B) = P(A) + P(B)$.
6. Nếu $A_1, A_2, \dots, A_n$ đôi một xung khắc thì $P(A_1 + \cdots + A_n) = P(A_1) + \cdots + P(A_n)$.
7. Với mọi biến cố A và B, ta có:
$$P(A+B) = P(A\bar{B} + AB + \bar{A}B)$$
$$= P(A\bar{B}) + P(AB) + P(\bar{A}B)$$
$$= P(A\bar{B}) + P(AB) + P(\bar{A}B) + P(AB) - P(AB)$$
$$= P(A) + P(B) - P(AB)$$
8. Với mọi biến cố A, B và C, ta có: $P(A+B+C)$ $= P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC)$ $+ P(ABC)$.

-**Nguyên lý xác suất nhỏ**: Một biến cố có xác suất rất nhỏ (xấp xỉ bằng 0) thì có thể cho rằng, trong thực tế, biến cố đó không xảy ra trong một (vài) phép thử.

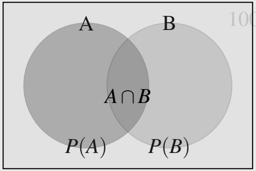
-**Nguyên lý xác suất lớn**: Một biến cố có xác suất rất lớn, xấp xỉ bằng 1 thì có thể cho rằng, trong thực tế, biến cố đó chắc chắn xảy ra trong một (vài) phép thử.

Ví dụ 1.20.

Qua điều tra sinh viên năm *I*, ta biết 40% sinh viên có chứng chỉ ngoại ngữ B1, 55% sinh viên có chứng chỉ tin học ứng dụng và 10% sinh viên có cả hai chứng chỉ này. Chọn ngẫu nhiên một sinh viên năm *I*. Tính xác suất để sinh viên đó không có chứng chỉ ngoại ngữ B1 và cũng không có chứng chỉ tin học ứng dụng.

Giải.

Xác suất có điều kiện



Cho trước hai biến cố A, B với $P(A) > 0$.

Xác suất để biến cố B xảy ra, khi đã biết A đã xảy ra (A: thông tin) được gọi là xác suất có điều kiện của biến cố B với điều kiện A, ký hiệu là $P(B \mid A)$, xác định bởi:

$$P(B \mid A) = \frac{P(AB)}{P(A)} \quad (P(A) > 0)$$

Tính chất

• $0 \leq P(B \mid A) = \dfrac{P(AB)}{P(A)} \leq 1$;

$0 \leq P(A \mid B) = \dfrac{P(AB)}{P(B)} \leq 1$;

$P(A \mid A) = 1 = P(B \mid B)$.

• $P(\bar{B} \mid A) = 1 - P(B \mid A)$

• $P\left((B_1 + B_2)A\right) = P\left(B_1A + B_2A\right)$

$$= P(B_1A) + P(B_2A) - P(B_1B_2A)$$

Do đó,

$$\frac{P\left((B_1 + B_2)A\right)}{P(A)} = \frac{P(B_1A)}{P(A)} + \frac{P(B_2A)}{P(A)} - \frac{P(B_1B_2A)}{P(A)},$$

hay

$P(B_1 + B_2 \mid A) = P(B_1 \mid A) + P(B_2 \mid A) - P(B_1B_2 \mid A)$.

Nếu B_1, B_2 xung khắc thì

$P(B_1 + B_2 \mid A) = P(B_1 \mid A) + P(B_2 \mid A)$

• Nếu $B_1, B_2, \ldots, B_n$ đôi một xung khắc thì

$P(B_1 + \cdots + B_n \mid A) = P(B_1 \mid A) + \cdots + P(B_n \mid A)$.

Ví dụ 1.21.

Cho một hộp đựng 32 bi, trong đó có 23 bi đỏ và 9 bi xanh. Lấy lần lượt 2 bi (lấy không hoàn lại).

a) Tính xác suất để lần thứ 1 lấy được bi đỏ?

b) Tính xác suất để lần thứ 2 lấy được bi đỏ biết lần thứ nhất lấy được bi đỏ?

c) Tính xác suất để lần thứ 2 lấy được bi đỏ biết lần thứ nhất không lấy được bi đỏ?

Quy tắc nhân

1. Cho A, B là hai biến cố với $P(A) > 0$, ta có: $$P(AB) = P(A)\,P(B \mid A).$$
2. Cho A, B, C là ba biến cố với $P(A) \geq P(AB) > 0$, ta có: $$P(ABC) = P(AB) \cdot P(C \mid AB) = P(A) \cdot P(B \mid A) \cdot P(C \mid AB).$$
3. Cho $A_1, A_2, \ldots, A_n$ là các biến cố sao cho $P(A_1A_2 \cdots A_{n-1}) > 0$, ta có: $$P(A_1A_2 \ldots A_{n-1}A_n) = P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1A_2) \cdots P(A_n \mid A_1A_2 \cdots A_{n-1}).$$

Sự độc lập của hai biến cố

Ta gọi hai biến cố A và B là **độc lập** nếu $P(AB) = P(A)\,P(B)$.

Khi $P(A) \neq 0$,
A,B độc lập $\quad \Leftrightarrow \quad P(B) = \dfrac{P(AB)}{P(A)} = P(B \mid A).$

Khi $P(B) \neq 0$,
A,B độc lập $\quad \Leftrightarrow \quad P(A) = \dfrac{P(AB)}{P(B)} = P(A \mid B).$

Trả lời nhanh

1. Tung một đồng xu hai lần.
   N_1: "lần tung thứ nhất, được mặt ngửa";
   N_2: "lần tung thứ hai, được mặt ngửa".
   N_1 và N_2 có độc lập hay không?

2. Tung một đồng xu ba lần.
   N_1: "lần tung thứ nhất, được mặt ngửa";
   A: "sau ba lần tung, được đúng hai lần ngửa".
   N_1 và A có độc lập hay không?

Ví dụ 1.22.

Trong ngăn kéo có 9 đôi vớ khác màu. Lấy ngẫu nhiên lần lượt 6 chiếc vớ. Hãy tính xác suất để có ít nhất một đôi vớ cùng màu.

Định Lý 1.23.

Cho A và B là cặp biến cố độc lập. Khi đó, $\bar{A}$ và B; A và $\bar{B}$; $\bar{A}$ và $\bar{B}$ cũng là các cặp biến cố độc lập.

Chứng minh.

- Vì $B = \left(\bar{A} \cap B\right) \cup (A \cap B)$ và $\left(\bar{A} \cap B\right)$, $(A \cap B)$ xung khắc nên $P(B) = P\left(\bar{A} \cap B\right) + P(A \cap B)$. Do đó,
  $$P\left(\bar{A} \cap B\right) = P(B) - P(A \cap B)$$
  $$= P(B) - P(A)P(B) \quad (\text{ vì } A \text{ và } B \text{ là độc lập })$$
  $$= (1 - P(A))P(B) = P\left(\bar{A}\right)P(B)$$
  Do vậy, $\bar{A}$ và B cũng là cặp biến cố độc lập.
- Vì $A = \left(A \cap \bar{B}\right) \cup (A \cap B)$ và $\left(A \cap \bar{B}\right)$, $(A \cap B)$ xung khắc nên $P(A) = P\left(A \cap \bar{B}\right) + P(A \cap B)$. Do đó,
  $$P\left(A \cap \bar{B}\right) = P(A) - P(A \cap B)$$
  $$= P(A) - P(A)P(B) \quad (\text{ vì } A \text{ và } B \text{ là độc lập })$$
  $$= P(A)(1 - P(B)) = P(A)P\left(\bar{B}\right)$$
  Vì thế, A và $\bar{B}$ cũng là cặp biến cố độc lập.
- Do $\bar{A} = \left(\bar{A} \cap \bar{B}\right) \cup \left(\bar{A} \cap B\right)$ và $\left(\bar{A} \cap \bar{B}\right)$, $\left(\bar{A} \cap B\right)$ xung khắc nên $P\left(\bar{A}\right) = P\left(\bar{A} \cap \bar{B}\right) P\left(\bar{A} \cap B\right)$. Do đó,
  $$P\left(\bar{A} \cap \bar{B}\right) = P\left(\bar{A}\right) - P\left(\bar{A} \cap B\right)$$
  $$= P\left(\bar{A}\right) - P\left(\bar{A}\right)P(B) \quad (\text{ vì } \bar{A} \text{ và } B \text{ là độc lập })$$
  $$= P\left(\bar{A}\right)(1 - P(B)) = P\left(\bar{A}\right)P\left(\bar{B}\right)$$
  Do vậy, $\bar{A}$ và $\bar{B}$ cũng là cặp biến cố độc lập. □

Công thức xác suất đầy đủ

Cho $A_1, A_2, \ldots, A_k$ là một hệ biến cố đầy đủ, trong đó, $P(A_i) > 0, \forall i = \overline{1,k}$. Với mọi biến cố B, ta có:

$$P(B) = \sum_{i=1}^{k} P(A_i)\, P(B \mid A_i) = P(A_1)\, P(B \mid A_1) + P(A_2)\, P(B \mid A_2) + \cdots + P(A_k)\, P(B \mid A_k).$$

Chứng minh. Vì $A_1, A_2, \ldots, A_k$ là một hệ biến cố đầy đủ nên $B = B\Omega = B(A_1 + A_2 + \cdots + A_k)$
$= BA_1 + BA_2 + \cdots + BA_k$ và $BA_1, BA_2, \ldots, BA_k$ là đôi một xung khắc. Do đó,

$$P(B) = P(BA_1 + B_2 + \cdots + BA_k) = P(BA_1) + P(BA_2) + \cdots + P(BA_k)$$
$$= P(A_1)\, P(B \mid A_1) + P(A_2)\, P(B \mid A_2) + \cdots + P(A_k)\, P(B \mid A_k) = \sum_{i=1}^{k} P(A_i)\, P(B \mid A_i)$$

Công thức Bayes

Cho $A_1, A_2, \ldots, A_k$ là một hệ biến cố đầy đủ, trong đó, $P(A_i) > 0, \forall i = \overline{1,k}$.
Với mọi biến cố B với $P(B) > 0$, ta có:

$$P(A_i \mid B) = \frac{P(A_i B)}{P(B)} = \frac{P(A_i)\, P(B \mid A_i)}{P(B)} = \frac{P(A_i)\, P(B \mid A_i)}{\sum_{j=1}^{k} P(A_j)\, P(B \mid A_j)}.$$

Ví dụ 1.24.

Tô B_1 chứa 3 viên sỏi màu đỏ và 7 viên sỏi màu đen. Tô B_2 chứa 8 viên sỏi màu đỏ và hai viên sỏi màu đen. Bạn Minh tung ngẫu nhiên một cục xúc xắc 6 mặt. Nếu Minh được ít nhất 5 nút thì Minh sẽ bốc ngẫu nhiên một viên sỏi từ tô B_2. Nếu Minh được ít hơn 5 nút thì Minh sẽ bốc ngẫu nhiên một viên sỏi từ tô B_1.
Biết rằng viên sỏi bạn minh bốc được màu đỏ, hãy tính xác suất để viên sỏi đó được bốc ra từ tô B_1.

Giải.

Ví dụ 1.25.

Có ba hộp đựng thuốc: hộp I có 10 lọ trong đó có 7 lọ tốt; hộp II có 15 lọ trong đó có 12 lọ tốt; hộp III có 20 lọ trong đó có 18 lọ tốt. Lấy ngẫu nhiên 1 hộp, rồi từ hộp đó lấy ngẫu nhiên 1 lọ thuốc.

a) Tính xác suất để lọ thuốc được lấy ra là lọ tốt.
b) Biết rằng lọ thuốc được lấy ra là lọ thuốc hỏng. Lọ thuốc hỏng có khả năng thuộc hộp nào cao nhất?

Giải.

Dãy các phép thử Bernoulli

Đối với thí nghiệm ngẫu nhiên nào đó, chúng ta thực hiện n lần thử lặp lại. Chúng ta gọi dãy các phép thử này là dãy các phép thử Bernoulli nếu thoả mãn các điều kiện sau:

1. i) Đây là dãy các phép thử độc lập,
2. ii) Biến cố A xảy ra với xác suất p như nhau ở phép thử thứ i bất kỳ.

Nếu biến cố A xảy ra ở phép thử thứ i, ta nói phép thử thứ i thành công.
Trái lại, nếu nó không xảy ra ở phép thử thứ i, ta nói phép thử thứ i thất bại.

Công thức Bernoulli

Xác suất để biến cố A xuất hiện đúng k lần trong dãy n phép thử Bernoulli, ký hiệu là $P_n(k)$, hay đầy đủ hơn $P_n(k,p)$ hoặc $P(n;k;p)$, được cho bởi : $P_n(k) = P_n(k;p) = P(n;k;p) = C_n^k \cdot p^k \cdot (1-p)^{n-k}$, với $k = 0, 1, \ldots, n$.

Ví dụ 1.26.

Bắn 5 phát súng vào mục tiêu, xác suất trúng đích của mỗi phát là 0,2 . Để phá huỷ mục tiêu cần từ 3 phát trúng đích trở lên. Tính xác suất phá huỷ mục tiêu.

Giải. Xem như chúng ta đã thực hiện dãy 5 phép thử độc lập (mỗi phép thử là một lần bắn). Biến cố mục tiêu bị phá huỷ là biến cố có ít nhất 3 phát trúng đích trở lên. Từ đó, ta có:

$$P = P_5(3;0{,}2) + P_5(4;0{,}2) + P_5(5;0{,}2) = C_5^3 \cdot 0{,}2^3 \cdot 0{,}8^2 + C_5^4 \cdot 0{,}2^4 \cdot 0{,}8^1 + C_5^5 \cdot 0{,}2^5 \cdot 0{,}8^0 = 0{,}05792$$

Ví dụ 1.27.

Một sinh viên thi trắc nghiệm môn Ngoại Ngữ gồm có 10 câu hỏi. Mỗi câu có 4 phương án lựa chọn, trong đó chỉ có 1 phương án đúng. Giả sử sinh viên làm bài bằng cách chọn ngẫu nhiên các câu hỏi. Tính xác suất:

a) Sinh viên vừa đủ 5 điểm.

b) Sinh viên chọn đúng ít nhất 1 câu hỏi.

Giải.

Lược dồ Bernoulli

Một lược dồ Bernoulli (mở rộng) gồm:

- Dãy n phép thử độc lập, đồng nhất.
- Với mỗi phép thử, hệ biến cố $\{A_1, A_2, \ldots, A_k\}$ là một hệ biến cố đầy đủ.

Chú ý. Nếu $P(A_1) = p_1, P(A_2) = p_2, \ldots, P(A_k) = p_k$ thì $p_1 + p_2 + \cdots + p_k = 1$.

Công thức Bernoulli mở rộng

Giả sử ta thực hiện n phép thử độc lập, đồng nhất. Với mỗi phép thử, hệ biến cố $\{A_1, A_2, \ldots, A_k\}$ là một hệ biến cố đầy đủ, trong đó, $P(A_1) = p_1,\ P(A_2) = p_2, \ldots,\ P(A_k) = p_k$ và $p_1 + p_2 + \ldots + p_k = 1$. Khi đó, xác suất để trong n phép thử độc lập, biến cố A_1 xảy ra m_1 lần, biến cố A_2 xảy ra m_2 lần, …, biến cố A_k xảy ra m_k lần, với $m_1 + m_2 + \ldots + m_k = n$, được tính theo công thức:

$$P(n; m_1, m_2, \ldots, m_k) = \frac{n!}{m_1! m_2! \ldots m_k!} p_1^{m_1} \cdot p_2^{m_2} \cdots p_k^{m_k}$$

Ví dụ 1.28.

Lô hàng có 100 sản phẩm trong đó có 30 sản phẩm loại A, 50 sản phẩm loại B và 20 sản phẩm loại C.
Lần lượt rút có hoàn lại 9 sản phẩm để kiểm tra. Tính xác suất để trong 9 lần rút đó có 3 lần rút được sản phẩm loại A, 4 lần rút được sản phẩm loại B và 2 lần rút được sản phẩm loại C.

Giải. Gọi A, B, C lần lượt là biến cố rút được sản phẩm loại A, B, C trong mỗi lần rút. Rõ ràng, hệ biến cố $\{A, B, C\}$ là một hệ đầy đủ các biến cố, với và $P(A) = \frac{30}{100}, P(B) = \frac{50}{100}, P(A) = \frac{20}{100}$.

Do đó: $\quad \mathrm{P}(9;\ 3\mathrm{A},\ 4\mathrm{B},\ 2\mathrm{C}) = \frac{9!}{3!4!2!}\left(\frac{30}{100}\right)^3\left(\frac{50}{100}\right)^4\left(\frac{20}{100}\right)^2 = 0.08505$

Bài tập 1.16.

Một lớp có 60 học sinh trong đó có 28 em giỏi toán, 30 em giỏi lý, 32 em giỏi ngoại ngữ, 15 em vừa giỏi toán vừa giỏi lý, 10 em vừa giỏi lý vừa giỏi ngoại ngữ, 12 em vừa giỏi toán vừa giỏi ngoại ngữ, 2 em giỏi cả 3 môn. Chọn ngẫu nhiên một học sinh của lớp. Tính xác suất để:

a) Chọn được em giỏi ít nhất 1 môn.

b) Chọn được em chỉ giỏi toán.

c) Chọn được em giỏi đúng 2 môn.

Bài tập 1.17.

Hai xạ thủ mỗi người bắn một phát đạn vào bia. Xác suất bắn trúng của người thứ nhất là $p = 0.9$; của người thứ hai là $p = 0.7$. Giả sử hai người bắn độc lập với nhau, tính xác suất để:

a) Cả hai đều bắn trúng.

b) Có đúng một viên đạn trúng bia.

c) Bia bị trúng đạn.

Bài tập 1.18.

Có 2 lô sản phẩm. Lô I có 6 chính phẩm và 4 phế phẩm. Lô II có 5 chính phẩm và 5 phế phẩm. Từ lô I lấy ngẫu nhiên ra 2 sản phẩm và từ lô II lấy ngẫu nhiên ra 1 sản phẩm. Sau đó, chọn ngẫu nhiên 1 sản phẩm từ 3 sản phẩm đó. Tìm xác suất chọn được phế phẩm.

Bài tập 1.19.

Một phân xưởng sản xuất chi tiết máy có hai máy: máy I sản xuất 60% sản phẩm của phân xưởng; máy II sản xuất 40% sản phẩm của phân xưởng. Tỷ lệ sản phẩm bị lỗi của máy I là 7% và tỷ lệ sản phẩm bị lỗi của máy II là 4% . Sản phẩm của phân xưởng sau khi sản xuất được đem trộn lẫn với nhau. Lấy ngẫu nhiên một sản phẩm của phân xưởng này thì thấy sản phẩm đó là sản phẩm bị lỗi, tính xác suất để sản phẩm đó do máy I sản xuất.

Bài tập 1.20.

Có 3 bao chứa cùng loại lúa. Bao 1 nặng 20 kg chứa 1% hạt lép, bao 2 nặng 30 kg chứa 1,2% hạt lép và bao 3 nặng 50 kg chứa 1,5% hạt lép. Trộn cả 3 bao lại rồi bốc ngẫu nhiên 1 hạt thì được hạt lép. Tính xác suất để hạt lép này là của bao thứ ba.

Bài tập 1.21.

Ba kiện hàng đều có 20 sản phẩm, với số sản phẩm tốt trong mỗi kiện tương ứng là 12, 15, 18 sản phẩm. Lấy ngẫu nhiên 1 kiện hàng (giả sử 3 kiện hàng có cùng khả năng), rồi từ kiện đó lấy tùy ý ra 1 sản phẩm.

a) Tính xác suất để sản phẩm chọn ra là tốt.

b) Giả sử sản phẩm chọn ra là tốt, tính xác suất để sản phẩm đó thuộc kiện hàng thứ hai.

Bài tập 1.22.

Hai nhà máy cùng sản xuất ra một loại chi tiết. Năng suất của nhà máy I gấp đôi nhà máy II. Tỷ lệ chi tiết đạt tiêu chuẩn của nhà máy I là 94%, của nhà máy II là 98%. Lấy ngẫu nhiên 1 chi tiết từ lô hàng do 2 nhà máy sản xuất thì được chi tiết đạt tiêu chuẩn. Tính xác suất để chi tiết đó do nhà máy I sản xuất.

Bài tập 1.23.

Theo kết quả điều tra, tỷ lệ bệnh lao ở một vùng là 0,2%. Tính xác suất để khi khám cho 10 người thì:

a) không người nào bị bệnh lao.
b) Có ít nhất 2 người bệnh lao.

Bài tập 1.24.

Một sinh viên thi trắc nghiệm môn ngoại ngữ gồm 20 câu hỏi. Mỗi câu có 4 phần để chọn, trong đó chỉ có 1 phần đúng. Giả sử sinh viên đó đã biết rõ 8 câu hỏi, còn lại thì chọn một cách ngẫu nhiên.

a) Tính xác suất để sinh viên đó làm đúng được toàn bài.
b) Nếu chọn đúng từ phân nữa trở đi thì sinh viên đó sẽ đậu. Tính xác suất để sinh viên đó đậu.

Biến ngẫu nhiên

Một số người cho rằng gọi tên biến ngẫu nhiên là một sự nhầm lẫn, do một biến ngẫu nhiên không phải là một biến mà là một hàm số ánh xạ các biến cố tới các số.
Cho A là một σ-đại số và Ω là không gian các biến cố liên quan tới thực nghiệm đang được tiến hành (không gian các biến cố sơ cấp).

Trong ví dụ thả súc sắc, không gian các biến cố chính là các kết quả có thể của một lần thả, nghãa là $\Omega = \{1,2,3,4,5,6\}$, và A sẽ là tập lũy thừa (tập hợp tất cả các tập con) của Ω.
Trong trường hợp này, một biến ngẫu nhiên thích hợp có thể là hàm đồng nhất (identity function) $X(\omega) = \omega$, sao cho nếu kết quả là 1 thì biến ngẫu nhiên cũng sẽ bằng 1.

Một biến ngẫu nhiên được định nghĩa như là một hàm đo được (measurable function) từ một không gian xác suất tới một *không gian đo được* nào đó. Không gian đo được này là một không gian của các giá trị có thể của biến, và nó thường được lấy là các số thực với Borel $\sigma-$đại số.

Cho không gian xác suất (Ω, A, P). Một hàm $X : \Omega \to \mathbf{R}$ là một biến ngẫu nhiên giá trị thực nếu với mọi tập con $\mathrm{A}_r = \{\omega : X(\omega) \leq r\}$ trong đó $r \in R$, ta cũng có $\mathrm{A}_r \in A$.
Định nghĩa này có tầm quan trọng ở chỗ nó cho phép ta xây dựng hàm phân bố của biến ngẫu nhiên.

BNN không phải là biến số độc lập; nó là hàm số, xác định trên không gian các biến cố sơ cấp Ω.

- Nếu tập giá trị hữu hạn hay vô hạn đếm được thì BNN được gọi là rời rạc.
- Nếu tập giá trị lấp đầy một hoặc một số khoảng thì BNN được gọi là liên tục.

Thường người ta ký hiệu BNN bởi chữ cái in hoa: $\mathrm{X}, \mathrm{Y}, \mathrm{Z}, \ldots$ hoặc có thêm chỉ số: $\mathrm{X}_1, \mathrm{X}_2, \ldots$

Nhận xét. Đòi hỏi tính đo được của BNN mang tính chất toán học thuần tuý, có thể bỏ qua trong các bài toán đơn giản.

Ví dụ 2.1.

Tung con xúc xắc cân đối; mỗi nút trên mặt con xúc xắc được thưởng 10USD. Đặt X bằng 10 lần số nút trên mặt con xúc xắc. Khi đó, X là một BNN.

Không gian các biến cố sơ cấp là $S = \{M_1, \ldots, M_6\}$, trong đó $M_i = \{$ xuất hiện mặt có i nút $\}$. Đặt $X : S \to \mathbb{R}$ với $X(M_1) = 10$, $X(M_2) = 20$, …, $\mathrm{X}(\mathrm{M}_6) = 60$. Tập giá trị của X là $\{10; \ldots; 60\}$; X là BNN rời rạc.

Ví dụ 2.2.

Số phế phẩm có trong lô hàng n sản phẩm. Không nói trước được số phế phẩm là bao nhiêu, đây là BNN rời rạc. Tập giá trị là $\{0; \ldots; \mathrm{n}\}$.

Ví dụ 2.3.

Tuổi thọ của một bóng đèn LED, đó là BNN liên tục; có thể, tập giá trị là $[0; 100000]$ (giờ).

Việc nghiên cứu BNN là tổng quát hơn so với chỉ nghiên cứu biến cố thông thường.
Việc biết tập giá trị của BNN là quan trọng. Tuy nhiên, hai BNN có tập giá trị giống nhau lại có thể hoàn toàn khác nhau. Tập giá trị của BNN cho ta rất ít thông tin về BNN. Điều quan trọng là biết BNN nhận các giá trị có thể của nó với xác suất tương ứng là bao nhiêu.

Luật phân bố của biên ngẫu nhiên

Mối quan hệ giữa các giá trị có thể của BNN với xác suất tương ứng được gọi là luật phân bố của biến ngẫu nhiên ấy.

Luật phân bố của BNN rời rạc

Giả sử $\{x_1, x_2, \dots\}$ là tập tất cả các giá trị của BNN rời rạc X.
Bộ số $p_1, p_2, \dots$ với $p_i = P(X = x_i), i = 1,2\dots$ được gọi là luật phân bố của BNN rời rạc X.

Rõ ràng rằng bộ số $p_1, p_2, \dots$ thoả mãn điều kiện: $p_i \geq 0$ và $\sum_i p_i = 1$ (tổng hữu hạn hay vô hạn).

Để thuận lợi, người ta sắp xếp bộ số $p_1, p_2, \dots$ thành bảng:

- dòng trên ghi các giá trị của BNN nhận (thường theo thứ tự tăng dần),
- dòng dưới ghi các xác suất tương ứng.

Bảng này gọi là bảng phân phối xác suất của BNN rời rạc X:

X	x_1	x_2	$\cdots$	x_n	$\cdots$
P	p_1	p_2	$\cdots$	p_n	$\cdots$

Khi đó, ta có các tính chất sau:

1. $0 \le p_i \le 1, \quad i = 1, 2, \ldots$.
2. $P(X \le x_i)$
$$= P\left[(X = x_1) + (X = x_2) + \ldots + (X = x_i)\right]$$
$$= P(X = x_1) + P(X = x_2) + \ldots + P(X = x_i)$$
$$= p_1 + p_2 + \cdots + p_i$$
$P(X \ge x_i)$
$$= P\left[(X = x_i) + (X = x_{i+1}) + \ldots + (X = x_n)\right]$$
$$= P(X = x_i) + P(X = x_{i+1}) + \ldots + P(X = x_n)$$
$$= p_i + p_{i+1} + \cdots + p_n$$
3. Nếu X nhận hữu hạn giá trị $\{x_1, x_2, \ldots, x_n\}$ thì
$$1 = \sum_{i=1}^{n} p_i = p_1 + p_2 + \cdots + p_n.$$
Nếu X nhận vô hạn đếm được các giá trị $\{x_1, x_2, \ldots\}$ thì
$$1 = \sum_{i=1}^{\infty} p_i = p_1 + p_2 + \cdots + p_n + \cdots$$
4. $$\mathrm{P}(a < X < b) = \sum_{a < x_i < b} \mathrm{P}(X = x_i)$$
$$\mathrm{P}(a \le X < b) = \sum_{a \le x_i < b} \mathrm{P}(X = x_i)$$
$$\mathrm{P}(a < X \le b) = \sum_{a < x_i \le b} \mathrm{P}(X = x_i).$$
$$\mathrm{P}(a \le X \le b) = \sum_{a \le x_i \le b} \mathrm{P}(X = x_i).$$
(chú ý **các giá trị biên**)

Ví dụ 2.4.

Lập bảng phân phối xác suất của số nút xuất hiện khi tung một cục xúc xắc.

Giải. Vì $p_i = P(X = i) = \frac{1}{6}, \forall i = \overline{1, 6}$ nên ta nhận được bảng:

X	1	2	3	4	5	6
P	1/6	1/6	1/6	1/6	1/6	1/6

Khi BNN RR X nhận n giá trị: $\{x_1, \ldots, x_n\}$ với các xác suất tương ứng $p_1 = \cdots = p_n = \frac{1}{n}$, X gọi là có phân bố đều rời rạc trên tập $\{x_1, \ldots, x_n\}$.

Luật phân bố này được sử dụng khi thiếu thông tin về X; khi ấy, ta coi các giá trị mà nó có thể nhận là **đồng khả năng**.

Ví dụ 2.5.

Tung một cục xúc xắc hai lần liên tiếp. Gọi biến ngẫu nhiên M là giá trị lớn nhất nhận được từ hai lần tung.

Ta nhận được bảng phân phối xác suất:

M	1	2	3	4	5	6
P	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

Ví dụ 2.6.

Hộp có 10 viên bi, trong đó có 6 viên màu đỏ, còn lại màu trắng. Rút đồng thời 4 viên bi và gọi X là số viên bi màu đỏ được rút ra. Lập bảng phân phối xác suất của X.

Giải.

Hàm mật độ xác suất

Giả sử biến ngẫu nhiên rời rạc X có bảng phân phối xác suất:

X	x_1	x_2	$\cdots$	x_n	$\cdots$
P	p_1	p_2	$\cdots$	p_n	$\cdots$

Khi đó, hàm mật độ xác suất (hay gọi tắt là hàm mật độ) của biến ngẫu nhiên X là một hàm số $f : \mathbb{R} \to [0;1]$

được xác định bởi: $$f_X(x) = \begin{cases} p_i = P(X = x_i) & \text{khi} \quad x = x_i, \quad i = 1,2,3,\ldots \\ 0 & \text{khi} \quad x \neq x_i, \quad \forall i \end{cases}$$

Chú ý: $\forall x \in \mathbb{R}, f_X(x) \geq 0$ và $\sum_i f_X(x_i) \sum_i p_i = 1$

Trong Ví dụ 3.3, ta có bảng phân phối xác suất:

X	0	1	2	3	4
P	0.004762	0.114286	0.428571	0.380952	0.071429

Hàm mật độ của biến ngẫu nhiên X là:

$$f_X(x) = \begin{cases} 0.004762, & \text{khi} \quad x = 0 \\ 0.114286, & \text{khi} \quad x = 1 \\ 0.428571, & \text{khi} \quad x = 2 \\ 0.380952, & \text{khi} \quad x = 3 \\ 0.071429, & \text{khi} \quad x = 4 \\ 0, & \text{khi} \quad x \notin \{0,1,2,3,4\} \end{cases}$$

Hàm phân phối xác suất (hàm phân phối tích lũy)

Giả sử biến ngẫu nhiên rời rạc X có bảng phân phối xác suất:

X	x_1	x_2	$\cdots$	x_n	$\cdots$
P	p_1	p_2	$\cdots$	p_n	$\cdots$

Khi đó, hàm phân phối xác suất (hay gọi tắt là hàm phân phối) của biến ngẫu nhiên X là một hàm số $F_X : \mathbb{R} \to [0;1]$ được xác định bởi:

$$F_X(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq x} f_X(x_i) = \sum_{x_i \leq x} p_i, \quad \forall x \in \mathbb{R}.$$

Tính chất:

1. $\forall k,\ F_X(x_k) = \sum_{x_i \leq x_k} P(X = x_i) = \sum_{x_i \leq x_k} f_X(x_i) = \sum_{i=1}^{k} p_i = p_1 + p_2 + \cdots + p_k.$
2. $\forall x \in \mathbb{R}, \quad 0 \leq F_X(x) \leq 1.$
3. $F_X(x)$ là một hàm không giảm: *nếu $x \leq y$ thì $F_X(x) \leq F_X(y)$. .*
4. $\lim_{x\to-\infty} F_X(x) = 0$ và $\lim_{x\to\infty} F_X(x) = 1$
5. $F_X(x)$ là hàm liên tục phải trên $\mathbb{R}$, nghĩa là $F_X(a) = \lim_{x\to a^+} F_X(x), \quad \forall a \in \mathbb{R}.$

Trong VD 2.5, ta có bảng phân phối xác suất:

M	1	2	3	4	5	6
P	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

Hàm phân phối tích lũy của biến ngẫu nhiên M là:

$$F_M(x) = \begin{cases} 0, & \text{khi} \quad x < 1 \\ \frac{1}{36}, & \text{khi} \quad 1 \le x < 2 \\ \frac{1}{36} + \frac{3}{36} = \frac{4}{36}, & \text{khi} \quad 2 \le x < 3 \\ \frac{1}{36} + \frac{3}{36} + \frac{5}{36} = \frac{9}{36}, & \text{khi} \quad 3 \le x < 4 \\ \frac{1}{36} + \frac{3}{36} + \frac{5}{36} + \frac{7}{36} = \frac{16}{36}, & \text{khi} \quad 4 \le x < 5 \\ \frac{1}{36} + \frac{3}{36} + \frac{5}{36} + \frac{7}{36} + \frac{9}{36} = \frac{25}{36}, & \text{khi} \quad 5 \le x < 6 \\ \frac{1}{36} + \frac{3}{36} + \frac{5}{36} + \frac{7}{36} + \frac{9}{36} + \frac{11}{36} = \frac{36}{36} = 1, & \text{khi} \quad x \ge 6 \end{cases}$$

Hàm mật độ và hàm phân phối của BNN liên tục

Hàm mật độ

$f_X(x)$ là hàm mật độ của BNN liên tục X nếu nó thỏa 2 điều kiện sau: $\begin{cases} f_X(x) \geq 0, & \forall x \in \mathbb{R} \\ \int_{-\infty}^{+\infty} f_X(x)\,\mathrm{d}x & = 1 \end{cases}$

Hàm phân phối xác suất

Hàm phân phối $F_X(x)$ của BNN liên tục X là hàm được xác định như sau: $F_X(x) = P(X \leq x), \quad \forall x \in \mathbb{R}$.

Nếu $f_X(x)$ là hàm mật độ của BNN liên tục X thì $F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t$.

$\forall a,b \in \mathbb{R} : a \leq b$, ta có: $P(a<X<b) = P(a<X\leq b) = P(a\leq X<b) = P(a\leq X\leq b) = \int_a^b f_X(x)\,\mathrm{d}x = F_X(b) - F_X(a)$.

Tính chất:

1. $\forall x \in \mathbb{R}, \quad 0 \leq F_X(x) \leq 1$.
2. $F_X(x)$ là một hàm không giảm trên $\mathbb{R}$: $\forall x_1, x_2 \in \mathbb{R}$, *nếu* $x_1 \leq x_2$ *thì* $F_X(x_1) \leq F_X(x_2)$.
3. $\lim_{x\to-\infty} F_X(x) = 0$ và $\lim_{x\to\infty} F_X(x) = 1$
4. $F_X(x)$ là hàm liên tục phải trên $\mathbb{R}$, nghĩa là $F_X(a) = \lim_{x\to a^+} F_X(x), \quad \forall a \in \mathbb{R}$.

Ví dụ 2.7.

Cho X là biến ngẫu nhiên liên tục có hàm mật độ là: $f(x) = \begin{cases} \dfrac{k}{x^2} & \text{nếu} \quad x \in [1,2] \\ 0 & \text{nếu} \quad x \notin [1,2] \end{cases}$, với k là một hằng số.

a) Tìm k

b) Tính $P(0 \leq X < 1.5)$

c) Tìm hàm phân phối $F_X(x)$.

Giải.

BNN rời rạc (hữu hạn giá trị)

1. BNN rời rạc X có bảng phân phối xác suất:

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

, trong đó

$p_i = P(X = x_i)$ và $\sum_{i=1}^{n} p_i = p_1 + p_2 + \cdots + p_n = 1$

2. Hàm mật độ xác suất của BNN rời rạc X:

$$f_X(x) = \begin{cases} p_i & \text{khi} \quad x = x_i, \quad i = 1,2,3,\ldots \\ 0 & \text{khi} \quad x \neq x_i, \quad \forall i \end{cases}$$

3. Hàm phân phối xác suất (hàm phân phối tích lũy)

$$F_X(x)=P(X\le x)=\begin{cases} 0, & \text{khi} \quad x<x_1 \\ p_1, & \text{khi} \quad x_1\le x<x_2 \\ p_1+p_2, & \text{khi} \quad x_2\le x<x_3 \\ \vdots & \\ p_1+p_2+\cdots+p_k, & \text{khi} \quad x_k\le x<x_{k+1} \\ \vdots & \\ p_1+p_2+\cdots+p_n=1, & \text{khi} \quad x\ge x_n \end{cases}$$

4. $p_k = P(X=x_k) = f_X(x_k) = F_X(x_k) - F_X(x_{k-1})$.

BNN liên tục

1. Thông thường, ta sử dụng:
   $P(X = a) = 0, \quad \forall a \in \mathbb{R}$.

2. $f_X(x)$ là hàm mật độ xác suất của BNN liên tục X:

$$\begin{cases} f_X(x) \ge 0, & \forall x \in \mathbb{R} \\ \int_{-\infty}^{+\infty} f_X(x)\,\mathrm{d}x & = 1 \end{cases}$$

3. Hàm phân phối của BNN liên tục X:

$$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t, \quad \forall x \in \mathbb{R}.$$

4. $f_X(x) = F_X'(x)$.

Hàm của các biến ngẫu nhiên

Trường hợp $Y = \varphi(X)$, với X là một BNN rời rạc, hữu hạn giá trị

Cho bảng phân phối xác suất của BNN rời rạc X (với hữu hạn giá trị):

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

Ta cần tìm bảng phân phối xác suất của $Y = \varphi(X)$.

Bước 1: Tính các giá trị cho biến Y:

X	x_1	x_2	$\cdots$	x_n
$Y = \varphi(X)$	$\varphi(x_1)$	$\varphi(x_2)$	$\cdots$	$\varphi(x_n)$

Bước 2: Do bảng các giá trị của Y được tính ở Bước 1 có thể trùng lặp nên ta cần tính xác suất tương ứng cho từng giá trị của Y:

$$P(Y = y_k) = \sum_{\varphi(x_i)=y_k} P(X = x_i) = \sum_{\varphi(x_i)=y_k} p_i$$

Sau đó, ta lập bảng phân phối xác suất của Y.

Ví dụ 2.8.

Cho biến ngẫu nhiên X có bảng phân phối xác suất như sau:

X	0	1	2	3
P	0.1	0.3	0.4	0.2

.

Hãy lập bảng phân phối xác suất của $Y = X^2 - 4X + 5$.

Giải.

Bài tập 2.1.

Hãy lập bảng phân phối xác suất của $Y = 2X^2 - 3X + 5$, biết X là một biến ngẫu nhiên rời rạc có bảng phân phối xác suất như sau:

X	1	3	5	6
P	0.1	0.4	0.3	0.2

.

Biến ngẫu nhiên n chiều

Biến ngẫu nhiên n chiều là một bộ gồm n biến ngẫu nhiên. Ký hiệu: $V = (X_1, X_2, \ldots, X_n)$, trong đó $X_1, X_2, \ldots, X_n$ là các BNN.

Ví dụ 2.9.

Người ta nghiên cứu về doanh thu và chi phí quảng cáo của một công ty tư nhân. Gọi X là doanh thu và Y là chi phí quảng cáo thì $V = (X, Y)$ tạo nên một biến ngẫu nhiên 2 chiều.

Ví dụ 2.10.

Một nhà máy chuyên sản xuất container. Nếu kích thước của các container được đo bằng chiều dài X, chiều rộng Y và chiều cao Z thì ta có biến ngẫu nhiên 3 chiều: $W = (X, Y, Z)$.

Chú ý:

- Nếu tất cả $X_1, X_2, X_3, \ldots, X_n$ đều là *BNN* rời rạc thì $\mathrm{V} = (\mathrm{X}_1, \mathrm{X}_2, \mathrm{X}_3, \ldots, \mathrm{X}_n)$ là BNN rời rạc.
- Nếu tất cả $X_1, X_2, X_3, \ldots, X_n$ đều là BNN liên tục thì $\mathrm{V} = (\mathrm{X}_1, \mathrm{X}_2, \mathrm{X}_3, \ldots, \mathrm{X}_n)$ là BNN liên tục.
- Trong học phần này, ta không xét trường hợp vừa có thành phần rời rạc vừa có thành phần liên tục.

Bảng phân phối xác suất của BNN 2 chiều rời rạc (X,Y), với hữu hạn giá trị

Giả sử X, Y là các BNN rời rạc, trong đó:

- ◇ X nhận các giá trị là $x_1 < x_2 < \cdots < x_m$;
- ◇ Y nhận các giá trị là $y_1 < y_2 < \cdots < y_n$.

• Cặp đại lượng ngẫu nhiên rời rạc được xét đồng thời (X,Y) được gọi là vector ngẫu nhiên rời rạc.

• Ký hiệu biến cố: $(X \le x)\cdot(Y \le y) := (X \le x;\ Y \le y),\ \forall x,y \in \mathbb{R}$.

• Hàm phân phối xác suất đồng thời của X và Y là: $F(x,y) = P(X \le x; Y \le y), \forall x,y \in \mathbb{R}$.

• X và Y được gọi là độc lập nếu
$F(x,y) = F_X(x)\cdot F_Y(y),\ \forall x,y \in \mathbb{R}$.

Nếu X,Y độc lập thì hàm phân phối xác suất đồng thời của (X,Y) được xác định **dễ dàng** qua các hàm phân phối của X và Y.

Bảng phân phối xác đồng thời của (X,Y)

X \ Y	y_1	y_2	$\ldots$	y_j	$\ldots$	y_n	P_X
x_1	p_{11}	p_{12}	$\ldots$	p_{1j}	$\ldots$	p_{1n}	p_1
x_2	p_{21}	p_{22}	$\ldots$	p_{2j}	$\ldots$	p_{2n}	p_2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_i	p_{i1}	p_{i2}	$\ldots$	p_{ij}	$\ldots$	p_{in}	p_i
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_m	p_{m1}	p_{m2}	$\ldots$	p_{mj}	$\ldots$	p_{mn}	p_m
P_Y	q_1	q_2	$\ldots$	q_j	$\ldots$	q_n	**1**

$\forall i \in \{1,\ldots m\}, \forall j \in \{1,\ldots,n\}$, $p_{i,j} = P(X = x_i, Y = y_j)$ là xác suất để $X = x_i$ và $Y = y_j$. Khi đó,

• $p_i := p_{i1} + \cdots + p_{in} = \sum_{j=1}^{n} p_{ij}$;

• $q_j := p_{1j} + \cdots + p_{mj} = \sum_{i=1}^{m} p_{ij}$;

• $1 = \sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij} = \sum_{i=1}^{m} p_i$
$= \sum_{j=1}^{n}\sum_{i=1}^{m} p_{ij} = \sum_{j=1}^{n} q_j$

Phân phối xác suất biên (lề)

Từ bảng phân phối xác suất đồng thời của (X,Y), ta có:

- Phân phối xác suất biên (lề) của X:

X	x_1	x_2	...	x_i	...	x_m
P_X	p_1	p_2	$\cdots$	p_i	...	p_m

,

trong đó $\quad p_i = P(X = x_i) = \sum_{j=1}^{n} P(X = x_i, Y = Y_j) = p + p_{i2} + p_{i3} + \cdots + p_{in} = \sum_{j=1}^{n} p_{ij}.$

- Phân phối xác suất biên (lề) của Y:

Y	y_1	y_2	...	y_j	...	y_n
P_Y	q_1	q_2	$\cdots$	q_j	...	q_n

,

trong đó $\quad q_j = P(Y = y_j) = \sum_{i=1}^{m} P(X = x_i, Y = Y_j) = p_{1j} + p_{2j} + p_{3j} + \cdots + p_{mj} = \sum_{i=1}^{m} p_{ij}.$

Tính chất

X và Y độc lập khi và chỉ khi $p_{ij} = p_i \cdot q_j, \forall i,j.$ $\quad P(X = x_i, Y = y_i) = P(X = x_i) \cdot P(Y = y_j), \forall i,j$

Hàm mật độ xác suất đồng thời của $V = (X,Y)$

Hàm mật độ xác suất đồng thời của $V = (X,Y)$ là: $f(x,y) = \begin{cases} p_{ij} & \text{khi} \quad (x,y) = (x_i,y_j) \\ 0 & \text{khi} \quad (x,y) \neq (x_i,y_j), \forall i,j \end{cases}$

Phân phối có điều kiện

- Bảng phân phối có điều kiện của X khi $Y = y_j$ với $(P(Y = y_j) = q_j \neq 0)$:

  Ta luôn có: $P(X = x_i \mid Y = y_j) = \dfrac{P(X = x_i; Y = y_j)}{P(Y = y_j)} = \dfrac{p_{ij}}{q_j}$. Do đó,

X	x_1	x_2	$\dots$	x_i	$\dots$	x_m
$P(X \mid Y = y_j)$	$\dfrac{p_{1j}}{q_j}$	$\dfrac{p_{2j}}{q_j}$	$\dots$	$\dfrac{p_{ij}}{q_j}$	$\dots$	$\dfrac{p_{mj}}{q_j}$

- Bảng phân phối có điều kiện của Y khi $X = x_i$ với $(P(X = x_i) = p_i \neq 0)$:

  Ta luôn có: $P(Y = y_j \mid X = x_i) = \dfrac{P(X = x_i; Y = y_j)}{P(X = x_i)} = \dfrac{p_{ij}}{p_i}$. Do đó,

X	y_1	y_2	$\dots$	y_j	$\dots$	y_n
$P(Y \mid X = x_i)$	$\dfrac{p_{i1}}{p_i}$	$\dfrac{p_{i2}}{p_i}$	$\dots$	$\dfrac{p_{ij}}{p_i}$	$\dots$	$\dfrac{p_{in}}{p_i}$

Ví dụ 2.11.

Cho bảng phân phối xác suất đồng thời của X và Y :

$X \backslash Y$	10	20	30	40
10	0,2	0,04	0,01	0
20	0,1	0,36	0,09	0
30	0	0,05	0,1	0
40	0	0	0	0,05

a) Tìm phân phối biên của X, của Y.

b) Xét xem X và Y có độc lập không ?

c) Hãy lập bảng phân phối có điều kiện của X khi $Y = 30$.

Giải.

Ví dụ 2.12.

Cho hai biến ngẫu nhiên rời rạc độc lập có bảng phân phối xác suất như sau:

X	1	2	3
P	$1/3$	$1/2$	$1/6$

Y	-1	1
P	$1/3$	$2/3$

Hãy lập bảng phân phối đồng thời của X và Y.

Giải.

Chú ý. Sinh viên tự nghiên cứu về biến ngẫu nhiên nhiều chiều liên tục

Hàm của các biến ngẫu nhiên

Trường hợp $Z = \varphi(X,Y)$, với X,Y là các BNN rời rạc, hữu hạn giá trị

Cho bảng phân phối xác suất đồng thời của (X,Y). Ta cần tìm bảng phân phối xác suất của $Z = \varphi(X,Y)$.

Bước 1: Tính các giá trị cho biến Z:

Z \ Y / X	y_1	y_2	$\dots$	y_j	$\dots$	y_n
x_1	$z_{11} = \varphi(x_1,y_1)$	$z_{12} = \varphi(x_1,y_2)$	$\dots$	$z_{1j} = \varphi(x_1,y_j)$	$\dots$	$z_{1n} = \varphi(x_1,y_n)$
x_2	$z_{21} = \varphi(x_2,y_1)$	$z_{22} = \varphi(x_2,y_2)$	$\dots$	$z_{2j} = \varphi(x_2,y_j)$	$\dots$	$z_{2n} = \varphi(x_2,y_n)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
x_i	$z_{i1} = \varphi(x_i,y_1)$	$z_{i2} = \varphi(x_i,y_2)$	$\dots$	$z_{ij} = \varphi(x_i,y_j)$	$\dots$	$z_{in} = \varphi(x_i,y_n)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
x_m	$z_{m1} = \varphi(x_m,y_1)$	$z_{m2} = \varphi(x_m,y_2)$	$\dots$	$z_{mj} = \varphi(x_m,y_j)$	$\dots$	$z_{mn} = \varphi(x_m,y_n)$

Bước 2: Do bảng các giá trị của Z được tính ở Bước 1 có thể trùng lặp nên ta cần tính xác suất tương ứng cho từng giá trị của Z:

$$P(Z = z_k) = \sum_{\varphi(x_i,y_j)=z_k} P(X = x_i, Y = y_j) = \sum_{\varphi(x_i,y_j)=z_k} p_{ij}$$

Sau đó, ta lập bảng phân phối xác suất của Z.

Ví dụ 2.13.

Cho bảng phân phối xác suất đồng thời:

X \ Y	-1	0	1
1	0,1	0,15	0,05
2	0,3	0,2	0,2

Hãy lập bảng phân phối xác suất của $Z = \varphi(X, Y) = 2X - Y + 5$.

Giải.

Bài tập 2.2.

X \ Y	-1	0	1
0	0,1	0,2	0,3
1	0,2	0,1	0,1

Bên trái là bảng phân phối xác suất đồng thời của X và Y. Hãy tìm bảng phân phối xác suất của $Z = X - Y + 1$.

Bài tập 2.3.

Cho X và Y là hai biến ngẫu nhiên có bảng phân phối xác suất đồng thời:

X \ Y	-1	0	1	2
1	0,10	0,00	0,10	0,20
2	0,00	0,20	0,10	0,05
3	0,05	0,05	0,00	0,15

a) Lập các bảng phânphối xác suất của X và của Y.

b) Xét tính độc lập của X và Y.

c) Lập bảng phân phối XS của $W = Y^2 - 2Y + 3$.

d) Lập bảng phân phối có điều kiện của X khi $Y = 2$.

e) Tính $P(X + Y \geq 3{,}5)$ và $P(Y = 1 \mid X = 1)$.

Bài tập 2.4.

Cho X và Y là hai biến ngẫu nhiên có bảng phân phối xác suất đồng thời:

X \ Y	1	2	3	4
3	0,05	0,15	0,1	0,2
4	0,02	0,06	0,04	0,08
5	0,03	0,09	0,06	0,12

a) Lập các bảng phânphối xác suất của X và của Y.

b) Xét tính độc lập của X và Y.

c) Lập bảng phân phối xác suất của $Z = 2X - Y$

d) Lập bảng phân phối xác suất có điều kiện của Y khi $X = 3$.

e) Tính $P(X + Y \geq 5{,}5)$ và $P(X = 5 \mid Y = 4)$.

Kỳ vọng (mean; expected value, or expectation)

Định nghĩa 2.14.

Kỳ vọng của biến ngẫu nhiên X, ký hiệu là $E[X]$ (hay μ_X), được xác định như sau:

- Nếu X là BNN rời rạc

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

, thì

$$E[X] = x_1 \cdot p_1 + x_2 \cdot p_2 + \cdots + x_n \cdot p_n = \sum_{i=1}^{n} x_i \cdot p_i = \sum_{i=1}^{n} x_i \cdot f(x_i)$$

- Nếu X là BNN liên tục với hàm mật độ $f_X(x)$ thì

$$E[X] = \int_{-\infty}^{+\infty} x \cdot f_X(x)\,\mathrm{d}x$$

Kỳ vọng $E[X]$ là giá trị trung bình (theo xác suất) mà X nhận được, nó phản ánh **giá trị trung tâm** của phân phối xác suất của X.

Ví dụ

1. Cho BNN rời rạc X có bảng phân phối xác suất:

X	1	3	4
P	0,1	0,5	0,4

Tính kỳ vọng của BNN X:

$E[X] = 1 \times 0{,}1 + 3 \times 0{,}5 + 4 \times 0{,}4 = 3{,}2.$

2. Cho BNN liên tục Y có hàm mật độ

$$f(x) = \begin{cases} \frac{3}{4}\left(x^2 + 2x\right), & \text{với} \quad x \in (0;1) \\ 0, & \text{với} \quad x \notin (0;1) \end{cases}$$

$$E[Y] = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 x \frac{3}{4}\left(x^2 + 2x\right) dx = \frac{11}{16}.$$

Tính chất của kỳ vọng

1. $E[C] = C$, với mọi hằng số C.
2. Với mọi BNN X, với mọi $a \in \mathbb{R}$, $E[aX] = a\,E[X]$
3. Nếu X và Y là những BNN có kỳ vọng thì BNN tổng $(X+Y)$ cũng có kỳ vọng và $E[X+Y] = E[X] + E[Y]$
4. Nếu các BNN X và Y độc lập và có kỳ vọng thì BNN tích $X \cdot Y$ cũng có kỳ vọng và $E[X \cdot Y] = E[X] \cdot E[Y]$
5. Giả sử $\varphi(\mathrm{x})$ là hàm số thông thường nào đó sao cho $\mathrm{Y} = \varphi(\mathrm{X})$ là BNN có kỳ vọng.

- X rời rạc nhận các giá trị $\{x_i \mid i \in \mathbb{Z}\}$:
$$E[Y] = \sum_{i=-\infty}^{+\infty} \varphi(x_i)\, p_i = \sum_{i=-\infty}^{+\infty} \varphi(x_i)\, f_X(x_i);$$
Nếu X chỉ nhận các giá trị $\{x_1, \ldots, x_n\}$ thì
$$E[Y] = \sum_{i=1}^{n} \varphi(x_i)\, p_i = \sum_{i=1}^{n} \varphi(x_i)\, f_X(x_i)$$
- X liên tục có hàm mật độ là $f_X(x)$:
$$E[Y] = \int_{-\infty}^{+\infty} \varphi(x)\, f_X(x)\, \mathrm{d}x$$

Ví dụ 2.15.

Tính $E[X^2]$ với X là BNN có bảng phân phối xác suất

X	-3	1	3	5	6
P	0,1	0,2	0,3	0,2	0,2

.

Giải.

Ví dụ 2.16.

Một thùng đựng 10 quả quả dưa khác nhau về trọng lượng: 5 quả nặng 3kg, 2 quả nặng 4kg, 3 quả nặng 5kg. Lấy ngẫu nhiên từ thùng ra 1 quả. Tìm trọng lượng trung bình của một quả dưa.

Giải.

Ví dụ 2.17.

Cho biến ngẫu nhiên X có hàm mật độ: $f_X(x) = \begin{cases} \dfrac{2}{x^2} & \text{khi} \quad x \in [1,2] \\ 0 & \text{khi} \quad x \notin [1,2] \end{cases}$.

a) Tìm $E[X]$.

b) Tính kỳ vọng của Y với $Y = X^2 - 2X$.

Giải.

Mode

Định nghĩa 2.18.

Mode của biến ngẫu nhiên X, ký hiệu là $Mod[X]$, là giá trị mà tại đó xác suất tương ứng đạt được giá trị lớn nhất:

- Nếu X là BNN RR $\begin{array}{c|cccc} X & x_1 & x_2 & \cdots & x_n \\ \hline P & p_1 & p_2 & \cdots & p_n \end{array}$, thì

$$x_k = Mod[X] \Leftrightarrow P(X = x_k) = p_k = \max\{p_1, \ldots, p_n\}$$
$$\Leftrightarrow f_X(x_k) = \max_{x\in\mathbb{R}} f_X(x).$$

- Nếu X là BNN liên tục với hàm mật độ $f_X(x)$ thì $x_0 = Mod[X] \Leftrightarrow f_X(x_0) = \max_{x\in\mathbb{R}} f_X(x)$.

Chú ý. $Mod[X]$ có thể nhận nhiều giá trị khác nhau. Nếu chỉ có đúng 1 mode thì BNN X được gọi là đơn mode.

Ví dụ

1. Cho BNN rời rạc X có bảng phân phối xác suất:

X	1	3	4
P	0,1	0,5	0,4

Tính Mode của BNN X: $Mod[X] = 3$ vì $P(X = 3) = 0{,}5 = \max\{0{,}1;\ 0{,}5;\ 0{,}4\}$.

2. Cho BNN liên tục Y có hàm mật độ $f_Y(x) = \begin{cases} \dfrac{3}{4}\left(x^2 + 2x\right), & \text{với} \quad x \in [0;1] \\ 0, & \text{với} \quad x \notin [0;1] \end{cases}$

Tính Mode của BNN Y: $Mod[Y] = 1$ vì $f_Y(1) = \max_{x\in\mathbb{R}} f_Y(x)$.

Median - trung vị

Định nghĩa 2.19.

Trung vị của biến ngẫu nhiên X, ký hiệu là $Med[X]$, được xác định bởi:

- Nếu X là BNN RR

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

, thì

$$x_k = Med[X] \Leftrightarrow \begin{cases} P(X \le x_k) & \ge & 1/2 \\ P(X \ge x_k) & \ge & 1/2 \end{cases}.$$

Chú ý. $Med[X]$ có thể nhận nhiều giá trị khác nhau.

- Nếu X là BNN liên tục với hàm mật độ $f_X(x)$ thì

$$x_0 = Med[X] \Leftrightarrow \begin{cases} \displaystyle\int_{-\infty}^{x_0} f_X(x)\,\mathrm{d}x \ge 1/2 \\ \displaystyle\int_{x_0}^{+\infty} f_X(x)\,\mathrm{d}x \ge 1/2 \end{cases}.$$

Ví dụ

1. Cho BNN rời rạc X có bảng phân phối xác suất:

X	1	3	4
P	0,1	0,5	0,4

. Tính Med của BNN X:

Do $P(X \le 3) = 0{,}6 \ge 0{,}5$ và $P(X \ge 3) = 0{,}9 \ge 0.5$ nên $Med[X] = 3$.

2. Cho BNN rời rạc Y có bảng phân phối xác suất:

Y	1	3	4	5
P	0,1	0,4	0,3	0,2

. Tính Med của BNN Y:

$$\text{Do} \begin{cases} P(X \le 3) = 0{,}5 \ge 0{,}5 \quad \text{và} \quad P(X \ge 3) = 0{,}9 \ge 0.5 \\ P(X \le 4) = 0{,}8 \ge 0{,}5 \quad \text{và} \quad P(X \ge 4) = 0{,}5 \ge 0.5 \end{cases}$$

nên $Med[X]$ nhận hai giá trị: 3 và 4.

Nhận xét

- Khi X là BNN RR:

$$x_k = Med[X] \quad \Leftrightarrow \quad F_X(x_{k-1}) = P(X < x_k) = 1 - P(X \geq x_k) \leq 0,5 \leq P(X \leq x_k) = F_X(x_k).$$

- Khi X là BNN liên tục với hàm mật độ $f_X(x)$: $\quad x_0 = Med[X] \quad \Leftrightarrow \quad F_X(x_0) = \int_{-\infty}^{x_0} f_X(x)\,\mathrm{d}x = \frac{1}{2}.$

Ví dụ 2.20.

a) Cho BNN liên tục X có hàm mật độ

$$f_X(x) = \begin{cases} 2x, & \text{với} \quad x \in [0;1] \\ 0, & \text{với} \quad x \notin [0;1] \end{cases}$$

Do $F_X\left(\frac{1}{\sqrt{2}}\right) = \int_{-\infty}^{1/\sqrt{2}} f_X(x)\,\mathrm{d}x = \left[x^2\right]_0^{1/\sqrt{2}} = \frac{1}{2}$

và $\frac{1}{\sqrt{2}}$ là giá trị duy nhất thỏa điều này nên

$Med[X] = \frac{1}{\sqrt{2}}$.

b) Cho BNN liên tục Y có hàm mật độ

$$f_Y(x) = \begin{cases} 0.5, & \text{nếu} \quad 1 \leq |x| \leq 2 \\ 0, & \text{ngược lại} \end{cases}$$

Với mọi $a \in [-1;1]$, ta có: $F_Y(a) = \int_{-\infty}^{a} f_Y(x)\,\mathrm{d}x$

$$= \int_{-\infty}^{-2} 0\,\mathrm{d}x + \int_{-2}^{-1} 0.5\,\mathrm{d}x + \int_{-1}^{a} 0\,\mathrm{d}x = 0.5$$

Do đó, $Med[Y]$ nhận tất cả các giá trị trong đoạn $[-1;1]$.

Phương sai (Variance)

Phương sai của biến ngẫu nhiên X, ký hiệu là $Var[X]$, được xác định bởi: $Var[X] = E\left[\left(X - E[X]\right)^2\right]$.

Chú ý. Phương sai chính là trung bình của bình phương độ lệch giữa giá trị X so với giá trị trung bình $E[X]$. Nó thể hiên mức độ phân tán của BNN X quanh giá trị kỳ vọng $E[X]$. Hơn nữa, ta có:

$$\begin{aligned} Var[X] &= E\left[\left(X - E[X]\right)^2\right] = E\left[X^2 - 2X \cdot E[X] + (E[X])^2\right] \\ &= E[X^2] - 2E\left[X \cdot E[X]\right] + (E[X])^2 \\ &= E[X^2] - 2E[X] \cdot E[X] + (E[X])^2 = E\left[X^2\right] - \left(E[X]\right)^2 \end{aligned}$$

- Khi X là BNN RR

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

, thì

$$Var[X] = E\left[X^2\right] - \left(E[X]\right)^2 = \sum_{i=1}^{n} x_i^2 p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2.$$

- Khi X là BNN liên tục với hàm mật độ $f_X(x)$ thì

$$Var[X] = E\left[X^2\right] - \left(E[X]\right)^2 = \int_{-\infty}^{+\infty} x^2 f_X(x)\,\mathrm{d}x - \left(\int_{-\infty}^{+\infty} x f_X(x)\,\mathrm{d}x\right)^2$$

Độ lệch chuẩn (standard deviation)

Căn bậc hai của phương sai của BNN X được gọi là độ lệch chuẩn của BNN X, ký hiệu là σ_X:

$$\sigma_X = \sqrt{Var[X]} = \sqrt{\sigma_X^2}.$$

Tính chất

Cho các BNN X,Y và a,C là các hằng số.

1. $Var[C] = E[C^2] - (E[C])^2 = C^2 - C^2 = 0;$
2. $Var[a\,X] = E[(aX)^2] - (E[aX])^2$
   $= a^2 E[X^2] - (a\,E[X])^2 = a^2\,Var[X];$
3. $Var[X+C] = E[(X+C)^2] - (E[X+C])^2$
   $= E[X^2 + 2XC + C^2] - (E[X]+C)^2$
   $= E[X^2] + 2E[XC] + E[C^2] - \left((E[X])^2 + 2E[X]C + C^2\right)$
   $= E[X^2] - (E[X])^2 = Var[X];$
4. Nếu X,Y là **độc lập** và có phương sai hữu hạn thì $(X+Y)$ cũng là BNN có phương sai hữu hạn và $Var[X+Y] = Var[X] + Var[Y]$. Thật vậy,

$Var[X+Y] = E[(X+Y)^2] - (E[X+Y])^2$
$= E[X^2 + 2XY + E^2] - (E[X]+E[Y])^2$
$= E[X^2] + 2E[XY] + E[Y^2] - \left((E[X])^2 + 2E[X]E[Y] + (E[Y])^2\right)$
$= E[X^2] + 2E[X]E[Y] + E[Y^2] - (E[X])^2 - 2E[X]E[Y] - (E[Y])^2$
(do X,Y **độc lập**)
$= E[X^2] - (E[X])^2 + E[Y^2] - (E[Y])^2 = Var[X] + Var[Y];$

Hệ quả

a) Nếu X,Y là **độc lập** và có phương sai hữu hạn thì $Var[X-Y] = Var[X] + Var[Y]$.

b) Nếu $X_1,\dots,X_n$ là các BNN **độc lập** và có phương sai hữu hạn thì

$$Var[X_1 + \cdots + X_n] = Var[X_1] + \cdots + Var[X_n].$$

Ví dụ 2.21.

Tính $Var[X]$ với X là BNN RR có bảng phân phối xác suất

X	-3	1	3	5	6
P	0,1	0,2	0,3	0,2	0,2

.

Giải.

Ví dụ 2.22.

Tính $Var[X]$ với X là BNN liên tục có hàm mật độ $f_X(x) = \begin{cases} 2x, & \text{với} \quad x \in [0;1] \\ 0, & \text{với} \quad x \notin [0;1] \end{cases}$

Giải.

Kì vọng của biến ngẫu nhiên 2 chiều

Cho biến ngẫu nhiên 2 chiều $V=(X,Y)$. Kì vọng của V là: $E[V]=\big(E[X],E[Y]\big)\in\mathbb{R}^2$.

Trường hợp (X,Y) rời rạc: Kỳ vọng của hàm 2 biến ngẫu nhiên $Z=\varphi(X,Y)$, kỳ vọng có điều kiện

Giả sử X, Y là các BNN rời rạc, trong đó: $\begin{cases} X \text{ nhận các giá trị là } x_1,x_2,\ldots,x_m, \text{ với } x_1<x_2<\cdots<x_m; \\ Y \text{ nhận các giá trị là } y_1,y_2,\ldots,y_n, \text{ với } y_1<y_2<\cdots<y_n. \end{cases}$

- Kỳ vọng của Z: $E[Z]=E[\varphi(X,Y)]=\sum_{i=1}^{m}\sum_{j=1}^{n}\varphi(x_i,y_j)P(X=x_i;Y=y_j)=\sum_{i=1}^{m}\sum_{j=1}^{n}\varphi(x_i,y_j)p_{ij}.$
- Kỳ vọng của BNN X với điều kiện $(Y=y_j)$ khi $P(Y=y_j)=q_j\neq 0$:

$$E[X\mid Y=y_j]=\sum_{i=1}^{m}x_i\frac{P(X=x_i;Y=y_j)}{P(Y=y_j)}=\sum_{i=1}^{m}x_i\frac{p_{ij}}{q_j}.$$

- Kỳ vọng của BNN Y với điều kiện $(X=x_i)$ khi $P(X=x_i)=p_i\neq 0$:

$$E[Y\mid X=x_i]=\sum_{j=1}^{n}y_j\frac{P(X=x_i;Y=y_j)}{P(X=x_i)}=\sum_{j=1}^{n}y_j\frac{p_{ij}}{p_i}.$$

Hiệp phương sai và hệ số tương quan

Biến ngẫu nhiên X có kỳ vọng là $E[X]=\mu_X$, biến ngẫu nhiên Y có kỳ vọng là $E[Y]=\mu_Y$.

- **Hiệp phương sai (Covariance)** giữa hai biến ngẫu nhiên X và Y được định nghĩa như sau:

$$Cov(X,Y) = E\left[(X-\mu_X)\cdot(Y-\mu_Y)\right] = E\left[X\cdot Y-\mu_X\cdot Y-\mu_Y\cdot X+\mu_X\mu_Y\right]$$
$$= E[X\cdot Y]-\mu_x E[Y]-\mu_Y E[X]+\mu_X\mu_Y = E[X\cdot Y]-\mu_X\mu_Y = E[X\cdot Y]-E[X]\cdot E[Y]$$

Khi X, Y là các BNN rời rạc, trong đó: $\begin{cases} X \text{ nhận các giá trị là } x_1,x_2,\dots,x_m, \text{ với } x_1<x_2<\cdots<x_m; \\ Y \text{ nhận các giá trị là } y_1,y_2,\dots,y_n, \text{ với } y_1<y_2<\cdots<y_n. \end{cases}$

$$Cov(X,Y) = E[X\cdot Y]-E[X]\cdot E[Y] = \sum_{i=1}^{m}\sum_{j=1}^{n} x_i\, y_j\, p_{ij} - \left(\sum_{i=1}^{m} x_i p_i\right)\left(\sum_{j=1}^{n} y_j q_j\right)$$

Chú ý. Nếu X và Y độc lập thì $Cov(X,Y)=E[X\cdot Y]-E[X]\cdot E[Y]=E[X]\cdot E[Y]-E[X]\cdot E[Y]=0$.
Do đó, nếu $Cov(X,Y)\neq 0$ thì X và Y không độc lập.

- **Hệ số tương quan** (Pearson) của hai biến X và Y: $\rho_{XY}=\dfrac{Cov(X,Y)}{\sigma_X\ \sigma_Y}$, trong đó σ_X, σ_Y lần lượt là độ lệch chuẩn của BNN X và Y.

Tính chất

1. $|\rho_{XY}| \leq 1.$
2. Với $a,b \in \mathbb{R}$, ta có:

$$
\begin{aligned}
Var[aX+bY] &= E\left[(aX+bY)^2\right] - \left(E[aX+bY]\right)^2 \\
&= E\left[a^2X^2 + 2abXY + b^2Y^2\right] - \left(aE[X]+bE[Y]\right)^2 \\
&= a^2E[X^2] + 2abE[XY] + b^2E[Y^2] - a^2\left(E[X]\right)^2 - 2ab\,E[X]\,E[Y] - b^2\left(E[Y]\right)^2 \\
&= a^2\left(E[X^2] - (E[X]^2)\right) + b^2\left(E[Y^2] - (E[Y])^2\right) + 2ab\left(E[XY] - E[X]\,E[y]\right) \\
&= a^2\,Var[X] + b^2\,Var[Y] + 2ab\,Cov(X,Y).
\end{aligned}
$$

Bài tập 2.5.

Có 3 lô hàng, mỗi lô hàng có 20 lọ thuốc. Số lọ thuốc tốt của mỗi lô hàng lần lượt là 5; 18; 10.

a) Lấy ngẫu nhiên mỗi lô 1 lọ. Gọi X là số lọ thuốc tốt trong 3 lọ lấy ra. Lập bảng phân phối xác suất cho X và tính $E[X], Var[X], Mod[X]$

b) Chọn ngẫu nhiên 1 lô, rồi từ lô đó lấy ngẫu nhiên 3 lọ. Gọi Y là số lọ thuốc tốt trong 3 lọ lấy ra. Lập bảng phân phối xác suất cho Y và tính X ? Tính $E[Y], Var[Y], Mod[Y]$

Bài tập 2.6.

Tuổi thọ của một loại côn trùng (đơn vị: tháng) nào đó là một biến ngẫu nhiên X có hàm mật độ:

$$f(x) = \begin{cases} ax^2(4-x) & \text{khi } x \in [0;4] \\ 0 & \text{khi } x \notin [0;4] \end{cases}$$

a) Xác định a.

b) Tìm hàm phân phối của X.

c) Tính $E[X], Var[X]$

d) Tìm $Mod[X], Med[X]$

e) Tính xác suất để côn trùng chết trước khi nó được 1,5 tháng tuổi.

f) Tính xác suất để tuổi thọ côn trùng sống ít nhất 1,25 tháng tuổi.

g) Tính xác suất để côn trùng sống từ 1,25 đến 2,75 tháng tuổi.

Bài tập 2.7.

Xét trò chơi, tung 1 xúc xắc 3 lần:

◇ nếu 3 lần xuất hiện mặt 1 nút thì người chơi được 6000đ,

◇ nếu 2 lần xuất hiện mặt 1 nút thì người chơi được 4000đ,

◇ nếu 1 lần xuất hiện mặt 1 nút thì người chơi được 2000đ.

1. Mỗi lần chơi, người chơi phải đóng 2000đ. Hỏi trung bình một lần chơi, người chơi lời (lỗ) bao nhiêu tiền? Nếu người chơi chơi 1000 lần, thì trung bình người chơi lời (lỗ) bao nhiêu tiền?
2. Mỗi lần chơi, người chơi phải đóng m đồng.
   i) Hỏi m bằng bao nhiêu thì lâu dài người chơi huề vốn?
   ii) Hỏi m bằng bao nhiêu thì trung bình mỗi lần chơi, người chơi mất 2000 đ?

Bài tập 2.8.

Có ba hộp thuốc: Hộp I có 10 lô thuốc trong đó có 2 lọ hỏng; Hộp II có 20 lọ thuốc trong đó có 8 lọ hỏng; Hộp III có 15 lọ thuốc trong đó có 5 lọ hỏng.

a) Lấy ngẫu nhiên mỗi hộp 1 lọ. Gọi X là số lọ hỏng trong 3 lọ được lấy.
   i) Lập bảng phân phối xác suất của X
   ii) Tính trung bình và phương sai của X

b) Lấy ngẫu nhiên 3 lọ ở hộp I. Gọi Y là số lọ hỏng trong 3 lọ được lấy.
   i) Lập bảng phân phối xác suất của Y
   ii) Tính trung bình và phương sai của Y

c) Lấy ngẫu nhiên 1 hộp; từ hộp đó lấy ngẫu nhiên 2 lọ. Gọi Z là số lọ hỏng trong 2 lọ được lấy.
   i) Lập bảng phân phối xác suất của Z
   ii) Tính trung bình và phương sai của Z

d) Lấy ngẫu nhiên từng lọ trong hộp I để kiểm tra (không hoàn lại) cho đến khi gặp lọ hỏng thì dừng. Gọi W là số lần kiểm tra.
   i) Lập bảng phân phối xác suất của W?
   ii) Tính trung bình và phương sai của W

e) Lấy ngẫu nhiên từng lọ trong hộp I để kiểm tra (không hoàn lại) cho đến khi gặp 2 lọ hỏng thì dừng. Gọi V là số lần kiểm tra.
   i) Lập bảng phân phối xác suất của V
   ii) Tính trung bình và phương sai của V
   iii) Tính xác suất để sau 4 lần kiểm tra thì dừng

Bài tập 2.9.

Thống kê về doanh số bán hàng (X) và chi phí quảng cáo (Y) (đơn vị: triệu đồng) của một công ty thu được bảng phân phối xác suất đồng thời như sau

$X \backslash Y$	100	200	300
1	0,15	0,1	0,04
1,5	0,05	0,2	0,15
2	0,01	0,05	0,25

a) Tìm chi phí quảng cáo trung bình và độ lệch chuẩn.

b) Tìm doanh số bán hàng trung bình khi chi phí quảng cáo là 1,5 triệu.

c) Nếu muốn doanh số bán hàng là 300 triệu đồng thì chi phí quảng cáo trung bình là bao nhiêu?

d) Doanh số bán hàng có phụ thuộc vào chi phí quảng cáo không?

Bài tập 2.10.

Cho biến ngẫu nhiên 2 chiều (X,Y) có bảng phân phối xác suất

$X \backslash Y$	1	3	4	7
2	0,15	0,05	0,2	0,1
5	0,3	0,1	0,05	0,05

a) Lập bảng phân phối xác suất của X và Y.

b) Tính $\mathrm{E}(\mathrm{X}), \mathrm{E}(\mathrm{Y}), \mathrm{Var}(\mathrm{X}), \mathrm{Var}(\mathrm{Y})$

c) Tính $\mathrm{Cov}(\mathrm{X},\mathrm{Y}), \rho(X,Y)$.

d) Xét tính độc lập của X và Y.

Phân phối Bernoulli: $B(p)$

Xét một phép thử Bernoulli τ, trong phép thử này ta chỉ quan tâm đến 2 biến cố A và $\bar{A}$ với $P(A) = \text{p} > 0$.

Đặt biến ngẫu nhiên X như sau: $X = \begin{cases} 1 & ;\ \text{nếu A xảy ra} \\ 0 & ;\ \text{nếu A không xảy ra} \end{cases}$

Khi đó, biến ngẫu nhiên X được gọi là có phân phối Bernoulli, ký hiệu $X \sim B(p)$

Định nghĩa 3.1.

Biến ngẫu nhiên rời rạc X nhận hai giá trị 0 và 1 được gọi là có phân phối Bernoulli, ký hiệu $X \sim B(p)$,

nếu hàm mật độ của X có dạng: $f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{khi} \quad x = 0,\ 1 \\ 0 & \text{khi} \quad x \notin \{0,1\} \end{cases}$

Khi đó, bảng phân phối xác suất của X như sau:

X	0	1
P	$\underbrace{1-p}_{=q}$	p

Tính chất

Giả sử $X \sim B(p)$. Khi đó, kỳ vọng $E[X] = p$ và phương sai $Var[X] = E[X^2] - (E[X])^2 = p(1-p) = pq$.

Phân phối nhị thức ($B(n;p)$)

Xét n phép thử Bernoulli độc lập và đồng nhất, trong mỗi phép thử ta chỉ xét 2 biến cố A và $\bar{A}$.
Xác suất xuất hiện của biến cố A trong mỗi phép thử không đổi và đều bằng p (tức là $P(A) = p$).
Gọi X là số lần xuất hiện biến cố A trong n phép thử trên. Khi đó, X là một biến ngẫu nhiên rời rạc, nhận các giá trị: $0,1,2,\ldots,n$. Đặt $\mathrm{X_i}$ là kết quả ở phép thử thứ i $(\mathrm{i} = 1,2,\ldots,\mathrm{n})$: $X_i = \begin{cases} 1 & ;\ \text{nếu A xảy ra} \\ 0 & ;\ \text{nếu A không xảy ra} \end{cases}$

Ta có thể xem $X = \sum_{i=1}^{n} X_i = X_1 + X_2 + \cdots + X_n$ với $X_1, \ldots X_n$ đôi một độc lập. Ta có:

$$
\begin{aligned}
P(X=0) &= P(X_1 = 0, X_2 = 0, \ldots, X_n = 0) = P(X_1 = 0)\, P(X_2 = 0)\, \ldots P(X_n = 0) && = C_n^0 p^0 (1-p)^{n-0} \\
P(X=1) &= \ldots && = C_n^1 p^1 (1-p)^{n-1} \\
P(X=2) &= \ldots && = C_n^2 p^2 (1-p)^{n-2} \\
&\vdots \\
P(X=k) &= \ldots && = C_n^k p^k (1-p)^{n-k} \\
&\vdots \\
P(X=n) &= \ldots && = C_n^n p^n (1-p)^{n-n}
\end{aligned}
$$

Định nghĩa 3.2.

Biến ngẫu nhiên rời rạc X được gọi là có phân phối nhị thức với hai tham số n và p, ký hiệu $X \sim B(n,p)$, nếu hàm mật độ của X là: $f_X(k) = P(X=k) = f(k,n,p) = \begin{cases} C_n^k p^k (1-p)^{n-k} & ;\text{ khi } \quad k = 0,1,2,\dots,n \\ 0 & ;\text{ khi } \quad k \notin \{0,1,2,\dots,n\} \end{cases}$

Bảng phân phối xác suất của X :

X	0	1	2	$\cdots$	n
P	$\underbrace{C_n^0 p^0 (1-p)^{n-0}}_{C_n^0 p^0 q^{n-0}}$	$\underbrace{C_n^1 p^1 (1-p)^{n-1}}_{C_n^1 p^1 q^{n-1}}$	$\underbrace{C_n^2 p^2 (1-p)^{n-2}}_{C_n^2 p^2 q^{n-2}}$	$\cdots$	$\underbrace{C_n^n p^n (1-p)^{n-n}}_{C_n^n p^n q^{n-n}}$

Tính chất

Giả sử $X \sim B(n;p)$. Khi đó:

1. Kỳ vọng:
   $E[X] = E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n] = np.$
2. Phương sai: Do $X_1, \dots, X_n$ đôi một độc lập nên
   $$Var[X] = Var[X_1 + X_2 + \cdots + X_n]$$
   $$= Var[X_1] + Var[X_2] \cdots + Var[X_n]$$
   $$= np(1-p) = npq$$
3. Mode: Ta xét $0 < p < 1$. So sánh

$\dfrac{f_X(k+1)}{f_X(k)} = \dfrac{(n-k)p}{(k+1)(1-p)}$ với 1, cho ta thấy rằng:

◇ khi $k > (n+1)p - 1$ thì $f(k) > f(k+1)$;

◇ khi $k = (n+1)p - 1$ thì $f(k) = f(k+1)$;

◇ khi $k < (n+1)p - 1$ thì $f(k) < f(k+1)$. Do đó:

- Nếu $(n+1)p - 1 \in \mathbb{Z}$ thì $(n+1)p - 1 = np - q$ và $(n+1)p$ là các giá trị của $Mod[X]$.
- Nếu $(n+1)p - 1 \notin \mathbb{Z}$ thì $Mod[X] = \lfloor (n+1)p \rfloor$, số nguyên lớn nhất không vượt quá $(n+1)p$.

Ví dụ 3.3.

Một đề thi có 20 câu hỏi, mỗi câu hỏi có 5 phương án trong đó chỉ có 1 phương án đúng. Sinh viên *A* trả lời ngẫu nhiên tất cả các câu hỏi. Gọi X là số câu trả lời đúng trong 20 câu hỏi trên.

a) Xác định phân phối xác suất của BNN RR X và tìm hàm mật độ của X.
b) Tính xác suất sinh viên *A* trả lời đúng 3 câu.
c) Tính xác suất sinh viên *A* trả lời đúng từ 3 đến 5 câu
d) Tính $E[X]$, $Var[X]$
e) Số câu hỏi sinh viên *A* có khả năng trả lời đúng lớn nhất là bao nhiêu?
f) Đề thi nên có ít nhất bao nhiêu câu để xác suất sinh viên *A* trả lời đúng ít nhất 1 câu không bé hơn 99% ?

Giải.

Ví dụ 3.4.

Tỷ lệ sản phẩm bị lỗi trong 1 lô hàng là 3%. Lấy ngẫu nhiên lần lượt 100 sản phẩm để kiểm tra. Tính xác suất để:

a) Có 3 sản phẩm bị lỗi.

b) Có không quá 3 sản phẩm bị lỗi.

Giải.

Ví dụ 3.5.

Một máy sản xuất được 200 sản phẩm trong một ngày. Xác suất để sản phẩm bị lỗi là 0.05. Tìm $E[X]$, $Var[X]$ và số sản phẩm bị lỗi có khả năng tin chắc của máy đó trong một ngày.

Giải.

Phân phối xác suất hình học (Geometric Distribution)

Xét một dãy các phép thử Bernoulli độc lập và đồng nhất, trong phép thử này ta chỉ quan tâm đến 2 biến cố A và $\bar{A}$ với $P(A) = p, 0 < p < 1$.

Đặt biến ngẫu nhiên X là số lần thực hiện phép thử cho đến khi biến cố A xảy ra, hàm mật độ:

$f_X(x) = (1-p)^{x-1}p, \quad x = 1,2,3\ldots$

Ví dụ 3.6.

Ta tung một đồng xu cho đến khi đạt được mặt ngửa. Gọi X là số lần đã tung đồng xu. Khi, đó, X có phân phối xác suất hình học. Ta có: $P(X = 10) = f_X(10) = (1-0.5)^9 \times 0.5 \simeq 0.001$

Tính chất

Giả sử X có phân phối xác suất hình học. Khi đó,

$E[X] = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = p\sum_{k=1}^{\infty} kq^{k-1}$ với $q = (1-p)$

Ta có: $p\sum_{k=1}^{\infty} q^k = pq\sum_{k=0}^{\infty} q^k = \frac{pq}{1-q} = p\left(-1+\frac{1}{1-q}\right)$.

Lấy đạo hàm hàm riêng hai vế theo biến q, ta được:

$\mu_X = E[X] = p\sum_{k=1}^{\infty} kq^{k-1} = \frac{p}{(1-q)^2} = \frac{1}{p}$.

Dùng công thức $Var[X] = E[X^2] - \big(E[X]\big)^2$ để tính toán, ta có thể tính được kết quả sau: $Var[X] = \frac{1-p}{p^2}$

Phân phối xác suất đều, rời rạc

Biến ngẫu nhiên rời rạc X nhận các giá trị: $x_1, x_2, \ldots, x_n$.

X có phân phối có phân phối xác suất đều, rời rạc nếu $P(X = x_i) = \dfrac{1}{n}, \forall i = 1,2,\ldots,n$.

Ví dụ 3.7.

Ta tung một cục xúc xắc. Gọi X là số nút thu được. Khi, đó, X có phân phối xác suất đều, rời rạc

Tính chất

Giả sử X X có phân phối xác suất đều, rời rạc và nhận các giá trị là các số nguyên liên tiếp: a, $a+1$, $a+2,\ldots$, b, với $a \le b$.
Khi đó, X nhận $(b-a+1)$ giá trị khác nhau, với xác suất cho từng giá trị là $1/(b-a+1)$.

Kỳ vong: $\mu_X = E[X] = \displaystyle\sum_{k=a}^{b} k\frac{1}{b-a+1} = \frac{b+a}{2}$.

Phương sai: $\sigma_X^2 = Var[X] = \dfrac{(b-a+1)^2-1}{12}$

Phân phối xác suất siêu bội (Hypergeometric Distribution): $H(N,M,n)$

Cho tập hợp có N phần tử, trong đó có M phần tử có tính chất A và $(N-M)$ phân tử không có tính chất A.
Lấy ngẫu nhiên ra n phần tử. Gọi X là số phần tử có tính chất A có trong n phần tử lấy ra.
Khi đó, X là BNN rời rạc có thể nhận các giá trị $0,1,2,..,n$ với các xác suất tương ứng là:

$$P(X=x) = \frac{C_M^x C_{N-M}^{n-x}}{C_N^n} \qquad \text{(gọi là công thức siêu bội)}$$

Định nghĩa 3.8.

BNN X có luật phân phối xác suất siêu bội là BNN rời rạc nhận các giá trị $0,1,2,..,n$ với các xác suất tương ứng được tính theo công thức siêu bội.

Ví dụ 3.9.

Một lô hàng gồm có 10 sản phẩm, trong đó có 4 loại A. Lấy ngẫu nhiên 4 sản phẩm từ lô hàng, tính xác suất để có 2 sản phẩm loại A

Giải. Gọi X là số sản phẩm loại A trong 4 sản phẩm lấy ra. X là BNN có phân phối siêu bội với tham số $N=10, M=4$ và $n=4$

$$P(X=2) = \frac{C_4^2 C_6^2}{C_{10}^4} = 0.4286$$

Tính chất

Cho $X \sim H(N,M,n)$, đặt $p = \dfrac{M}{N}$ và $q = 1 - p$. Ta có:

1. Kỳ vọng: $\mu_X = E[X] = np$
2. Phương sai: $\sigma_X^2 = Var[X] = npq\dfrac{N-n}{N-1}$

Ví dụ 3.10.

Một lô hàng gồm có 1000 sản phẩm, trong đó có 400 loại A. Lấy ngẫu nhiên 4 sản phẩm từ lô hàng, tính xác suất để có 2 sản phẩm loại A

Giải. Gọi X là số sản phẩm loại A trong 4 sản phẩm lấy ra. X là BNN có phân phối siêu bội với tham số $N = 1000, M = 400$ và $n = 4$

$$P(X=2) = \frac{C_{400}^2 C_{600}^2}{C_{1000}^4} = 0.346235 \approx C_4^2 \times \left(\frac{400}{1000}\right)^2 \times \left(\frac{600}{1000}\right)^2.$$

Chú ý. Nếu N khá lớn và $n << N$ thì $\dfrac{C_M^x \cdot C_{N-M}^{n-x}}{C_N^n} \approx C_n^x p^x (1-p)^{n-x}$ với $p = \dfrac{M}{N}$.

Như vậy: khi N khá lớn và $n << N$, ta có thể xem như $X \sim B(n;p)$ với $p = \dfrac{M}{N}$

Ví dụ 3.11.

Trong một lô hàng có 50 lọ thuốc trong đó có 10 lọ hỏng. Lấy ngẫu nhiên 6 lọ thuốc từ lô hàng trên. Gọi X là số lọ thuốc tốt trong 6 lọ thuốc được lấy.

a) Xác định phân phối xác suất của X ? Viết hàm mật độ của X ?

b) Tính xác suất lấy được 2 lọ thuốc tốt?

c) Tính xác suất lấy được ít nhất 1 lọ thuốc tốt?

d) Tính xác suất lấy được không quá 2 lọ thuốc tốt?

e) Tính $E[X], Var[X]$?

Giải.

Ví dụ 3.12.

Một trường gồm có 10000 sinh viên, trong đó có 1000 học kém. Một đoàn thanh tra đến trường, chọn ngẫu nhiên 100 sinh viên để kiểm tra. Tính xác suất để có 15 sinh viên học kém.

Giải.

Chú ý. Dù ta dùng phân phối nhị thức để xấp xỉ phân phối siêu bội, nhưng ta cũng có thể gặp khó khăn khi tính toán C_n^k trong công thức phân phối nhị thức với n lớn và k không quá bé.

Phân phối Poisson ra đời gắn liền với quá trình Poisson. Nó được áp dụng cho nhiều hiện tượng (có tính rời rạc), ví dụ như:

- Số lượng xe hơi đi ngang qua 1 điểm trên con đường trong một khoảng thời gian cho trước.
- Số lần gõ bị sai của khi đánh máy một trang giấy.
- Số cuộc điện thoại đến một số tổng đài trong mỗi phút.
- Số lần truy cập vào một website trong mỗi phút.
- Số lần đột biến xảy ra trên một đoạn DNA sau khi chịu một lượng bức xạ.
- Số lượng cây thông trên mỗi đơn vị diện tích rừng hỗn hợp.
- Số lượng ngôi sao trong một thể tích không gian vũ trụ.
- Số lượng bóng đèn bị cháy trong một khoảng thời gian xác định.
- Số lượng virut có thể lây nhiễm lên một tế bào trong cấu trúc tế bào.
- ....

Trong các ví dụ trên: số lần xuất hiện trong một khoảng (thời gian, không gian) cho trước đó phải là số nguyên, với một "xác suất để sự kiện (hiện tượng) đó xảy ra" là không đổi trong suốt khoảng (thời gian, không gian) đó; λ: Số lần biến cố xuất hiện "trung bình" trong khoảng thời gian t hay không gian h.

Phân phối xác suất Poisson $P(\lambda)$

- Biến ngẫu nhiên X có phân phối Poisson với tham số $\lambda > 0$ (trung bình số lần xuất hiện A) nếu X nhận các giá trị $0,1,2,\ldots,n,\ldots$ với xác suất tương ứng là:

$$p_k = P(X = k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!}.$$

Tính chất

$E[X] = \mathrm{Var}[X] = \lambda;$ $\mathrm{Mod}[X] = \mathrm{x}_0$, với x_0 là số nguyên : $\lambda - 1 \leq \mathrm{x}_0 \leq \lambda.$

Chú ý

Biến ngẫu nhiên X là số lần xuất hiện của biến cố A tại thời điểm ngẫu nhiên trong khoảng thời gian $(\mathrm{t}_1;\mathrm{t}_2)$ phải thỏa 2 điều kiện sau:

(i) Số lần xuất hiện của biến cố A trong khoảng $(\mathrm{t}_1;\mathrm{t}_2)$ không ảnh hưởng đến xác suất xuất hiện của biến cố A trong khoảng thời gian kế tiếp.

(ii) Số lần xuất hiện của biến cố A trong khoảng thời gian bất kỳ tỷ lệ với độ dài của khoảng thời gian đó.

Ví dụ 3.13.

Ở một tổng đài điện thoại, các cuộc điện thoại gọi đến ngẫu nhiên và độc lập với trung bình 6 cuộc gọi 1 phút. Biết rằng số cuộc gọi đến tổng đài điện thoại là biến ngẫu nhiên có phân phối Poisson.

a) Tính xác suất có đúng 5 cuộc gọi trong 2 phút
b) Tính xác suất có tối đa 2 cuộc gọi trong 30 giây
c) Tính xác suất có ít nhất 1 cuộc gọi trong 10 giây
d) Số cuộc gọi có khả năng lớn nhất trong 5 phút

Giải.

Ví dụ 3.14.

Kiểm tra ngẫu nhiên các mét vuông của một lô vải, kết quả thấy có 90% vải có lỗi. Hãy xác định số lỗi trung bình trong mỗi mét vuông vải, biết rằng số lỗi trên một mét vuông vải được xem là một biến ngẫu nhiên có phân phối Poisson.

Giải.

Tổng của các biến ngẫu nhiên độc lập có phân phối Poisson

Cho $X_1, X_2, \ldots, X_k$ là các biến ngẫu nhiên đôi một độc lập, trong đó: $X_i \sim P(\lambda_i), \forall i = \overline{1,k}$. Khi đó, $X = \sum\limits_{i=1}^{k} X_i$ cũng là biến ngẫu nhiên có phân phối Poisson và

$$X = \sum_{i=1}^{k} X_i \quad \sim \quad P\left(\sum_{i=1}^{k} \lambda_i\right)$$

Liên hệ giữa phân phối nhị thức và phân phối Poisson

Giả sử X là BNN có phân phối nhị thức với tham số $n > 50$ và $p < 0.1$. Đặt $np = \lambda$ (hằng số), ta có:

$$P(X = k) = C_n^k p^k q^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}$$

Điều náy có nghĩa là: X có phân phối Poisson $P(\lambda)$ với $\lambda = np$.

Ví dụ 3.15.

Trong một lô thuốc, tỉ lệ thuốc hỏng là 0,01. Kiểm tra 1000 ống.

a) Tính xác suất để gặp 5 ống bị hỏng.

b) Tính xác suất để gặp 50 ống bị hỏng.

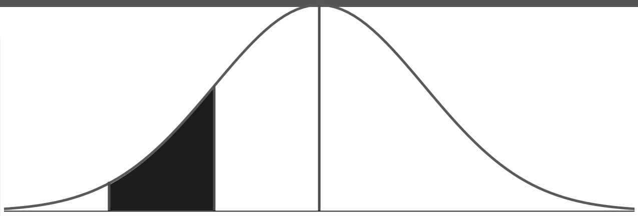
Giải.

Định nghĩa 3.16.

Biến ngẫu nhiên X được gọi là có phân phối chuẩn tham số μ và σ^2, ký hiệu $X \sim N\left(\mu;\sigma^2\right)$, nếu hàm mật độ của X có dạng

$$X \sim N\left(\mu;\sigma^2\right) \Leftrightarrow f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

với mọi $-\infty < x < +\infty$



Chú ý: Dạng đồ thị của hàm mật độ phân phối chuẩn

- Cực đại : $M\left(\mu;\dfrac{1}{\sigma\sqrt{2\pi}}\right)$
- Trục đối xứng: $x = \mu$
- Tiệm cận ngang: trục hoành
- Đồ thị có dạng hình quả chuông.

Tính chất

Nếu $X \sim N(\mu;\sigma^2)$ thì $E[X] = Mod[X] = Med[X] = \mu;\quad Var[X] = \sigma^2$

Phân phối chuẩn tắc

Biến ngẫu nhiên X được gọi là có phân phối chuẩn tắc (phân phối Gauss), ký hiệu $X \sim N(0;1)$, nếu hàm mật độ của X có dạng: $X \sim N(0;1) \quad \Leftrightarrow \quad f(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ với mọi $-\infty < x < +\infty$

Hàm mật độ của phân phối chuẩn được gọi là **hàm Gauss**. Đây là một hàm chẵn, i.e. $f(x) = f(-x), \forall x \in \mathbb{R}$.

Cách tìm giá trị của hàm Gauss tại 1 điểm x_0

1. Tính trực tiếp bằng cách thay $x = x_0$ trong công thức: $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
2. Tra bảng giá trị hàm Gauss.

Ví dụ 3.17.

Tìm các giá trị của: $f(0), f(1.96), f(4.1), f(-1), f(8)$.

Hàm phân phối của biến chuẩn tắc

Cho $X \sim N(0;1)$. Khi đó, hàm phân phối của X là

$$F_X(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,\mathrm{d}t = 0.5 + \int_{0}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,\mathrm{d}t, \forall x \in \mathbb{R}.$$

Hàm Laplace $\varphi(x)$

Hàm Laplace $\varphi(x)$ là hàm số xác định bởi $\varphi(x) = \int_{0}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,\mathrm{d}t, \forall x \in \mathbb{R}$.

Đây là một hàm lẻ, $\varphi(-x) = -\varphi(x), \forall x \in \mathbb{R}$.

Ta cũng có 2 cách tính giá trị hàm Laplace:

- thay giá trị $x = x_0$ vào công thức trên
- tra bảng giá trị hàm Laplace.

Ví dụ 3.18.

Tính giá trị $\varphi(0)$, $\varphi(0.45)$,$\varphi(-2.58)$, $\varphi(4.1)$,$\varphi(6)$.

Định lý (chuẩn tắc hóa phân phối chuẩn)

Nếu $X \sim N\left(\mu, \sigma^2\right)$ thì $Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$.

Hệ quả

Cho $X \sim N\left(\mu, \sigma^2\right)$, với x_1,$x_2 \in \mathbb{R}$, ta có:

1. $P(X \leq x_1) = P\left(\dfrac{X-\mu}{\sigma} < \dfrac{x_1-\mu}{\sigma}\right) = 0.5 + \varphi\left(\dfrac{x_1-\mu}{\sigma}\right)$
2. $P(X > x_1) = 0.5 - \varphi\left(\dfrac{x_1-\mu}{\sigma}\right)$
3. $P\left(x_1 \leq X \leq x_2\right) = \varphi\left(\dfrac{x_2-\mu}{\sigma}\right) - \varphi\left(\dfrac{x_1-\mu}{\sigma}\right)$.
4. $P(|X-\mu| \leq \varepsilon) = 2\varphi\left(\dfrac{\varepsilon}{\sigma}\right)$

   $P(|X-\mu| \leq k\sigma) = 2\varphi(k)$

   Từ đó, ta có:

   $P(|X-\mu| \leq \sigma) \approx 68.27\%$;

   $P(|X-\mu| \leq 2\sigma) \approx 95.45\%$;

   $P(|X-\mu| \leq 3\sigma) \approx 99.73\%$

Ví dụ 3.19.

Thống kê điểm thi môn toán X trong kỳ thi HK1, của học sinh lớp 10 tại trường X cho thấy X là BNN liên tục với $X \sim N(7{,}5; 2{,}25)$. Tính $P(X > 9{,}3)$

Giải.

Ví dụ 3.20.

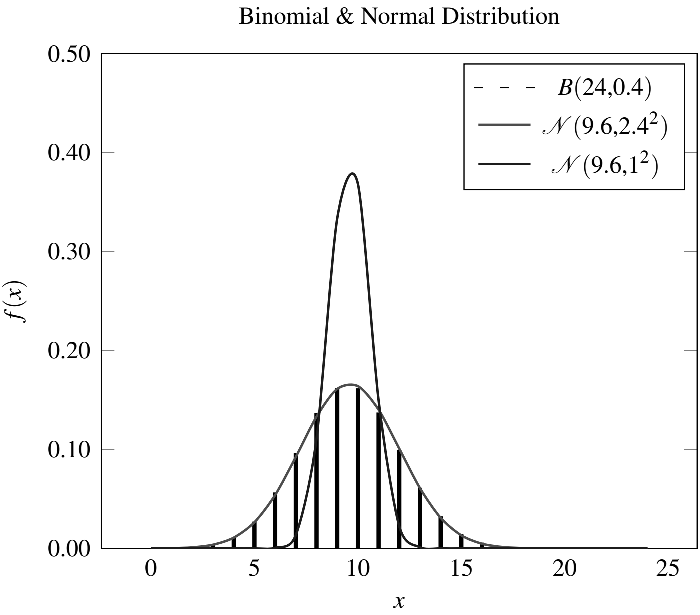
Lãi suất đầu tư vào Công ty B là BNN có phân phối chuẩn $N\left(\mu, \sigma^2\right)$, biết xác suất để đạt được lãi suất trên $20\%/$ năm là 0.2 và dưới $10\%/$ năm là 0.1.

1. Tìm kỳ vọng μ và phương sai σ^2.
2. Tính xác suất để khi đầu tư vào công ty B đó được lãi suất ít nhất $14\%/$năm.

Giải.

Xấp xỉ PP nhị thức bởi PP chuẩn

Cho $X \sim B(n;p)$ với $np \geq 5$ và $n(1-p) = nq \geq 5$. Ta có:
$\mu_X = E[X] = np$ và
$\sigma_X^2 = Var[X] = np(1-p) = npq$.
Khi đó, X có thể được xấp xỉ bởi phân phối chuẩn $N(\mu_X, \sigma_X^2)$.
Để tiện cho việc tính toán, ta đưa về phân phối chuẩn tắc qua phép biến đổi: $Z = \dfrac{X - np}{\sqrt{np(1-p)}} = \dfrac{X - \mu_X}{\sigma_X}$.



Để việc xấp xỉ này chính xác, ta áp dụng sự điều chỉnh nhỏ:

$$P(X \le x) = P(X \le x+0.5) \approx P\left(Z \le \frac{x+0.5-np}{\sqrt{np(1-p)}}\right) = 0.5 + \varphi\left(\frac{x+0.5-\mu_X}{\sigma_X}\right)$$

$$P(x \le X) = P(x-0.5 \le X) \approx P\left(\frac{x-0.5-np}{\sqrt{np(1-p)}} \le Z\right) = 0.5 - \varphi\left(\frac{x-0.5-\mu_X}{\sigma_X}\right)$$

$$P(a < X \le b) = P(X \le b) - P(X \le a) \approx \varphi\left(\frac{b+0.5-\mu_X}{\sigma_X}\right) - \varphi\left(\frac{a+0.5-\mu_X}{\sigma_X}\right)$$

$$P(a \le X < b) = P(X \ge a) - P(X \ge b) \approx \varphi\left(\frac{b-0.5-\mu_X}{\sigma_X}\right) - \varphi\left(\frac{a-0.5-\mu_X}{\sigma_X}\right)$$

$$P(X=a) = P(a-1<X\le a) \approx \varphi\left(\frac{a+0.5-\mu_X}{\sigma_X}\right)-\varphi\left(\frac{a-1+0.5-\mu_X}{\sigma_X}\right) = \int_{\frac{a-0.5-\mu_X}{\sigma_X}}^{\frac{a+0.5-\mu_X}{\sigma_X}} f(x)\,\mathrm{d}x \approx \frac{1}{\sigma_X} f\left(\frac{a-\mu_X}{\sigma_X}\right)$$

Chú ý. Với $a,b \in \mathbb{N}$, ta có thể chuyển: $(a \le X \le b) = (a-1 < X \le b) = (a \le X < b+1)$ và $(a < X < b) = (a < X \le b-1) = (a+1 \le X < b)$

Ví dụ 3.21.

Xác suất sinh được 1 em bé gái là 0,48. Tính xác suất sao cho trong 300 em bé sắp sinh

a) có 170 bé gái.

b) số bé gái vào khoảng từ 150 đến 170.

c) số bé gái ít nhất là 170.

Giải.

Xấp xỉ PP Poisson bởi PP chuẩn

Cho $X \sim P(\lambda)$ với $\lambda \geq 5$. Ta có: $\mu_X = E[X] = \lambda$ và $\sigma_X^2 = Var[X] = \lambda$.
Khi đó, X có thể được xấp xỉ bởi phân phối chuẩn $\mathcal{N}(\mu_X, \sigma_X^2) = \mathcal{N}(\lambda, \lambda)$.
Để tiện cho việc tính toán, ta đưa về phân phối chuẩn tắc qua phép biến đổi: $Z = \dfrac{X - \mu_X}{\sigma_X} = \dfrac{X - \lambda}{\sqrt{\lambda}}$.

Để việc xấp xỉ này chính xác, ta áp dụng sự điều chỉnh nhỏ:

$$P(X \leq x) = P(X \leq x+0.5) \approx P\left(Z \leq \frac{x+0.5-\lambda}{\sqrt{\lambda}}\right) = 0.5 + \varphi\left(\frac{x+0.5-\mu_X}{\sigma_X}\right)$$

$$P(x \leq X) = P(x-0.5 \leq X) \approx P\left(\frac{x-0.5-\lambda}{\sqrt{\lambda}} \leq Z\right) = 0.5 - \varphi\left(\frac{x-0.5-\mu_X}{\sigma_X}\right)$$

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) \approx \varphi\left(\frac{b+0.5-\mu_X}{\sigma_X}\right) - \varphi\left(\frac{a+0.5-\mu_X}{\sigma_X}\right)$$

$$P(a \leq X < b) = P(X \geq a) - P(X \geq b) \approx \varphi\left(\frac{b-0.5-\mu_X}{\sigma_X}\right) - \varphi\left(\frac{a-0.5-\mu_X}{\sigma_X}\right)$$

$$P(X = a) = P(a-1 < X \leq a) \approx \varphi\left(\frac{a+0.5-\mu_X}{\sigma_X}\right) - \varphi\left(\frac{a-1+0.5-\mu_X}{\sigma_X}\right) = \int_{\frac{a-0.5-\mu_X}{\sigma_X}}^{\frac{a+0.5-\mu_X}{\sigma_X}} f(x)\,\mathrm{d}x \approx \frac{1}{\sigma_X} f\left(\frac{a-\mu_X}{\sigma_X}\right)$$

Phân phối đều $U(a,b)$

Biến ngẫu nhiên X được gọi là có phân phối đều trong đoạn $[a,b]$, ký hiệu $X \sim U(a,b)$, nếu hàm mật độ của X có dạng:

$$X \sim U(a,b) \Leftrightarrow f(x) = \begin{cases} \dfrac{1}{b-a} & \text{khi} \quad x \in [a,b] \\ 0 & \text{khi} \quad x \notin [a,b] \end{cases}$$

Chú ý. Nếu $X \sim U(a,b)$ thì

$$E[X] = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \int_a^b \frac{x}{b-a}\,\mathrm{d}x = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2}.$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x = \int_a^b \frac{x^2}{b-a}\,\mathrm{d}x = \frac{b^3-a^3}{3(b-a)} = \frac{b^2+ab+a^2}{3}$$

và $Var[X] = \dfrac{b^2+ab+a^2}{3} - \left(\dfrac{a+b}{2}\right)^2 = \dfrac{(b-a)^2}{12}$

Ví dụ 3.22.

Giả sử một xe buýt chỉ ghé trạm đón khách trong khoảng thời gian từ 10 giờ đến 10 giờ 30 và thời điểm ghé trạm là biến ngẫu nhiên có phân phối đều. Nếu bạn đến trạm lúc 10 giờ thì

a) Thời gian trung bình bạn phải chờ là bao nhiêu?

b) Xác suất bạn phải chờ xe buýt hơn 10’ là bao nhiêu?

Giải.

Phân phối mũ $E(\lambda)$

Biến ngẫu nhiên X được gọi là có phân phối mũ với tham số $\lambda > 0$, ký hiệu $X \sim E(\lambda)$, nếu hàm mật độ của X có dạng:

$$X \sim E(\lambda) \Leftrightarrow f(x) = \begin{cases} 0 & \text{khi} \quad x < 0 \\ \lambda\, e^{-\lambda x} & \text{khi} \quad x \geq 0 \end{cases}$$

Chú ý. Nếu $X \sim E(\lambda)$ thì

$E[X] = \dfrac{1}{\lambda}$ và $Var[X] = \dfrac{1}{\lambda^2}$

Chú ý. Tuổi thọ của các thiết bị, tuổi thọ của nhiều loài sinh vật... có phân phối mũ.

Nếu $X \sim E(\lambda)$, với $a \geq 0$, ta có: $P(X \leq a) = \int_{-\infty}^{a} f(x)\,\mathrm{d}x = \int_0^a \lambda e^{-\lambda x}\,\mathrm{d}x = \left[-e^{-\lambda x}\right]_0^a = 1 - e^{\lambda a}$.

Ví dụ 3.23.

Tuổi thọ X(năm) của một mạch điện tử trong máy tính là biến ngẫu nhiên có phân phối mũ, trung bình 6,25. Thời gian bảo hành của mạch điện tử là 5 năm. Tính tỉ lệ mạch điện tử bán ra phải thay thế.

Giải.

Bài tập 3.1.

Một bài trắc nghiệm có 10 câu hỏi. Mỗi câu hỏi có 4 phương án trả lời trong đó chỉ có 1 phương án trả lời đúng. Một sinh viên làm bài trắc nghiệm này bằng cách chọn ngẫu nhiên một trong 4 phương án trả lời cho mọi câu hỏi. Biết rằng mỗi câu trả lời đúng được 2 điểm, mỗi câu trả lời sai bị trừ 1 điểm. Tính xác suất để số điểm của sinh viên là 14 điểm.

Bài tập 3.2.

Theo dõi trong một khoảng thời gian dài người ta nhận thấy rằng số tai nạn trong một ngày tuân theo luật Poisson có trung bình là 2 đối với những ngày trong tuần và là 3 đối với những ngày cuối tuần (là 2 ngày thứ bảy và chủ nhật). Quan sát ngẫu nhiên 1 ngày.

1. Tính xác suất có đúng 3 tai nạn xảy ra trong ngày đó.
2. Nếu trong ngày đó không có tai nạn, tính xác suất ngày đó là ngày trong tuần.

Bài tập 3.3.

Khi tiêm truyền một loại vaccine, trung bình có 2 trường hợp phản ứng trên 1000 trường hợp. Người ta dùng loại huyết thanh này tiêm cho 2500 người. Tính xác suất để có 7 trường hợp phản ứng

Bài tập 3.4.

Một bệnh viện lớn có 2000 bệnh nhân. Qua thống kê cho biết tỉ lệ bệnh nhân cần trợ giúp bằng thiết bị đặc biệt trong một ngày là 2,50%. Tính xác suất để trong một ngày không có quá 100 bệnh nhân cần trợ giúp bằng thiết bị đặc biệt.

Bài tập 3.5.

Đĩa cứng của một công ty có tuổi thọ là biến ngẫu nhiên có phân phối mũ với tuổi thọ trung bình là 8 năm. Khi bán một đĩa cứng ra thị trường, nếu không phải bảo hành công ty sẽ lãi 10USD, trong trường hợp ngược lại, chi phí bảo hành trung bình cho một đĩa cứng là 8USD. Hãy tính lợi nhuận trung bình của công ty đối với mỗi đĩa cứng bán ra thị trường nếu thời gian bảo hành là 1 năm.

Douglas C. Montgomery, George C. Runger, *Applied Statistics and Probability for Engineers*, Wiley (2018).