LÝ THUYẾT XÁC SUẤT CƠ BẢN

Phần I

Nhắc lại về tổ hợp và giải tích

1.1. Tập hợp

Định nghĩa 1.1.1. Cho trước hai tập A, B ta có:

$$\begin{aligned} A \cup B &= \{x : x \in A \text{ hoặc } x \in B\} \\ A \cap B &= \{x : x \in A \text{ và } x \in B\} \\ A \backslash B &= \{x : x \in A, x \notin B\} \\ \bigcup_{i=I} A_i &= \{x : \exists i \in I, x \in A_i\} \\ \bigcap_{i=I} A_i &= \{x : \forall i \in I, x \in A_i\} \end{aligned}$$

Định nghĩa 1.1.2. Cho $A_1 \subset A_2 \subset \ldots \subset A_n \subset \ldots$

$$\lim_{n\to\infty} A_n = \bigcup_{n\geq 1} A_n$$

Tương tự với dãy $(A_n)_n \geq 1$ thỏa $A_1 \supset A_2 \ldots$

$$\lim_{n\to\infty} A_n = \bigcap_{n\geq 1} A_n$$

Công thức De - Morgan

$$\overline{\left(\bigcup_{i\in I} A_i\right)} = \bigcap_{i\in I} \overline{A_i}$$

$$\overline{\left(\bigcap_{i=I} A_i\right)} = \bigcup_{i\in I} \overline{A_i}$$

1.2. Giải tích tổ hợp

Định nghĩa 1.2.1. (Quy tắc cộng)

Định nghĩa 1.2.2. (Quy tắc nhân)

Ví dụ: Từ các chữ số $0, 1, 2, 3, 4, 5$. Có bao nhiêu số có 4 chữ số.

a. là khác nhau.

b. là các số lẻ và các chữ số có thể khác nhau.

Giải

Gọi số có 4 chữ số cần tìm là: $\overline{abcd}$.

a. Ta có thể chia công việc thành lập số thỏa yêu cầu bài toán thành 4 giai đoạn:

Giai đoạn 1: a có 5 cách chọn $a \neq 0$

Giai đoạn 2: b có 5 cách chọn $b \in \{0, 1, \ldots, 5\} \setminus \{a\}$

Giai đoạn 3: c có 4 cách chọn $c \in \{0, 1, \ldots, 5\} \setminus \{a, b\}$

Giai đoạn 2: d có 3 cách chọn $b \in \{0, 1, \ldots, 5\} \setminus \{a, b, c\}$

Vậy có $5 \times 5 \times 4 \times 3 = 300$ số thỏa yêu cầu bài toán.

Định nghĩa 2.3. (Hoán vị) Cho tập A, có $|A| = n, 1$ nhóm hoán vị của các phần tử thuộc tập A là một song ánh từ A vào A. Số hoán vị của các phần tử thuộc A là:

$$P(n) = n! = 1.2.3 \ldots n$$

Chú ý: $(0! = 1)$.

Định nghĩa 2.4. (Chỉnh hợp) Chỉnh hợp chập k của n phần tử thuộc A là k bộ $(a_1, \ldots, a_k)$ thuộc A^k thỏa mãn $a_i \neq a_j; \forall i, j \in \overline{1, k}$. Số chỉnh hợp chập k của n phần tử thuộc A.

$$A_n^k = \frac{n!}{(n-k)!} \quad (0 \leq k \leq n)$$

Định nghĩa 2.5. Tổ hợp chập k của n phần tử thuộc A là một tập con gồm k phần tử của A. Số tổ hợp chập k của n phần tử thuộc

$$C_n^k = \frac{n!}{(n-k)!k!} = \begin{pmatrix} n \\ k \end{pmatrix}$$

Ví dụ: Một hồm gồm 8 bi đỏ, 6 bị trắng, 4 bi vàng. Người ta chọn ra 6 bi từ hộp đó. Hỏi có bao nhiêu cách chọn nếu:

a. Không yêu cầu gì thêm.

b. Trong 6 bi chọn ra có 2 bi đỏ, 2 bi trắng, 2 bi vàng.

c. Đúng 2 vàng.

d. Trong cách chọn có 2 vàng.

Phần II

Biến cố và xác suất

2.1. Phép thử.

2.1.1. Phép thử.

– Một thí nghiệm, một quan sát về một hiện tượng nào đó. Phép thử được gọi là một phép thử ngẫu nhiên nếu ta không biết trước kết quả của phép thử, nhưng vẫn có thể xác định đươc một tập hợp các kết quả có thể chứa nó.

– Tập hợp các kết quả có thể có của phép thử được gọi là không gian mẫu (không gian các sự kiện sơ cấp). Ký hiệu Ω.

– Mỗi phần tử của Ω được gọi là một sự kiện sơ cấp.

Ví dụ:

– Gieo đồng xu hai mặt (S, N): $\Omega = \{S, N\}$.

– Gieo đồng xu n lần: $\Omega = \{\omega = (\omega_1, \ldots, \omega_n) : \omega_i \in \{S, N\}\}$.

– Gieo con xúc sắc: $\Omega = \{1, 2, \ldots, 6\}$.

2.1.2. Biến cố.

1. Biến cố là tập hợp con của không gian mẫu.
2. Biến cố bất khả là biến cố không thể xảy ra khi thực hiện phép thử.
3. Biến cố chắc chắn là biến cố chắc chắn sẽ xảy ra trong khi thực hiện một phép thử.

2.1.3. Quan hệ giữa các biến cố.

– Quan hệ kéo theo

– Quan hệ bằng nhau (tương đương)

2.1.4. Các phép toán trên các biến cố.

– Biến cố tổng.

– Biến cố tích.

– Biến cố kí hiệu.

– Biến cố xung khắc.

– Biến cố đối lập.

2.2. Đại số và xích ma đại số

Định nghĩa 2.2.1. Họ các tập con $\mathcal{A}$ của Ω được gọi là một đại số nếu thỏa các tính chất sau:

1. $\Omega \in \mathcal{A}$.
2. $\forall A, B \in \mathcal{A} : A \cup B \in \mathcal{A}$.
3. $\forall A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$.

Chú ý: Nếu $\mathcal{A}$ là một đại số.

- $\emptyset \in \mathcal{A} : \Omega \in \mathcal{A} \Rightarrow \overline{\Omega} = \emptyset = \Omega$.
- $\forall A, B \in \mathcal{A} : A \cap B \in \mathcal{A} : A \cap B = \left(\overline{A} \cup \overline{B}\right)$.
- $A_1, \ldots, A_n \in \mathcal{A} \Rightarrow \begin{cases} \bigcup\limits_{i=1}^{n} A_i & \in \mathcal{A} \\ \bigcap\limits_{i=1}^{n} A_i & \in \mathcal{A} \end{cases}$

Ví dụ: $\mathcal{A} = \{\Omega, \emptyset\}; \mathcal{A} = \left\{\emptyset, A, \overline{A}, \Omega\right\}$.

Định nghĩa 2.2.2. Họ các tập con $\mathcal{F}$ của Ω được gọi là σ đại số nếu thỏa các tính chất sau:

1. $\Omega \in \mathcal{F}$.
2. $\forall A \in \mathcal{F}; \overline{A} \in \mathcal{F}$.
3. $\forall A_1, \ldots, A_n \in \mathcal{F} : \bigcup\limits_{i=1}^{n} A_i \in \mathcal{F}$.

2.3. Xác suất

Định nghĩa 2.3.1. (Độ đo xác suất) Cho ánh xạ $P : \mathcal{F} \longrightarrow \mathbb{R}$ được gọi là một độ đo xác suất nếu thỏa các tính chất sau:

1. $P(A) \geq 0, \forall A \in \mathcal{F}$.
2. $P(\Omega) = 1$.
3. $\forall A_1, \ldots, A_n \in \mathcal{F}, A_i \cap A_j = \emptyset, \forall i \neq j$

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

Định nghĩa 2.3.2. (Không gian xác suất) Bộ thứ tự $(\Omega, \mathcal{F}, P)$ trong đó Ω là một tập, $\mathcal{F}$ là một σ đại số trên Ω, P là một độ đo xác suất được gọi là một không gian xác suất.

Tính chất 2.3.3.

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

2. Nếu $A, B \in \mathcal{F}, A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B).$

3. $\forall A, B \in \mathcal{F} : P(A \cup B) = P(A) + P(B) - P(A \cap B).$

4. $\forall A_1, A_2, \ldots, A_n \in A$

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$$

Chứng minh:

1. $A_i = \emptyset, \forall i \geq 1$

$$\begin{aligned}
P\left(\bigcup_{i=1}^{\infty} A_i\right) &= \sum_{i=1}^{\infty} P(A_i) \\
P(\emptyset) &= \sum_{i=1}^{n} P(\emptyset) \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} P(\emptyset) \\
&= \lim_{n\to\infty} nP(\emptyset) \\
\Rightarrow \lim_{n\to\infty} [(n-1)P(\emptyset)] &= 0 \\
&\Rightarrow P(\emptyset) = 0
\end{aligned}$$

2. Tự cm

3. Tự cm

4. Ta đặt

$$\begin{aligned}
B_1 &= A_1 \\
B_2 &= (A_1 \cup A_2) \backslash B_1 \\
&\vdots \\
B_{n+1} &= \bigcup_{k=1}^{n+1} A_k \backslash B_n
\end{aligned}$$

$$\begin{cases} \bigcup_{n=1}^{\infty} A_n &= \bigcup_{n-1}^{\infty} B_n \\ B_i \cap B_j &= \emptyset, \forall i \neq j, B_n \subseteq A_n \end{cases}$$

$$\begin{aligned}
\Rightarrow P\left(\bigcup_{n=1}^{\infty} A_n\right) &= P\left(\bigcup_{n=1}^{\infty} B_n\right) \\
&= \sum_{n=1}^{\infty} P(B_n) \\
&\leq \sum_{n=1}^{\infty} P(A_n)
\end{aligned}$$

$$B_n \subset A_n \Rightarrow P(B_n) \le P(A_n)$$

2.4. Không gian xác suất hữu hạn

Định nghĩa 2.4.1. Trường hợp $|\Omega| < \infty$, ta có thể định nghĩa độ đo xác suất P trên $\mathcal{F} = P(\Omega)$

$$\mathbb{P} : \mathcal{F} \longrightarrow \mathbb{R}$$

$$A \longrightarrow \frac{|A|}{N}; \quad N = |\Omega|$$

$$P(\omega) = \frac{1}{N} \quad (\omega \in \Omega)$$

Ta gọi độ đo xác xuất này là độ đo xác suất rời rạc.

Ví dụ: Rút ngẫu nhiên không hoàn lại 3 lá bài từ một bộ bài, tính xác suất trong 3 lá bài từ 1 bộ bài, tính xác suất trong 3 lá bài vừa rút không có lá bài nào chất cơ.

Giải

Gọi A là : "Lấy được 3 lá bài không phải là cơ"
Số phần tử của không gian mẫu là C_{52}^3.
Số phần tử của biến cố A là C_{39}^3

$$P(A) = \frac{C_{39}^3}{C_{52}^3}$$

Ví dụ: Nếu 1 số có 3 chữ số 000 → 999 được chọn 1 cách ngẫu nhiên. Tính xác suất trong số đó có đúng 1 chữ số lớn hơn 5

Giải

Gọi $A_i = \left\{\text{Chọn được số có chữ số thứ } i > 5 \text{ và các chữ số còn lại } \le 5\right\}; i = 1, 2, 3$
Ta có $A = A_1 + A_2 + A_3; A_1 \cap A_2 \neq \emptyset$
$P(A) = P(A_1 + A_2 + A_3) = P(A_1) + P(A_2) + P(A_3)$.

$$P(A_1) = \frac{4.6.6}{10^3}$$

$$P(A_1) = P(A_2) = P(A_3)$$

$$P(A) = 3P(A_1) = 0.432$$

Ví dụ: Trong một hộp có 100 bóng đèn trong đó có 75 bóng đèn tốt và 25 bóng đèn hư. Lấy ngẫu nhiên không hoàn lại 15 bóng đèn, tính xác suất trong số 15 bóng đèn lấy ra có ít nhất một bóng đèn bị hư

Giải

A : "Trong 15 bóng đèn lấy ra có ít nhất 1 bóng bị hư" $\Rightarrow \overline{A}$: "Trong 15 bóng đèn không có bóng nào bị hư"

$$\begin{aligned} P(A) &= 1 - P\left(\overline{A}\right) \\ &= 1 - \frac{C_{75}^{15}}{C_{100}^{15}} \end{aligned}$$

Ví dụ: Gọi E, F, G là 3 biến cố. Hãy biểu diễn các biến cố sau thông qua 3 biến cố kể trên

1. Chỉ E xảy ra.
2. Cả $E \& G$ xảy ra, F không xảy ra.
3. Có ít nhất 1 trong 3 sự kiện xảy ra.
4. Có ít nhất 2 sự kiện xảy ra.
5. Có đúng 2 sự kiện xảy ra.
6. Có nhiều nhất 2 sự kiện xảy ra.

Giải

1. $E.\overline{F}.\overline{G}$
2. $E.G.\overline{F}$
3. $E + G + F$
4. $E.F.\overline{G} + E.\overline{F}.G + \overline{E}.F.G + E.F.G$
5. $\overline{E}.F.G + E.\overline{F}.G + E.F.\overline{G}$
6. $\overline{E}.\overline{F}.\overline{G} + E.\overline{F}.\overline{G} + \overline{E}.F.\overline{G} + \overline{E}.\overline{F}.G + \overline{E}.F.G + E.F.\overline{G} + E.\overline{F}.G$

Ví dụ: 2 con xúc sắc được gieo liên tục cho tới khi tổng điểm ở mặt phía trên của chúng là 5 hoặc 7 thì dừng lại. Tính xác suất tổng điểm của 2 xúc sắc bằng 5 xảy ra trước

Giải

Ta đặt $A_n = \{$Tổng 2 xúc sắc bằng 5 ở lần gieo thứ n và tổng 2 xúc sắc $\neq 5\&7$ ở $n-1$ lần gieo trước đó $\}$

$$A = A_1 + A_2 +, \ldots \bigcup_{n=1}^{\infty} A_n$$

$$A_n \cap A_m = \emptyset; \forall m \neq n$$

$$P(A) = \sum_{n=1}^{\infty} P(A_n)$$

2.5. Xác suất có điều kiện

Định nghĩa 2.5.1. Xác suất của sự kiện A được tính với giả thiết sự kiện B đã xảy ra được gọi là xác suất có điều kiện của A với điều kiện B, kí hiệu là $P(A|B)$ và được tính theo công thức

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

Tính chất 2.5.2.

1. $0 \leq P(A|B) \leq 1$
2. $P(B|B) = 1$
3. Nếu $AC = \emptyset$ thì $P[(A + C)|B] = P(A|B) + P(C|B)$
4. $P(\overline{A}|B) = 1 - P(A|B)$

Ví dụ: Gieo ngẫu nhiên 2 con xúc sắc cân đối & đồng chất

1. Tính xác suất khi biết tổng điểm 2 mặt của xúc sắc bằng 8 khi biết rằng xúc sắc thứ nhất đã ra 3.
2. Tính xác suất tổng điểm ra bằng 6 khi biết rằng xức sắc thứ nhất I ra bằng 3.

Giải

Ta có $\Omega = \{(\omega_i, \omega_j) : 1 \leq \omega_i, \omega_j \leq 6\} = \{(1,1), \ldots, (6,6)\}$
Gọi $A = \{\text{Tổng 2 mặt xúc sắc bằng 8}\}; B = \{\text{ xúc sắc thứ nhất ra mặt bằng 3}\}$
$P(A) = \frac{5}{36}; P(B) = \frac{6}{36} = \frac{1}{6}; P(A.B) = \frac{1}{36}$

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

Ví dụ: Trong một hộp gồm 8 bi đỏ, 4 bi trắng. Lấy ngẫu nhiên không hoàn lại 2 bi từ hộp, tính xác suất để cả 2 bi đều là bi đỏ

Giải

Đặt $A_i = \{\text{Bi đỏ thứ i}\}, \quad i = 1, 2$
Ta có

$$\begin{aligned} P(A_1A_2) &= P(A_2|A_1).P(A_1) \\ &= P(A_2|A_1).P(A_1) \\ &= \frac{7}{11}.\frac{8}{12} \\ &= \frac{14}{33} \end{aligned}$$

Định nghĩa 2.5.3. Công thức nhân xác suất.
Cho A và B là hai biến cố, nếu $P(B) > 0$ thì

$$\begin{aligned} P(AB) &= P(A|B).P(B) \\ &= P(B|A).P(A) \end{aligned}$$

Cho $A_1, A_2, ..., A_n$ là n biến cố

$$P(A_1, A_2, \ldots, A_n) = P(A_1).P(A_2|A_1).P(A_3|A_1A_2)\ldots P(A_n|A_1\ldots A_{n-1})$$

Ví dụ: Một bộ bài gồm 52 lá được chia ngẫu nhiên thành 4 phần bằng nhau cho người chơi. Tính xác suất mỗi người chơi nhận được đúng một con Át.

Giải

$A_1 = \Big\{\text{Con Át cơ nằm 1 trong 4 phần}\Big\}$
$A_2 = \Big\{\text{Con Át cơ và con Át rô nằm trong 2 phần khác nhau}\Big\}$
$A_3 = \Big\{\text{Con Át cơ, rô, nhép nằm trong 3 phần khác nhau}\Big\}$
$A_4 = \Big\{\text{4 con Át nằm trong 4 phần khác nhau}\Big\}$

$$\begin{aligned} P(A_4) &= P(A_1A_2A_3A_4) \\ &= P(A_1).P(A_2|A_1).P(A_3|A_1A_2).P(A_4|A_1A_2A_3) \\ &= 1.\frac{39}{51}.\frac{26}{50}.\frac{13}{49} \\ &= 0.106 \end{aligned}$$

2.6. Công thức BAYES

Định nghĩa 2.6.1. Hệ các biến cố đầy đủ
Hệ các biến cố $(A_i)_{i\in I}$ được gọi là một hệ đầy đủ nếu

$$\begin{cases} \bigcup_{i\in I} A_i &= \Omega \\ A_i \cap A_j &= \emptyset, \forall i \neq j \end{cases}$$

Ví dụ: Hệ $A, \overline{A}$ là hệ đầy đủ.

Định lý 2.6.2. Công thức xác suất toàn phần.
Cho $A_1, A_2, \ldots, A_n$ là hệ đầy đủ và B là biến cố bất kì

$$\begin{aligned} P(B) &= P(B|A_1).P(A_1) + \cdots + P(B|A_n).P(A_n) \\ &= \sum_{i=1}^{n} P(A_i)P(B|A_i) \end{aligned}$$

Ví dụ: Một người có hai đồng xu, đồng xu thứ nhất là đồng xu cân đối, đồng xu thứ 2 có xác suất mặt ngửa gấp 3 lần mặt sấp. Anh ta chọn ngẫu nhiên một đồng xu và gieo. Tính xác suất mặt hiện là mặt ngửa. Nếu gieo được mặt ngửa thì xác suất chọn được đồng xu cân đối là bao nhiêu?

Giải

$A_i = \left\{\text{Lấy được đồng xu thứ i để gieo}\right\}, \quad i = 1, 2$
$N = \left\{\text{Xuất hiện mặt ngửa sau khi gieo}\right\}$
Ta có

$$\begin{cases} A_1 A_2 & = \Omega \\ A_1 A_2 & = \emptyset \end{cases}$$

Áp dụng công thức xác suất toàn phần:

$$\begin{aligned} P(N) &= P(N|A_1)\,P(A_1) + P(N|A_2)\,P(A_2) \\ &= \frac{1}{2}.\frac{1}{2} + \frac{3}{4}.\frac{1}{2} \\ &= \frac{5}{8} \end{aligned}$$

$$\begin{aligned} P(A_1|N) &= \frac{P(A_1 N)}{P(N)} \\ &= \frac{P(N|A_1)\,P(A_1)}{P(N)} \\ &= \frac{\frac{1}{2}.\frac{1}{2}}{\frac{5}{8}} \\ &= \frac{2}{5} \end{aligned}$$

Định lý 2.6.3. Công thức Bayes.
Cho $A_1, A_2, \ldots, A_n$ là một hệ đầy đủ, B là một biến cố bất kỳ

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

Tính chất 2.6.4. Tính chất của xác suất có điều kiện
Cho A, B, C là 3 biến cố và $P(C) > 0$

1. $0 \leq P(A|B) \leq 1$
2. $P(B|B) = 1,\ P(\emptyset|C) = 0$
3. Nếu $AC = \emptyset$ thì $P[(A + C)|B] = P(A|B) + P(C|B)$
4. $P\left(\overline{A}|B\right) = 1 - P(A|B)$

Chú ý:$(\Omega, \mathcal{F}, \mathbb{P})$, $\mathcal{F} \cap C = \{F \cap C : F \in \mathcal{F}\}$
Hàm tập:

$$\begin{aligned} P(\cdot|C) : \mathcal{F} &\longrightarrow \mathbb{R} \\ A &\longmapsto P(A|C) \end{aligned}$$

Là một độ đo xác suất.

$$\begin{aligned} P(A \cup B|C) &= \frac{P[(A \cup B) C]}{P(C)} \\ &= \frac{P(AC \cup BC)}{P(C)} \\ &= \frac{P(AC) + P(BC)}{P(C)} \\ &= P(A|C) + P(B|C) \end{aligned}$$

2.7. Sự độc lập

Định nghĩa 2.7.1. Hai biến cố A và B được gọi là độc lập nếu:

$$P(AB) = P(A) P(B)$$

Lưu ý: Nếu A, B độc lập

$$\begin{aligned} P(A|B) &= P(A) \\ P(B|A) &= P(B) \end{aligned}$$

Ví dụ: Gieo hai con xúc sắc và theo dõi mặt phía trên của chúng:
$A = \{\text{Tổng hai xúc sắc bằng 7}\}$
$B = \{\text{Xúc sắc thứ nhất ra mặt 5}\}$

$$\begin{aligned} P(B) &= \frac{6}{36} = \frac{1}{6} \\ P(A) &= \frac{6}{36} = \frac{1}{6} \\ P(AB) &= \frac{1}{36} \end{aligned}$$

Định nghĩa 2.7.2. Họ các biến cố A_i, $i = \overline{1, n}$ được gọi là độc lập nếu:

$$\forall 2 \leq k \leq n, \ \forall i_1, i_2, \ldots, i_k \in \{1, \ldots, n\}$$

$$P(A_{i_1} A_{i_2} \ldots A_{i_k}) = P(A_{i_1}) \ldots P(A_{i_k})$$

Ví dụ:

$$A, B, C \text{ độc lập} \Leftrightarrow \begin{cases} P(AB) & = P(A)P(B) \\ P(AC) & = P(A)P(C) \\ P(BC) & = P(B)P(C) \\ P(ABC) & = P(A)P(B)P(C) \end{cases}$$

Ví dụ: Gieo hai con xúc sắc và quan sát mặt trên của chúng

– Tổng điểm trên hai xúc sắc bằng 9

$A = \{$Xúc sắc 1 ra mặt 1, 2 hoặc 3$\}$
$B = \{$Xúc sắc 1 ra mặt 4, 5 hoặc 6$\}$
$C = \{$Tổng điểm trên hai xúc sắc bằng 9$\}$
Ta có:

$$P(AB) = \frac{1}{6}, \quad P(AC) = \frac{1}{36}, \quad P(BC) = \frac{1}{12}$$
$$P(A) = \frac{18}{36}, \quad P(B) = \frac{18}{36}, \quad P(C) = \frac{4}{36}$$

$$P(AB) \neq P(A)P(B)$$
$$P(AC) \neq P(A)P(C)$$
$$P(BC) \neq P(B)P(C)$$

– Tổng điểm trên hai xúc sắc bằng 7

$A = \{$Xúc sắc 1 ra mặt 1, 2 hoặc 3$\}$
$B = \{$Xúc sắc 1 ra mặt 4, 5 hoặc 6$\}$
$C = \{$Tổng điểm trên hai xúc sắc bằng 7$\}$
Ta có:

$$P(A) = \frac{18}{36}, \quad P(B) = \frac{18}{36}, \quad P(C) = \frac{4}{36}$$

$$P(AB) = \frac{9}{36} = P(A)P(B)$$
$$P(BC) = \frac{3}{36} = P(C)P(D)$$
$$P(AC) = \frac{3}{36} = P(A)P(C)$$
$$P(ABC) = \frac{3}{36} \neq P(A)P(B)P(C)$$

Ví dụ: Giải lại ví dụ: 2 con xúc sắc được gieo liên tục cho tới khi tổng điểm ở mặt phía trên của chúng là 5 hoặc 7 thì dừng lại. Tính xác suất tổng điểm của 2 xúc sắc bằng 5 xảy ra trước.

$$P\left(E_n\right)=\left(1-\frac{10}{36}\right)^{n-1}\frac{4}{36}=\frac{1}{9}\left(\frac{13}{18}\right)^{n-1}$$

$B_i=\left\{\text{Gieo được tổng điểm bằng 5 ở lần thứ i}\right\}$
$A_i=\left\{\text{Gieo được tổng điểm bằng 5 hoặc 7 ở lần thứ i}\right\}$

$$E_i=A_1\ldots A_{n-1}B_i$$

$$\begin{aligned}
P\left(E_n\right)&=P\left(A_1\ldots A_{n-1}B_n\right)\\
&=P\left(A_1\right)\ldots P\left(A_{n-1}\right)P\left(B_n\right)\\
&=\left[P\left(A_1\right)\right]^{n-1}P\left(B_n\right)\\
P\left(E\right)&=\sum_{n=1}^{\infty}P\left(E_n\right)\\
&=\sum_{n=1}^{\infty}\frac{1}{9}\left(\frac{13}{18}\right)^{n-1}\\
&=\frac{1}{9}\sum_{n=1}^{\infty}\left(\frac{13}{18}\right)^{n}\\
&=\frac{1}{9}.\frac{1}{1-\frac{13}{18}}\\
&=\frac{1}{9}.\frac{18}{5}\\
&=\frac{2}{5}
\end{aligned}$$

Phần III

Biến ngẫu nhiên

3.1. Biến ngẫu nhiên

Định nghĩa 3.1.1. Cho hàm số

$$\begin{aligned} X\!:\!\Omega &\longrightarrow R \\ \omega &\longmapsto X(\omega) \end{aligned}$$

Thỏa $X^{-1}(B) \in \mathcal{F}, \forall B \in \mathcal{F}$ được gọi là một biến ngẫu nhiên. Trong đó: $\mathcal{F}$ là σ- đại số trên Ω.

Ví dụ: Tung hai đồng xu, gọi X là số mặt ngửa

$$\Omega = \{\text{SS, NS, SN, NN}\}$$
$$\mathcal{F} = \mathcal{P}(\Omega)$$

ω	SS	NS	SN	NN
$X(\omega)$	0	1	1	2

Chú ý:

$$(X \in B) = \{\omega : X(\omega) \in B\} = X^{-1}(B)$$
$$(X \le 5) = \{\omega \in \Omega : X(\omega) \le 5\}$$
$$(a \le X \le b) = \{\omega \in \Omega : a \le X(\omega) \le b\}$$

3.2. Phân phối của xác suất

Định nghĩa 3.2.1. Phân phối xác suất của biến ngẫu nhiên X là một độ đo cảm sinh trên $\mathbb{R}$, được xác định như sau:

$$\begin{aligned} P_X : B &\longrightarrow \mathbb{R} \\ B &\longmapsto P\left(X^{-1}(B)\right) = P(X \in B) \end{aligned}$$

$$\begin{aligned} P_X\left(\left[0,\frac{3}{2}\right]\right) &= P\left(X^{-1}\left(\left[0,\frac{3}{2}\right]\right)\right) \\ &= P\left(X \in \left(\left[0,\frac{3}{2}\right]\right)\right) \\ &= P\left(\{\text{SS, NS, SN}\}\right) \\ &= P\left(\{\text{SS}\}\right) + P\left(\{\text{NS}\}\right) + P\left(\{\text{SN}\}\right) \\ &= \frac{3}{4} \end{aligned}$$

Chú ý:

- Mọi ánh xạ trên Ω, trong trường hợp $|\Omega| < \infty$ là biến ngẫu nhiên.
- Mọi ánh xạ liên tục từ Ω vào $\mathbb{R}$ đều là biến ngẫu nhiên.

Định nghĩa 3.2.2. Hàm phân phối xác suất
Cho biến ngẫu nhiên X, hàm

$$\begin{aligned} F_X : \mathbb{R} &\longrightarrow [0,1] \\ x &\longrightarrow P\left(X \le x\right) \end{aligned}$$

được gọi là hàm phân phối xác suất của X.

Ví dụ: Tung hai đồng xu, gọi X là số mặt ngửa

$$\begin{aligned} \Omega &= \{\text{SS, NS, SN, NN}\} \\ \mathcal{F} &= \mathcal{P}\left(\Omega\right) \end{aligned}$$

- $x < 0$:

$$\begin{aligned} F_X\left(x\right) &= P\left(X \le x\right) \\ &= P\left(\{\omega : X\left(\omega\right) \le x\}\right) \\ &= 0 \end{aligned}$$

- $0 \le x < 1$:

$$\begin{aligned} F_X\left(X \le x\right) &= P\left(\omega \in \Omega, X\left(\omega\right) \le x\right) \\ &= P\left(\omega \in \Omega : X\left(\omega\right) = 0\right) \\ &= P\left(\text{SS}\right) \\ &= \frac{1}{4} \end{aligned}$$

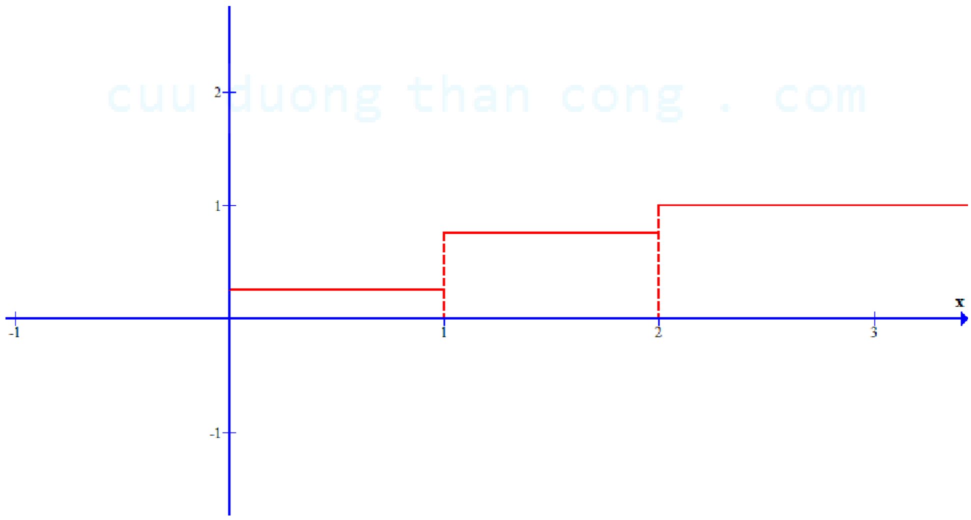
$$\begin{aligned} F_X\left(0\right) &= P\left(X \le 0\right) \\ &= P\left(X = 0\right) \\ &= P\left(\{\text{SS}\}\right) \\ &= \frac{1}{4} \end{aligned}$$

$-\ 1 \leq x < 2:$

$$\begin{aligned}F_X(x) &= P(X \leq x)\\ &= P(\omega \in \Omega : X(\omega) \leq x)\\ &= P\big(\omega \in \Omega : X(\omega) = 0 \text{ hoặc } X(\omega) = 1\big)\\ &= P(\omega \in \Omega : X(\omega) = 0) + P(\omega \in \Omega : X(\omega) = 1)\\ &= \frac{1}{4} + \frac{2}{4}\\ &= \frac{3}{4}\end{aligned}$$

$-\ x > 2:$

$$\begin{aligned}F_X(x) &= P(X \leq x)\\ &= P((X = 0) \cup (X = 1) \cup (X = 2))\\ &= P(X = 0) + P(X = 1) + P(X = 2)\\ &= \frac{1}{4} + \frac{2}{4} + \frac{1}{4}\\ &= 1\end{aligned}$$



Mệnh đề 3.2.3. Hàm phân phối xác suất của F_X của biến ngẫu nhiên X thỏa:

1. Hàm F_X không giảm
2. F_X liên tục phải.
3. $F_X(+\infty) = \lim_{x \to +\infty} F_X(x) = 1$

   $F_X(-\infty) = \lim_{x \to -\infty} F_X(x) = 0$
4. $F(a \leq X \leq b) = F_X(b) - F_X(a)$

Ví dụ: Chứng minh: $\forall x_1, x_2,\ x_1 \leq x_2$ suy ra $F_X(x_1) \leq F_X(x_2)$.

Giải

$$F_X(x_1) = P(X \leq x_1)$$
$$F_X(x_2) = P(X \leq x_2)$$

$$\begin{aligned}(X \leq x_1) &= \{\omega \in \Omega : X(\omega) \leq x_1\} \\ &\subset \{\omega \in \Omega : X(\omega) \leq x_2\} \\ &= (X \leq x_2) \\ \Rightarrow P(X \leq x_1) &\leq P(X \leq x_2) \\ \Leftrightarrow F_X(x_1) &\leq F_X(x_2)\end{aligned}$$

Vậy ta có điều phải chứng minh.

3.3. Phân loại biến ngẫu nhiên

Định nghĩa 3.3.1. Biến ngẫu nhiên rời rạc:
Biến ngẫu nhiên X được gọi là rời rạc nếu nó chỉ nhận hữu hạn hoặc vô hạn đếm được giá trị.
Giả sử X là biến ngẫu nhiên rời rạc nhận giá trị $(x_i)_{i\in I} \subset \mathbb{R}$
Đặt

$$\begin{aligned} f_X : \mathbb{R} &\longrightarrow [0,1] \\ x &\longmapsto f_X(x) = \begin{cases} P(X = x) & , x \in \{x_i : i \in I\} \\ 0 & , x \notin \{x_i : i \in I\} \end{cases}\end{aligned}$$

Hàm f_X được gọi là hàm xác suất của biến ngẫu nhiên X.

Mệnh đề 3.3.2. Cho X là biến ngẫu nhiên rời rạc với hàm xác suất f_X thì:

1. $F_X(x) = \sum\limits_{x_i \leq x} f_X(x_i) = \sum\limits_{x_i \leq x} p_i$
2. $\sum\limits_{i\in I} f_X(x_i) = \sum\limits_{i \in I} p_i = 1$
3. $P(a \leq X \leq b) = \sum\limits_{a<x_i\leq b} f_X(x_i) = \sum\limits_{a<x_i\leq b} p_i$

Ví dụ: Cho X là biến ngẫu nhiên rời rạc với bảng phân phối xác suất như sau:

X	-2	-1	0	1	2
P	1/8	1/4	1/4	1/4	1/8

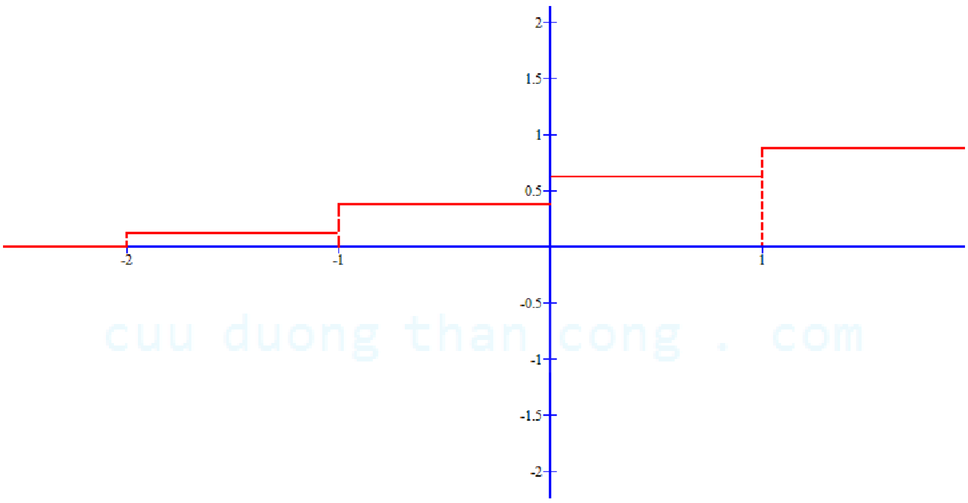
a) Xác định hàm phân phối xác suất F của X.

b) Tính $P(-1,5 \leq X \leq 1,5)$.

Giải

a)

$$F_X(x) = \begin{cases} 0 & , x < -2 \\ \frac{1}{8} & , -2 \leq x < -1 \\ \frac{3}{8} & , -1 \leq x < 0 \\ \frac{5}{8} & , 0 \leq x < 1 \\ \frac{7}{8} & , 1 \leq x < 2 \\ 1 & , \geq 2 \end{cases}$$



b) Tính $P(-1,5 \leq X \leq 1,5)$.

$$\begin{aligned} P(-1,5 \leq X \leq 1,5) &= F(1,5) - F(-1,5) \\ &= \frac{7}{8} - \frac{1}{8} \\ &= \frac{3}{4} \end{aligned}$$

Định nghĩa 3.3.3. Biến ngẫu nhiên liên tục tuyệt đối
Biến ngẫu nhiên X được gọi là biến ngẫu nhiên liên tục tuyệt đối nếu tồn tại f_X không âm, xác định trên $\mathbb{R}$ và thỏa:

$$P(X \in B) = \int_B f_X(x)\, dx$$

Hàm f_X được gọi là hàm mật độ xác suất của biến ngẫu nhiên X.
Chú ý: Nếu X là biến ngẫu nhiên liên tục với hàm mật độ f_X thì:

1. $P(X = a) = 0$, a là một hằng số.
2. $P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b) = \int_a^b f_X(x)dx$

3. Nếu f_X liên tục tại x thì

$$f_X(x) = \frac{d}{dx}(F_X(x)) = F'_X(x)$$

4. $F_X(x) = \int_{-\infty}^{+\infty} f_X(x)\,dx$

5. Mọi hàm f không âm và thỏa $\int_{-\infty}^{+\infty} f(x)\,dx = 1$ đều là hàm mật độ của biến ngẫu nhiên X nào đó.

Ví dụ: Cho X là biến ngẫu nhiên liên tục có hàm mật độ

$$f_X = \begin{cases} ax(x-2) & , x \in [0,2] \\ 0 & , x \notin [0,2] \end{cases}$$

a) Xác định $f_X(x)$.
b) Xác định F_X, tính $P(-1,5 \leq X \leq 1,5)$

Giải

a) Do f_X là hàm mật độ

$$\Leftrightarrow \begin{cases} f_X(x) \geq 0 & \forall x \\ \int_{-\infty}^{+\infty} f_X(x)\,dx & = 1 \end{cases}$$

$$\Leftrightarrow \begin{cases} ax(x-2) \geq 0 & , \forall x \in [0,2] \\ a\int_0^2 x(x-2)\,dx & = 1 \end{cases}$$

$$\Leftrightarrow \begin{cases} a & \leq 0 \\ a\left(\frac{x^3}{3} - x^2\right)\Big|_0^2 & = 1 \end{cases}$$

$$\Leftrightarrow \begin{cases} a & \leq 0 \\ a\left(\frac{-4}{3}\right) & = 1 \end{cases}$$

$$\Leftrightarrow a = -\frac{3}{4}$$

$$\Rightarrow f_X(x) = \begin{cases} -\frac{3}{4}x(x-2) & , x \in [0,2] \\ 0 & , x \notin [0,2] \end{cases}$$

b)
Tìm F_X:

$-\ x < 0:$

$$F_X(x) = 0$$

– $0 \leq x < 2$:

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{+\infty} f_X(t)\, dt \\ &= \int_0^x -\frac{3}{4} t(t-2)\, dt \\ &= -\frac{3}{4}\left(\frac{x^3}{3} - x^2\right) \end{aligned}$$

– $x \geq 2$:

$$F_X(x) = 1$$

Vậy ta có hàm phân phối F_X:

$$F_X(x) = \begin{cases} 0 & , x < 0 \\ -\frac{3}{4}\left(\frac{x^3}{3} - x^2\right) & , 0 \leq x < 2 \\ 1 & , x \geq 2 \end{cases}$$

$$P(-1,5 \leq x \leq 1,5) = F_X(1,5) - F_X(-1,5) = \frac{27}{2}$$

3.4.Hàm của một biến ngẫu nhiên.

Định nghĩa 3.4.1. Hàm của biến ngẫu nhiên rời rạc
Cho biến ngẫu nhiên X với bảng phân phối xác suất:

X	x_0	x_1	...
P	p_0	p_1	...

và f là một hàm đo được từ $\mathbb{R}$ vào $\mathbb{R}$ thì $Y = f(X)$ là một biến ngẫu nhiên rời rạc có bảng phân phối xác suất

Y	x_0	x_1	...
P	p_0	p_1	...

Với quy ước $f(x_i) = f(x_j)$ thì trong bảng ta chỉ ghi $f(x)$ và $p_i + p_j$.

Ví dụ:

X	-2	-1	0	1	2
P	1/8	1/4	1/4	1/4	1/8

Xét $Y = f(X) = X^2$

Y	4	1	0	1	4
P	1/8	1/4	1/4	1/4	1/8

Y	0	1	4
P	1/4	1/2	1/4

Định nghĩa 3.4.2 Hàm của biến ngẫu nhiên liên tục tuyệt đối.
Sử dụng hàm phân phối.

- Từ định nghĩa của $Y = f(X)$, ta tìm hàm phân phối của Y.
- Nếu F_Y khả vi thì $f_Y(y) = F'(y)$.

Ví dụ: Cho

$$f_X(x) = \begin{cases} -\frac{3}{4}x(x-2) & x \in [0,2] \\ 0 & x \notin [0,2] \end{cases}$$

Hãy tìm hàm mật độ của biến ngẫu nhiên $Y = X^2$.

$$F_Y(y) = P(Y \le y) = P\left(X^2 \le y\right)$$

- Nếu $y < 0$: $F_Y(y) = 0$
- Nếu $y \ge 0$:

$$\begin{aligned} F_Y(y) &= P\left(X^2 \le y\right) \\ &= P\left(-\sqrt{y} \le X \le \sqrt{y}\right) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \end{aligned}$$

$$\begin{aligned} \Rightarrow f_Y(y) &= [F_Y(y)]' \\ &= [F_X(\sqrt{y})]' - [F_X(-\sqrt{y})]' \\ &= \frac{1}{2\sqrt{y}} F'_X(\sqrt{y}) - \left(-\frac{1}{2\sqrt{y}}\right) F'_X(-\sqrt{y}) \\ &= \frac{1}{2\sqrt{y}} \left[F'_X(\sqrt{y}) + F'_X(-\sqrt{y})\right] \\ &= \frac{1}{2\sqrt{y}} \left[f_X(\sqrt{y}) + f_X(-\sqrt{y})\right] \\ &= \frac{1}{2\sqrt{y}} \left(-\frac{3}{4}\right) \sqrt{y}\left(\sqrt{y} - 2\right) \\ &= -\frac{3}{8}\left(\sqrt{y} - 2\right) \end{aligned}$$

Định lý 3.4.3. Giả sử g là một hàm có hàm ngược, g xác định trên một tập mở chứa giá trị của f_X và $h = g^{-1}$ là hàm khả vi, thì biến ngẫu nhiên $Y = g(X)$ có hàm mật độ:

$$\begin{aligned} f_Y(y) &= f_X\left(g^{-1}(y)\right) . \left[g^{-1}(y)\right]' \\ &= f_X(h(y)) . h'(y) \end{aligned}$$

Ví dụ: Cho biến ngẫu nhiên X có hàm mật độ f_X như dưới đây và $Y = g(X) = X^2$. Tìm f_Y.

$$f_X(x) = \begin{cases} -\frac{3}{4}x(x-2) & x \in (0,2) \\ 0 & x \notin (0,2) \end{cases}$$

$$\begin{aligned} g : (0,2) &\longrightarrow \mathbb{R} \\ x &\longmapsto x^2 \end{aligned}$$

$$\begin{aligned} h(x) &= g^{-1}(x) = \sqrt{x} \\ \Rightarrow f_Y(y) &= f_X(h(y)).h'(y) \\ &= -\frac{3}{4}\sqrt{y}(\sqrt{y}-2)\frac{1}{2\sqrt{y}} \\ &= -\frac{3}{8}(\sqrt{y}-2) \end{aligned}$$

Định lý 3.4.4. Nếu $g : U \longrightarrow V$ là một ánh xạ khả vi, trong đó U, V là các tập mở. Cho X là biến ngẫu nhiên có hàm mật độ f_X và U sai khác so với $\sup(f_X)$ một tập có độ đo không. Giả sử h_i $\left(i = \overline{1,n}\right)$ là các khả nghịch phải của g, h_i khả vi và có đạo hàm khác không, $h_i(V)$ tạo thành một phân hoạch của U thì $Y = g(X)$ có hàm mật độ là:

$$f_Y(y) = \sum_{i=1}^{n} f_X(h_i(y)).h_i'(y)$$

Ví dụ: Cho X là biến ngẫu nhiên có hàm mật độ $f_X(x)$. Hãy xác định hàm mật độ của $Y = X^2$.

$$\begin{aligned} g : R &\longrightarrow \mathbb{R}^+ \\ x &\longmapsto x^2 \end{aligned}$$

$$\begin{aligned} h_1 : \mathbb{R}^+ &\longrightarrow \mathbb{R} \\ x &\longmapsto \sqrt{x} \end{aligned}$$

$$\begin{aligned} h_2 : \mathbb{R}^+ &\longrightarrow \mathbb{R} \\ x &\longmapsto -\sqrt{x} \end{aligned}$$

h_1, h_2 là 2 khả nghịch phải của g.
$h_1(\mathbb{R}^+) = \mathbb{R}^+; h_2(\mathbb{R}^+) = \mathbb{R}^-; \mathbb{R} = \mathbb{R}^+ \cup \mathbb{R}^-$

Hàm mật độ của $Y = X^2$ là:

$$\begin{aligned} f_Y(y) &= f_X(h_1(y))).h_1'(y) + f_X(h_2(y)).h_2'(y) \\ &= f_X(\sqrt{y})\frac{1}{2\sqrt{y}} - f_X(-\sqrt{y})\frac{1}{2\sqrt{y}} \end{aligned}$$

Ví dụ: Cho $\lambda > 0$ và hàm f được định nghĩa như sau:

$$f(x) = \begin{cases} \frac{1}{2}e^{-\lambda x} & x \geq 0 \\ \frac{1}{2}e^{\lambda x} & x < 0 \end{cases}$$

1. Chứng tỏ f là hàm mật độ của một biến ngẫu nhiên nào đó.
2. Nếu X là biến ngẫu nhiên có hàm mật độ f. Hãy xác định hàm mật độ của $Y = |X|$.

Giải

1. Kiểm tra hàm mật độ:

$$\begin{cases} f(x) \geq 0 & \forall x \\ \displaystyle\int_{-\infty}^{+\infty} f(x)dx & = 1 \end{cases}$$

Ta có: $f(x) \geq 0, \forall x \in \mathbb{R}$.

Mặt khác:

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x)\,dx &= \frac{1}{2}\int_{-\infty}^{0} \lambda e^{\lambda x}dx + \frac{1}{2}\int_{0}^{+\infty} \lambda e^{-\lambda x}dx \\ &= \frac{1}{2}\lim_{t\to\infty}\int_{t}^{0} \lambda e^{\lambda x}dx + \frac{1}{2}\lim_{t\to\infty}\int_{0}^{t} \lambda e^{-\lambda x}dx \\ &= \frac{1}{2}\lim_{t\to-\infty}\left(e^{\lambda x}\right)\Big|_t^0 + \frac{1}{2}\lim_{t\to\infty}\left(-e^{-\lambda x}\right)\Big|_0^t \\ &= \frac{1}{2}\lim_{t\to-\infty}\left(1 - e^{\lambda t}\right) + \frac{1}{2}\lim_{t\to\infty}\left(-e^{-\lambda t} + 1\right) \\ &= \frac{1}{2} + \frac{1}{2} = 1. \end{aligned}$$

2. $Y = |X|$
$P(Y \leq y) = P(|X| \leq y)$

- $y < 0 \Rightarrow P(Y \leq y) = 0$
- $y \geq 0$

$$
\begin{aligned}
P(Y \le y) &= P(|X| \le y) \\
&= P(-y \le X \le y) \\
&= \int_{-y}^{0} \frac{1}{2} e^{\lambda x} \lambda dx + \int_{0}^{y} \frac{1}{2} e^{-\lambda x} \lambda dx \\
&= \left(\frac{1}{2} e^{\lambda x}\right)\Big|_{-y}^{0} - \left(\frac{1}{2} e^{-\lambda x}\right)\Big|_{0}^{y} \\
&= \frac{1}{2}\left(1 - e^{-\lambda y}\right) - \frac{1}{2}\left(e^{-\lambda y} - 1\right) \\
&= 1 - e^{-\lambda y}
\end{aligned}
$$

3.5. Một số tính chất của biến ngẫu nhiên.

Định nghĩa 3.5.1. Kỳ vọng

Trường hợp rời rạc: Cho X là biến ngẫu nhiên rời rạc có phân phối xác suất:

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

thì kỳ vọng của X là:

$$E(X) = \sum_{i=1}^{\infty} x_i p_i$$

Trường hợp liên tục tuyệt đối: Cho X là biến ngẫu nhiên liên tục tuyệt đối với hàm mật độ f_X, kỳ vọng của X là:

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

Ví dụ: Cho biến ngẫu nhiên X nhận giá trị trên $\mathbb{N}$ với:

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, k = 0, 1, \dots$$

Tính $E(X)$.

Giải

$$
\begin{aligned}
E(X) &= \sum_{k=0}^{+\infty} k e^{-\lambda} \frac{\lambda^k}{k!} \\
&= \sum_{k=1}^{+\infty} e^{-\lambda} \frac{\lambda^k}{(k-1)!} \\
&= e^{-\lambda} \lambda \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&= e^{-\lambda} \lambda \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} \\
&= e^{-\lambda} \lambda e^{\lambda} \\
&= \lambda
\end{aligned}
$$

Ví dụ: Cho

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

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} x f_X(x) dx \\
&= \int_0^1 x.2x dx \\
&= \left(\frac{2x^3}{3} \right) \Big|_0^1 \\
&= \frac{2}{3}
\end{aligned}
$$

3.5.2. Tính chất của kỳ vọng

1. $E(c) = c$, với c là hằng số.
2. $E(X+Y) = E(X) + E(Y)$, $E(cX) = cE(X)$.
3. $E(XY) = E(X)E(Y)$, nếu X, Y là 2 biến ngẫu nhiên độc lập.

Kỳ vọng của hàm của một biến ngẫu nhiên.
Trường hợp rời rạc: Cho X là biến ngẫu nhiên với bảng phân phối:

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

và $Y = g(X)$ thì

$$
E(Y) = \sum_{i=1}^{\infty} g(x_i) p_i
$$

Trường hợp liên tục tuyệt đối: Cho X là biến ngẫu nhiên với hàm mật độ f_X và $Y = g(X)$, thì:

$$E(Y) = \int_{-\infty}^{+\infty} g(x) f_X(x) \, dx$$

Ví dụ: Cho X là biến ngẫu nhiên có hàm mật độ:

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

Tính $E\left(X^2\right)$.

$$\begin{aligned} E\left(X^2\right) &= \int_{-\infty}^{+\infty} x^2 f_X(x) dx \\ &= \int_0^1 x^2 2x dx \\ &= \left(\frac{x^4}{2}\right)\Bigg|_0^1 \\ &= \frac{1}{2} \end{aligned}$$

Định nghĩa 3.5.3. Phương sai

Cho X là biến ngẫu nhiên với kỳ vọng $E(X)$. Phương sai của X, ký hiệu $Var(X)$ là giá trị:

$$Var(X) = E(X - E(X))^2$$

$$Var(X) = \begin{cases} \sum_i (x_i - E(X))^2 p_i \\ \int_{-\infty}^{+\infty} (x_i - E(X))^2 f_X(x) \, dx \end{cases}$$

Ví dụ: Tính phương sai của biến ngẫu nhiên X có hàm mật độ:

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

Ta có $E(X) = \dfrac{2}{3}$

$$\begin{aligned} Var(X) &= E(X - E(X))^2 \\ &= \int_{-\infty}^{+\infty} \left(x - \frac{2}{3}\right)^2 f_X(x) \, dx \\ &= \int_0^1 (x - \frac{2}{3})^2 2x dx \\ &= 2\int_0^1 \left(x^3 - \frac{4}{3}x^2 + \frac{4}{9}x\right) dx \\ &= \frac{1}{18} \end{aligned}$$

Chú ý: Trong tính toán chúng ta hay sử dụng công thức sau:

$$Var\left(X\right) = E\left(X^2\right) - \left(E\left(X\right)\right)^2$$

Tính chất 3.5.4.

1. $Var\left(c\right) = 0$.
2. $Var\left(cX\right) = c^2 Var\left(X\right)$.
3. $Var\left(X+Y\right) = Var\left(X\right) + Var\left(Y\right)$ nếu X, Y độc lập.

Định nghĩa 3.5.5. Độ lệch chuẩn
Cho biến ngẫu nhiên X có phương sai $Var\left(X\right)$, độ lệch chuẩn của X được định nghĩa là giá trị: $\sigma_X = \sqrt{Var\left(X\right)}$

Định nghĩa 3.5.6. Mode (Giá trị chắc chắn).
Mode của X là giá trị mà hàm xác suất (trường hợp X là biến ngẫu nhiên rời rạc), hoặc hàm mật độ xác suất (trường hợp X là biến ngẫu nhiên liên tục) nhận giá trị lớn nhất.
Chú ý: Xét

$$f_X\left(x\right) = \begin{cases} 2x & x \in [0,1] \\ 0 & x \notin [0,1] \end{cases}$$

$Mod(X)$ không tồn tại.

Định nghĩa 3.5.7. Trung vị.
Trung vị của biến ngẫu nhiên X là giá trị m của biến ngẫu nhiên X thỏa:

$$\begin{cases} P\left(X \le m\right) & \ge \frac{1}{2} \\ P\left(X \ge m\right) & \ge \frac{1}{2} \end{cases} \text{ hay } P\left(X \le m\right) = P\left(X \ge m\right)$$

Phần IV

Một số phân phối đặc biệt

4.1. Phân phối Bernoulli

Định nghĩa 4.1.1. Biến ngẫu nhiên X gọi là có phân phối Bernoulli với tham số p , ký hiệu $X \sim B(1,p)$ nếu nó có bảng phân phối xác suất như sau:

X	0	1
P	q	p

$(q = 1 - p)$

Chú ý: Ta có thể hình dung một biến ngẫu nhiên có phân phối Bernoulli được sử dụng để mô tả kết quả của phép thử ngẫu nhiên như sau:
Thực hiện một phép thử bất kỳ, ta quan tâm đến biến cố A có xảy ra hay không.
Nếu A xảy ra thì biến ngẫu nhiên X nhận giá trị 1, ngược lại nhận giá trị 0.
Giả sử xác suất xảy ra biến cố A là $P(A) = p$, và $P\left(\overline{A}\right) = 1 - P(A) = 1 - p$ thì rõ ràng X có phân phối Bernoulli với tham số p.

4.2. Phân phối nhị thức.

Định nghĩa 4.2.1. Biến ngẫu nhiên X có phân phối nhị thức với tham số n, p ký hiệu $B(n,p)$ nếu nó nhận các giá trị nguyên từ 0 đến n với xác suất tương ứng là:

$$P(X = k) = C_n^k q^{n-k} p^k, 0 \le k \le n, (q = 1 - p)$$

Chú ý: Nếu $X \sim B(n,p)$ thì X có thể biểu diễn như sau:

$$X = X_1 + X_2 + ... + X_n$$

trong đó: $X_i \sim B(1,p)$ và X_i độc lập.
Như vậy ta có thể hình dung X chính là biến ngẫu nhiên mô tả số lần biến cố A mà ta quan tâm xảy ra trong n lần tiến hành thí nghiệm liên tiếp và độc lập với nhau.

X	0	$1\ldots$	$k\ldots$	n
P	$C_n^0 q^n$	$C_n^1 q^{n-1} p^1 \ldots$	$C_n^k q^{n-k} p^k \ldots$	$C_n^n p^n$

Ví dụ: Một bài thí nghiệm gồm 10 câu hỏi, mỗi câu hỏi có 4 phương án trả lời trong đó chỉ có 1 phương án đúng. Giả sử mỗi câu trả lời đúng được 4 điểm, mỗi câu trả lời sai trừ 2 điểm. Một sinh viên làm bài bằng cách chọn ngẫu nhiên đáp án cho các câu hỏi. Tính xác suất:

a. Để sinh viên được 4 điểm.

b. Để sinh viên được điểm âm

Giải

Gọi X là số câu trả lời đúng của sinh viên trong 10 câu hỏi.

$$X \sim B\left(10, \frac{1}{4}\right)$$

a. Sinh viên được 4 điểm.

$$\begin{aligned} 4X - 2(10 - X) &= 4 \\ \Leftrightarrow 6X &= 24 \\ \Leftrightarrow X &= 4 \end{aligned}$$

$$P(X = 4) = C_{10}^{4}\left(\frac{1}{4}\right)^{4}\left(\frac{3}{4}\right)^{6}$$

b. Sinh viên được điểm âm.

$$\begin{aligned} 4X - 2(10 - X) &< 0 \\ \Leftrightarrow X &< \frac{10}{3} \end{aligned}$$

$$\begin{aligned} P\left(X < \frac{10}{3}\right) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= C_{10}^{0}\left(\frac{3}{4}\right)^{10} + C_{10}^{1}\frac{1}{4}\left(\frac{3}{4}\right)^{9} + C_{10}^{2}\left(\frac{1}{4}\right)^{2}\left(\frac{3}{4}\right)^{8} + C_{10}^{3}\left(\frac{1}{4}\right)^{3}\left(\frac{3}{4}\right)^{7} \end{aligned}$$

Định lý 4.2.2. Cho $X \sim B(n, p)$.

1. $E(X) = np$.
2. $Var(X) = npq$.
3. $np - q \leq Mod(X) \leq np + q$.

4.3. Phân phối siêu bội.

Cho tập hợp gồm N phần tử, trong đó có N_A phần tử có tính chất $A, N - N_A$ phần tử không có tính chất A. Lấy ngẫu nhiên đồng thời n phần tử từ tập hợp trên. Gọi X là số phần tử có tính chất A trong N phần tử lấy ra . Khi đó X là biến ngẫu nhiên rời rạc nhận các giá trị k thỏa:

$$\begin{cases} 0 \leq k \leq n \\ n - (N - N_A) \leq k \leq N_A \end{cases}$$

Ký hiệu:

$$\begin{aligned} k_1 &= \max\{0, n - (N - N_A)\} \\ k_2 &= \min\{n, N_A\} \end{aligned}$$

Thì X nhận các giá trị nguyên k thỏa $k_1 \leq k \leq k_2$ với xác suất tương ứng là:

$$P\left(X = k\right) = \frac{C_{N_A}^{k} C_{N-N_A}^{n-k}}{C_N^n}$$

Biến ngẫu nhiên X nhận các giá trị nguyên k thỏa $k_1 \leq k \leq k_2$ với xác suất tương ứng là:

$$P\left(X = k\right) = \frac{C_{N_A}^{k} C_{N-N_A}^{n-k}}{C_N^n}$$

được gọi là có phân phối siêu bội với các tham số N, N_A, n.
Ký hiệu: $X \sim H\left(N, N_A, n\right)$

Ví dụ: Một lớp có 50 sinh viên trong đó có 30 nữ. Cần chọn ra 10 bạn để tham gia vào công tác chuẩn bị cho 1 hoạt động sắp tới của trường. Nếu ta chọn các bạn trên một cách ngẫu nhiên, xác suất để số sinh viên nữ được chọn không quá 3 là bao nhiêu. Xác suất để chọn được ít nhất 1 sinh viên nữ là bao nhiêu?

Giải

Gọi X là số sinh viên nữ số 10 sinh viên được chọn.

$$X \sim H\left(50, 30, 10\right)$$

Số sinh viên nữ được chọn không quá 3: $X \leq 3$.

$$\begin{aligned} P\left(X \leq 3\right) &= P\left(X = 0\right) + P\left(X = 1\right) + P\left(X = 2\right) + P\left(X = 3\right) \\ &= \frac{C_{30}^{0} C_{20}^{10}}{C_{50}^{10}} + \frac{C_{30}^{1} C_{20}^{9}}{C_{50}^{10}} + \frac{C_{30}^{2} C_{20}^{8}}{C_{50}^{10}} + \frac{C_{30}^{3} C_{20}^{7}}{C_{50}^{10}} \end{aligned}$$

Có ít nhất 1 nữ: $X \geq 1$

$$\begin{aligned} P\left(X \geq 1\right) &= 1 - P\left(X < 1\right) \\ &= 1 - P\left(X = 0\right) \\ &= 1 - \frac{C_{30}^{0} C_{20}^{10}}{C_{50}^{10}} \end{aligned}$$

Định lý 3.6.6. Cho $X \sim H\left(N, N_A, n\right)$, đặt $p = \dfrac{N_A}{N}$

1. $E\left(X\right) = np$.
2. $Var\left(X\right) = npq\dfrac{N-n}{N-1}, q = 1 - p$.

Ví dụ: Một lớp gồm 50 sinh viên thi vấn đáp cuối khóa. Giáo viên phụ trách môn học cho 50 câu hỏi, trong đó có 40 câu hỏi dễ và 10 câu hỏi khó. Các sinh viên bốc thăm và chọn câu hỏi cho mình. Sau khi sinh viên cuối cùng của lớp bốc thăm xong thì giáo viên mới cho sinh viên biết nội dung câu hỏi tương ứng với số thứ tự sinh viên đó bốc.

Trước khi tiến hành bốc thăm thì xác suất các sinh viên bốc được câu hỏi dễ có phụ thuộc vào thứ tự bốc thăm không?

4.4. Phân phối Possion

Định nghĩa 4.4.1. Biến ngẫu nhiên X lấy giá trị nguyên, không âm được gọi là có phân phối Poisson với tham số λ nếu:

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!},\ k = 0, 1, \dots$$

Và ta kí hiệu:

$$X \sim P(\lambda)$$

Tính chất 4.4.2. Cho $X \sim P(\lambda)$

1. $\mathbb{E}(X) = \lambda$
2. $Var(X) = \lambda$

Chú ý:Biến ngẫu nhiên có phân phối Poisson thường được sử dụng để "đếm" số lần "xảy ra" của một sự kiện hay hiện tượng trong một khoảng thời gian.
Nếu $X \sim P(\lambda)$ là số lần xảy ra của một sự kiện nào đó trong khoảng thời gian T thì số lần xảy ra của sự kiện đó trong khoảng thời gian T_1 sẽ được mô tả bằng biến ngẫu nhiên Y với $Y \sim P\left(\lambda\frac{T_1}{T}\right)$

Ví dụ: Một trung tâm bưu điện nhận trung bình 150 cuộc điện thoại trong một giờ, tìm xác suất để trung tâm bưu điện này nhận không quá hai cuộc gọi trong một phút.

Giải

Gọi X là số cuộc gọi đến trung tâm bưu điện trong một giờ thì $X \sim P(150)$
Gọi Y là số cuộc gọi đến trung tâm bưu điện trong một phút thì $Y \sim P\left(150 \times \frac{1}{60}\right) = P(2.5)$
Khi đó:

$$\begin{aligned}P(Y \leq 2) &= P(Y = 0) + P(Y = 1) + P(Y = 2)\\ &= e^{-2.5}\left(\frac{2.5}{0!} + \frac{2.5^1}{1!} + \frac{2.5^2}{2!}\right)\end{aligned}$$

Tính chất 4.4.3. Xấp xỉ phân phối nhị thức bằng phương pháp Possion.
Nếu $X \sim B(n, p)$ với $n \geq 100, p \leq 0.01, np \leq 20$ thì ta có thể xấp xỉ phân phối của X bằng phân phối Poisson với tham số $\lambda = np$.

4.5. Phân phối đều.

Định nghĩa 4.5.1. Biến ngẫu nhiên liên tục X, nhận giá trị trong $[a, b]$ được gọi là có phân phối đều nếu:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & x \notin [a, b] \end{cases}$$

Ta kí hiệu: $X \sim U([a, b])$

Tính chất 4.5.2. Cho $X \sim U([a, b])$ thì:

1. $E(X) = \dfrac{a+b}{2}$
2. $Var(X) = \dfrac{(a-b)^2}{12}$

Ví dụ: Tại một trạm xe buýt khoảng cách giữa các tuyến liên tiếp của, một tuyến xe buýt bất kì là 15 phút. Các tuyến bắt đầu hoạt động từ 7 giờ sáng. Nếu một hành khách tới trạm xe buýt vào một thời điểm có phân phối đều từ 7 giờ tới 7 giờ 30 để bắt một tuyến xe buýt nào đó
Tính xác suất để anh ta đợi:
a) ít hơn hoặc bằng 5 phút
b) ít hơn hoặc bằng 10 phút
c) từ 6 đến 12 phút

Giải

X là thời điểm mà hành khách tới trạm.
$Y = (X - 7) \times 60$ là độ dài khoảng thời gian (phút) từ lúc 7 giờ tới lúc anh ta đến trạm
$Y \sim U([0, 30])$
a)

$$\begin{aligned} &P((Y = 0) \cup (10 \le Y \le 15) \cup (25 \le Y \le 30)) \\ =&P(Y = 0) + P(10 \le Y \le 15) + P(25 \le Y \le 30) \\ =&0 + (F_X(15) - F_X(10)) + (F_X(30) - F_X(25)) \end{aligned}$$

4.6. Phân phối chuẩn.

Định nghĩa 4.6.1. Biến ngẫu nhiên X nhận giá trị trên $\mathbb{R}$ được gọi là có phân phối chuẩn với tham số $\mu\,(\mu \in \mathbb{R})$ và $\sigma\,(\sigma > 0)$ nếu:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \times \exp\left(-\frac{1}{2} \times \frac{(x-\mu)^2}{\sigma^2}\right)$$

Và ta kí hiệu $X \sim \mathcal{N}(\mu, \sigma^2)$

Định lý 4.6.2. Cho $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$

1. $E\left(X\right) = \mu$
2. $Var\left(X\right) = \sigma^2$

Định lý 4.6.3. Cho $X_i \sim \mathcal{N}\left(\mu_i, \sigma_i^2\right), X, Y$ độc lập, $i = \overline{1, n}, a_i \in \mathbb{R}, b \in \mathbb{R}$

$$\begin{aligned} Y &= a_1X_1 + a_2X_2 + \ldots + a_nX_n + b \\ &\sim \mathcal{N}\left(a_1\mu_1 + a_2\mu_2 + \ldots + a_n\mu_n + b, a_1^2\sigma_1^2 + \ldots + a_n^2\sigma_n^2\right) \end{aligned}$$

Ví dụ: Cho

$$X_1 \sim \mathcal{N}\left(1, 2\right); X_2 \sim \mathcal{N}\left(0, 4\right)$$

X, Y độc lập, suy ra:

$$Y = 2X_1 + 3X_2 \sim \mathcal{N}\left(2 \times 1 + 3 \times 0, 2^2 \times 2 + 3^2 \times 4\right) = \mathcal{N}\left(2, 44\right)$$

Nếu

$$X = \frac{X - \mu}{\sigma} \Rightarrow Y \sim \mathcal{N}\left(\frac{1}{\sigma}\mu - \frac{\mu}{\sigma}, \frac{1}{\sigma^2} \times \sigma^2\right) = \mathcal{N}\left(0, 1\right)$$

Định nghĩa 4.6.4. Biến ngẫu nhiên X được gọi là có phân phối chuẩn tắc (hay chuẩn hóa) nếu$X \sim \mathcal{N}\left(0, 1\right)$. Ta kí hiệu hàm phân phối của biến ngẫu nhiên có phân phối chuẩn hóa là Φ

Chú ý: Nếu

$$\begin{aligned} X \sim \left(\mathcal{N}\mu, \sigma^2\right) &\Rightarrow Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}\left(0, 1\right) \\ &\Rightarrow P\left(X \le a\right) = P\left(\frac{X - \mu}{\sigma} \le \frac{a - \mu}{\sigma}\right) = P\left(Z \le \frac{a - \mu}{\sigma}\right) = \Phi\left(\frac{a - \mu}{\sigma}\right) \\ &\quad P\left(a < X \le b\right) = P\left(\frac{a - \mu}{\sigma} < X \le \frac{b - \mu}{\sigma}\right) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right) \end{aligned}$$

Phần V

Vector ngẫu nhiên

5.1. Một số khái niệm

Định nghĩa 5.1.1. Một bộ có thứ tự $(X_1, \ldots, X_n)$ của n biến ngẫu nhiên $X_1, \ldots, X_n$ được gọi là một vector ngẫu nhiên n chiều.

Định nghĩa 5.1.2. Phân phối của vector ngẫu nhiên
Độ đo $P_{X_1,\ldots X_n}$ thỏa:$P_{X_1\ldots X_n}(B) = P((X_1, \ldots, X_n) \in B)$
Trong đó: $B \in \mathcal{B}(\mathbb{R}^n)$ được gọi là phân phối xác suất của vector ngẫu nhiên $(X_1, \ldots, X_n)$

Định nghĩa 5.1.3. Hàm phân phối đồng thời
Hàm phân phối đồng thời của vector ngẫu nhiên $(X_1, \ldots, X_n)$ là hàm được định nghĩa như sau:

$$\begin{aligned} F_{X_1,\ldots,X_n} : \mathbb{R}^n &\longrightarrow [0,1] \\ (x_1, \ldots, x_n) &\longrightarrow P(X_1 \le x, \ldots X_n \le x_n) = P(\omega \in \Omega : X_1(\omega) \le x_1, \ldots, X_n \le x_n) \end{aligned}$$

5.2. Phân phối xác suất của vector ngẫu nhiên.

Ta sử dụng bảng sau:

$Y \backslash X$	x_1	x_2	x_3	$\ldots$	x_n	$f(.,y)$
y_1	$f(x_1, y_1)$					$f(.,y_1)$
y_2						$f(.,y_2)$
y_3						$f(.,y_3)$
$\ldots$						
y_n	$f(x_1, y_n)$				$f(x, y_n)$	$f(.,y_n)$
$f(x_1,.)$						1

Trong đó:

$$f(x,y) = \begin{cases} P(X = x_i, Y = y_i) & \text{nếu } (x,y) = (x_i, y_i) \\ 0 & \text{ngược lại} \end{cases}$$

Ví dụ: Tung ngẫu nhiên 3 đồng xu 1,2,3. Gọi X là số mặt ngửa xuất hiên ở hai đồng xu 1, 2, Y là số mặt ngửa xuất hiện ở cả ba đồng xu
Lập bảng phân phối xác suất (đồng thời) của (X, Y)

$Y\backslash X$	0	1	2	$f(.,y)$
0	1/8	0	0	1/8
1	1/8	2/8	0	3/8
2	0	2/8	1/8	3/8
3	0	0	1/8	1/8
$f(x,.)$	2/8	4/8	2/8	1

$$\begin{aligned} f(0,0) &= P(X=0,Y=0) \\ &= P(SSS) = \frac{1}{2}\times\frac{1}{2}\times\frac{1}{2} = \frac{1}{8} \\ f(0,1) &= P(SSN) = \frac{1}{8} \\ f(1,1) &= P(NSS) + P(SNS) = \frac{2}{8} \end{aligned}$$

5.3. Phân phối xác suất của biến ngẫu nhiên liên tục tuyệt đối.

Định nghĩa 5.3.1. Vector ngẫu nhiên $(X_1,\dots,X_n)$ được gọi là liên tục tuyệt đối nếu tồn tại hàm $f_{X_1,\dots,X_n}$không âm thỏa:

$$P[(X_1,\dots,X_n)\in B] = \iint_B f_{X_1,\dots,X_n}(x_1,\dots,x_n)\,dx_1\dots dx_n; \forall B\in\mathcal{B}(\mathbb{R}^2)$$

$$P(X_1\le x_1,\dots,X_n\le x_n) = \int_{-\infty}^{x_n}\dots\int_{-\infty}^{x_1} f_{X_1,\dots,X_n}(x_1,\dots,x_n)\,dx_1\dots dx_n$$

Hàm $f_{X_1,\dots,X_n}$ được gọi là hàm mật độ xác suất của vector ngẫu nhiên $(X_1,\dots,X_n)$
Chú ý:Mọi hàm$f(x,y)$ không âm và thỏa $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$ đều là hàm mật độ xác suất của một vector ngẫu nhiên (X,Y) nào đó.

Nếu ta biết $F_{X,Y}$ và $F_{X,Y}$ khả vi thì hàm mật độ xác suất:

$$f_{X,Y}(x,y) = \frac{\partial^2 F}{\partial x\partial y}(x,y)$$

Ví dụ: Cho vector ngẫu nhiên (X,Y) có hàm mật độ

$$f_{X,Y}(x,y) = \begin{cases} 1 & (x,y)\in[0,1]\times[0,1] \\ 0 & (x,y)\notin[0,1]\times[0,1] \end{cases}$$

Tính

$$P\left(\frac{1}{2}\le X+Y\le\frac{3}{2}\right)$$

Giải

Có:

$$\begin{aligned}
P\left(\frac{1}{2} \leq X+Y \leq \frac{3}{2}\right) &= P\left((X,Y) \in B\right) \\
&= \iint_B f_{X,Y}(x,y)\, dxdy \\
&= \int_0^{\frac{1}{2}} \int_{\frac{1}{2}-x}^{1} 1dydx + \int_{\frac{1}{2}}^{1} \int_0^{\frac{3}{2}-x} 1dydx \\
&= \int_0^{2} \left(x+\frac{1}{2}\right) dx + \int_{\frac{1}{2}}^{1} \left(\frac{3}{2}-x\right) dx \\
&= \left(\frac{x^2}{2}-\frac{1}{2}x\right)\Big|_0^2 + \left(\frac{3x}{2}-\frac{1}{2}x^2\right)\Big|_{\frac{1}{2}}^1 \\
&= \frac{3}{4}
\end{aligned}$$

Ví dụ:

$$f_{X,Y}(x,y) = \begin{cases} e^{-(x+y)} & x,y \geq 0 \\ 0 & \text{ngược lại} \end{cases}$$

Tính $P(X \geq Y \geq 2)$

Giải

$$\begin{aligned}
\iint_B f_{X,Y}(x,y)\, dxdy &= \int_2^{\infty} \int_2^{x} e^{-(x+y)} dydx \\
&= \int_2^{\infty} \left(\left(-e^{-x-y}\right)\Big|_{y=2}^{y=x}\right) dx \\
&= \lim_{m\to\infty} \left(-e^{-x-2} + \frac{1}{2}e^{-2x}\right)\Big|_{x=2}^{x=m} \\
&= \frac{1}{2}e^{-4}
\end{aligned}$$

5.4. Liên hệ giữa phân phối đồng thời và phân phối thành phần, sự độc lập của các biến ngẫu nhiên.

Nếu biết dược phân phối xác suất của một biến ngẫu nhiên thì ta cũng xác định được phân phối của các biến ngẫu nhiên thành phần. Các phân phối này gọi là phân phối lề

Định nghĩa 5.4.1. Trường hợp vector ngẫu nhiên rời rạc.
Giả sử vector ngẫu nhiên (X,Y) có hàm xác suất $f(x,y)$ thì hàm xác suất lề là:

$$\begin{aligned}
P(X=x) &= f_X(x) \\
&= \sum_y f(x,y) \\
&= \sum_y P(X=x, Y=y)
\end{aligned}$$

$$f_Y(y) = \sum_x f(x,y)$$

Định nghĩa 5.4.2. Trường hợp vector ngẫu nhiên liên tục tuyệt đối.
Cho vector ngẫu nhiên (X, Y) có hàm mật độ xác suất $f_{X,Y}(x,y)$, hàm mật độ lề, phân phối lề được xác định như sau:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$$
$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$$
$$F_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(u,v)\,dvdu$$
$$F_Y(y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(u,v)\,dudv$$

Ví dụ: cho vector ngẫu nhiên (X, Y) có:

$$f_{X,Y}(x,y) = \begin{cases} e^{-(x+y)} & ,x,y \geq 0 \\ 0 & ,\text{ngược lại} \end{cases}$$

Tìm f_X ,f_Y,F_X,F_Y

Định nghĩa 5.4.3. Các biến ngẫu nhiên $(X_1, \dots, X_n)$ được định nghĩa trên cùng không gian xác suất$(\Omega, \mathcal{F}, P)$ được gọi là độc lập khi:

$$P(X_1 \in B_1, \dots, X_n \in B_n) = P(X_1 \in B_1) \dots P(X_n \in B_n); \forall B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$$

Chú ý:

- Nếu $X_1, \dots, X_n$ là độc lập thì mọi con của nó là độc lập.
- Họ bất kì các biến ngẫu nhiên là độc lập $\Leftrightarrow$ mọi họ con hữu hạn là độc lập.

Định lý 5.4.4. Cho $X_1, \dots, X_n$ là các biến ngẫu nhiên độc lập được định nghĩa trên cùng không gian xác suất $(\Omega, \mathcal{F}, P)$;$X_1, \dots, X_n$ độc lập

$$\Leftrightarrow f_{X_1,\dots,X_n} = f_{X_1} \dots f_{X_n}$$
$$\Leftrightarrow f_{X_1,\dots,X_n}(x_1, \dots, x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n)$$

X_1 : biến ngẫu nhiên liên tục bất kì

$$X_2 = X_1\,(X_2(\omega) = X_1(\omega), \forall \omega \in \Omega)$$
$$X = (X_1, X_2) = 1$$
$$0 = \iint_B f_{X,Y}(x,y)$$

Định lý 5.4.5. Nếu X_1,X_2,. . .,X_n là các biến ngẫu nhiên độc lập. g_1,g_2,. . .,g_n là các hàm đo được thì $g_1(X_1)$,$g_2(X_2)$,. . .,$g_n(X_n)$là các biến ngẫu nhiên độc lập.

Ví dụ: Cho vectơ ngẫu nhiên $Z = (X, Y)$, hàm mật độ đồng thời $f(x, y) = \dfrac{3(2-2x-y)}{2}$ trong miền giới hạn bởi $x = 0, y = 0, y = 2-2x$. Tìm f_X. Tính $P\left(X > \dfrac{1}{2}\right)$, $P(X > Y)$.

Giải

$$f(x, y) = \frac{3(2-2x-y)}{2}$$

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{+\infty} f(x, y)\,dy \\
&= \int_{-\infty}^{+\infty} \frac{3(2-2x-y)}{2} dy \\
&= \frac{3}{2}\int_0^{2-2x} (2-2x-y)\,dy \\
&= \frac{3}{2}\left(2y - 2xy - \frac{y^2}{2}\right)\Bigg|_{y=0}^{y=2-2x} \\
&= \frac{3}{4}(2-2x)^2 \quad (0 \le x \le 1)
\end{aligned}$$

$$\begin{aligned}
P\left(X > \frac{1}{2}\right) &= \int_{\frac{1}{2}}^{+\infty} f_X(x)\,dx \\
&= \frac{3}{4}\int_{\frac{1}{2}}^{1} (2-2x)^2\,dx
\end{aligned}$$

Đặt $t = 2-2x \Rightarrow dt = -2dx$, ta có:

$$\begin{aligned}
P\left(X > \frac{1}{2}\right) &= \frac{3}{8}\int_0^1 t^2 dt \\
&= \frac{1}{8}
\end{aligned}$$

Ta cũng có:

$$\begin{aligned}
\mathbb{P}(X > Y) &= \int\limits_{\{(x,y):x>y\}} f(x, y)\,dxdy \\
&= \int\limits_0^{\frac{2}{3}} \int_y^{1-\frac{y}{2}} \frac{3}{2}(2-2x-y)\,dxdy \\
&= \frac{3}{2}\int\limits_0^{\frac{2}{3}} \left[-x^2 + (2-y)\,x\right]_y^{1-\frac{y}{2}} dy
\end{aligned}$$

Ví dụ: Véctơ ngẫu nhiên $Z = (X, Y)$ có hàm mật độ đồng thời là hằng số xác định trên miền $0 \le y \le x$, $y \le 1$, $x \le 2$.

a) Xác định $f_{X,Y}$.

b) Tìm f_X, f_Y.

c) Tính $P(X > 1)$,$P\left(X \le \frac{1}{2}\right)$,$P(X + Y < 1)$.

Giải

a) Xác định hàm mật độ đồng thời:

$$f(x, y) = \begin{cases} a & (x, y) \in D \\ 0 & (x, y) \notin D \end{cases}$$

Miền $D = \{(x, y) : y \le x \le 2; 0 \le y \le 1\}$

$$\begin{aligned} 1 &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)\, dxdy \\ \Leftrightarrow 1 &= \int_0^1 \int_{-\infty}^{+\infty} adxdy \\ \Leftrightarrow 1 &= \int_0^1 ax_y^2 dy \\ \Leftrightarrow 1 &= \int_0^1 (2a - ay)\, dy \\ \Leftrightarrow 1 &= \left(2ay - a\frac{y^2}{2}\right)\Bigg|_0^1 \\ \Leftrightarrow 1 &= 2a - \frac{a}{2} \\ \Leftrightarrow a &= \frac{2}{3} \end{aligned}$$

b) Xác định f_X, f_Y.

− $0 \le x \le 1$:

$$\begin{aligned} f_X(x) &= \int_0^x f_{X,Y}(x, y)\, dy \\ &= \int_0^x \frac{2}{3} dy \\ &= \frac{2x}{3} \end{aligned}$$

− $1 < x \leq 2$:

$$
\begin{aligned}
f_X(x) &= \int_0^1 f_{X,Y}(x,y)\,dy \\
&= \int_0^1 \frac{2}{3} dy \\
&= \frac{2}{3}
\end{aligned}
$$

$$
\begin{aligned}
f_Y(y) &= \int_y^2 f_{X,Y}(x,y)\,dy \\
&= \int_y^2 \frac{2}{3} dx \\
&= \frac{2}{3}(2-y), \quad (0 \leq y \leq 1)
\end{aligned}
$$

c) Tính $P(X+Y<1)$.

$$
\begin{aligned}
P(X+Y<1) &= \int_{\{(x,y):x+y<1\}} f_{X,Y}(x,y)\,dxdy \\
&= \int_0^{1/2}\int_y^{1-y} \frac{2}{3} dxdy \\
&= \int_0^{1/2} \frac{2}{3}(1-2y)\,dy \\
&= \frac{2}{3}\left(y-y^2\right)\Big|_0^{\frac{1}{2}} \\
&= \frac{1}{6}
\end{aligned}
$$

5.5. Hàm của vectơ ngẫu nhiên:

Định nghĩa 5.5.1. Biến ngẫu nhiên liên tục.

Định lý 5.5.2. Giả sử g là 1 ánh xạ khả nghịch định nghĩa trên 1 tập mở của $\mathbb{R}^n$ chứa giá trị của vectơ ngẫu nhiên$(X_1, X_2, X_3, \ldots, X_n)$. Trong đó,$(X_1, X_2, X_3, \ldots, X_n)$là vectơ liên tục tuyệt đối với hàm mật độ $f_{X_1,X_2,X_3,\ldots,X_n}$. Giả sử $h = g^{-1}$là ánh xạ khả vi liên tục với Jacobian$\neq 0$. Khi đó vectơ ngẫu nhiên $(Y_1, Y_2, Y_3, \ldots, Y_n) = g(X_1, X_2, X_3, \ldots, X_n)$có hàm mật độ là:

$$
\begin{aligned}
f_{X_1,X_2,X_3,\ldots,X_n}(y_1,y_2,y_3,\ldots,y_n) =& f_{X_1,X_2,X_3,\ldots,X_n}(h(y_1,y_2,y_3,\ldots,y_n)) \\
&\times \left|\det \nabla h(y_1,y_2,y_3,\ldots,y_n)\right|
\end{aligned}
$$

Trong đó:

$$\nabla h\left(y_1, y_2, y_3, \ldots, y_n\right) = \begin{pmatrix} \dfrac{\partial x_1}{\partial y_1} & \cdots & \dfrac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial x_n}{\partial y_1} & \cdots & \dfrac{\partial x_n}{\partial y_n} \end{pmatrix}$$

$$\begin{aligned} g : U &\longrightarrow \mathbb{R}^n \\ (x_1, x_2, x_3, \ldots, x_n) &\longmapsto (y_1, y_2, y_3, \ldots, y_n) \end{aligned}$$

$$\begin{aligned} h : \mathbb{R}^n &\longrightarrow U \\ (y_1, y_2, y_3, \ldots, y_n) &\longmapsto (x_1, x_2, x_3, \ldots, x_n) \end{aligned}$$

Ví dụ: Cho X, Y là các biến ngẫu nhiên độc lập, cùng phân phối, $X \sim U(0,1)$. Tìm hàm mật độ đồng thời của $(U, V) = (X, XY)$

Giải

$$f_X(x) = \begin{cases} 1 & x \in (0,1) \\ 0 & x \notin (0,1) \end{cases}$$

Vì X, Y độc lập, cùng phân phối nên:

$$\begin{aligned} f_{X,Y}(x,y) &= f_X(x) f_Y(y) \\ &= \begin{cases} 1 & x \in (0,1) \times (0,1) \\ 0 & x \notin (0,1) \times (0,1) \end{cases} \end{aligned}$$

Xác định:

$$\begin{aligned} g : (0,1) \times (0,1) &\longrightarrow \mathbb{R}^n \\ (x,y) &\longmapsto (u,v) = (x, xy) \end{aligned}$$

$$\begin{aligned} h : \mathbb{R}^n \setminus \{0\} \times \mathbb{R} &\longrightarrow \mathbb{R}^n \\ (x,y) &\longmapsto \left(u, \frac{v}{u}\right) \end{aligned}$$

$$
\begin{aligned}
\nabla h(u,v) &= \begin{pmatrix} 1 & 0 \\ -\dfrac{v}{u^2} & \dfrac{1}{u} \end{pmatrix} \\
\Rightarrow |\det \nabla h(u,v)| &= \left|\frac{1}{u}\right| \\
\Rightarrow f_{U,V}(u,v) &= f_{X,Y}(h(u,v))\left|\frac{1}{u}\right| \\
&= f_{X,Y}\left(u,\frac{v}{u}\right)\frac{1}{|u|}
\end{aligned}
$$

$$
f_{u,v}(u,v) = \begin{cases} \frac{1}{|u|} & 0<u<1, 0<\frac{v}{u}<1 \\ 0 & \text{ngược lại} \end{cases}
$$

Ví dụ: Cho X, Y là các biến ngẫu nhiên độc lập, cùng phân phối, $E \sim Exp(1)$. CMR: $X/(X+Y)$và $X+Y$ là 2 biến ngẫu nhiên độc lập. Tìm phân phối của chúng.

Giải

$$X \sim Exp(\lambda)$$

$$
f_X(\lambda) = \begin{cases} \lambda.e^{-\lambda x} & x \geq 0 \\ 0 & x<0 \end{cases}
$$

Đặt: $U = \dfrac{X}{X+Y}$ và $V = X+Y$. Ta có:

$$
\begin{aligned}
f_{X,Y}(x,y) &= f_X(x) f_Y(y) \\
&= \begin{cases} e^{-x}\lambda.e^{-y} & x \geq 0,\, y \geq 0 \\ 0 & \text{ngược lại} \end{cases}
\end{aligned}
$$

Xét:

$$
\begin{aligned}
g: (0,+\infty)\times(0,+\infty) &\longrightarrow \mathbb{R}^2 \\
(x,y) &\longmapsto (u,v) = \left(\frac{x}{x+y}, x+y\right)
\end{aligned}
$$

cuu duong than cong . com

$$
\begin{aligned}
h: \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\
(u,v) &\longmapsto (x,y) = (uv, v-uv)
\end{aligned}
$$

$$\nabla h(u,v) = \begin{pmatrix} v & u \\ -v & 1-u \end{pmatrix}$$

$$\Rightarrow |\det \nabla h(u,v)| = |-vu - v + vu| = |-v| = |v|$$

$$\Rightarrow f_{U,V}(u,v) = f_{X,Y}(h(u,v))\,|v|$$

$$= f_{X,Y}(uv, v-uv)\,|v|$$

$$= \begin{cases} e^{-uv}.e^{-(v-uv)} & uv \geq 0,\ v-uv \geq 0 \\ 0 & \text{ngược lại} \end{cases}$$

$$= \begin{cases} e^{-v}\,|v| & uv \geq 0,\ v-uv \geq 0 \\ 0 & \text{ngược lại} \end{cases}$$

$$v(1-u) \geq 0 \Leftrightarrow \left[\begin{array}{l} \begin{cases} v \geq 0 \\ 1-u \geq 0 \end{cases} \\ \begin{cases} v \leq 0 \\ 1-u \leq 0 \end{cases} \end{array}\right.$$

$$\Leftrightarrow \left[\begin{array}{l} \begin{cases} v \geq 0 \\ u \leq 1 \end{cases} \\ \begin{cases} v \leq 0 \\ u \geq 1 \end{cases} \end{array}\right.$$

$$f_{u,v}(u,v) = \begin{cases} e^{-v}v & v \geq 0, 0 \leq u \leq 1 \\ 0 & \text{ngược lại} \end{cases}$$

$$= 1_{\{0 \leq u \leq 1\}} \times e^{-v} v_{\{v \geq 0\}}$$

$$= f_U(u) \times f_V(v)$$

Trong đó:

$$f_U(u) = \begin{cases} 1 & 0 \leq u \leq 1 \\ 0 & \text{ngược lại} \end{cases}$$

$$f_V(v) = \begin{cases} e^{-v}v & v \geq 0 \\ 0 & \text{ngược lại} \end{cases}$$

Vậy U, V độc lập.$U \sim U(0,1)$, $V \sim \Gamma(2,1)$.

5.5. Kỳ vọng

Định lý 5.5.1. Nếu $X_1, X_2, \ldots, X_n$ là các biến ngẫu nhiên rời rạc, $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ là một hàm đo được thì:

$$\begin{aligned} E\left(g\left(X_1, X_2, \ldots, X_n\right)\right) &= \sum g\left(x_1, x_2, \ldots, x_n\right) f_{X_1, X_2, \ldots, X_n}\left(x_1, x_2, \ldots, x_n\right) \\ &= \sum g\left(x_1, x_2, \ldots, x_n\right) P\left(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\right) \end{aligned}$$

Chú ý: điều này đúng khi và chỉ khi

$$\left[\begin{array}{l} g\left(x_1, x_2, \ldots, x_n\right) \geq 0 \\ \text{chuỗi bên phải liên tục tuyệt đối} \end{array}\right.$$

Ví dụ: Cho biến ngẫu nhiên thỏa:

$$P\left(X = n\right) = \left(\frac{1}{2}\right)^n, n = 1, 2, \ldots$$

Đặt $Y = g\left(X\right)$, trong đó:

$$g\left(n\right) = \left(-1\right)^{n+1} \cdot \frac{2^n}{n}, n = 1, 2, \ldots$$

Tính $E\left(Y\right)$?

Giải

Tính theo định lý:

$$\begin{aligned} E\left(Y\right) &= \sum_{i=1}^{+\infty} g\left(x_i\right) f_X\left(x_i\right) \\ &= \sum_{n=1}^{+\infty} g\left(n\right) f_X\left(n\right) = \sum_{n=1}^{+\infty} \left(-1\right)^{n+1} \cdot \frac{2^n}{n} \cdot \left(\frac{1}{2}\right)^n \\ &= \sum_{n=1}^{+\infty} \left(-1\right)^{n+1} \cdot \frac{1}{n} \end{aligned}$$

Chuỗi này hội tụ có điều kiện (chuỗi đan dấu)
Cách khác:

$$\begin{aligned} Y\left(n\right) &= g\left(X\right)\left(n\right) = g\left(X\left(n\right)\right) = g\left(n\right) \\ &= \left(-1\right)^{n+1} \cdot \frac{2^n}{n} \\ &= \begin{cases} -\frac{2^n}{n} & \text{nếu } n \text{ chẵn} \\ \frac{2^n}{n} & \text{nếu } n \text{ lẻ} \end{cases} \end{aligned}$$

Do:

$$\begin{aligned} Y^+(n) &= \max\{Y, 0\} \\ Y^-(n) &= \min\{-Y, 0\} \end{aligned}$$

Ta có:

$$\begin{aligned} Y^+(n) &= \begin{cases} 0 & \text{nếu } n \text{ chẵn} \\ \frac{2^n}{n} & \text{nếu } n \text{ lẻ} \end{cases} \\ E\left(Y^+\right) &= \sum_{k=0}^{+\infty} \frac{2^{2k+1}}{2k+1} \left(\frac{1}{2}\right)^{2k+1} \\ &= \sum_{k=0}^{+\infty} \frac{1}{2k+1} \\ &= 1 + \frac{1}{3} + \frac{1}{5} + \cdots = +\infty \end{aligned}$$

Chọn chuỗi $\sum_{k=1}^{+\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots$ phân kì. Ta có:

$$\lim \frac{\frac{1}{2k+1}}{\frac{1}{k}} = \frac{1}{2} > 0 \Rightarrow \text{Chuỗi phân kì}$$

$$\begin{aligned} Y^- &= \begin{cases} \frac{2}{n} & \text{nếu } n \text{ chẵn} \\ 0 & \text{nếu } n \text{ lẻ} \end{cases} \\ E\left(Y^-\right) &= \sum_{k=1}^{+\infty} \frac{2^{2k}}{2k} \left(\frac{1}{2}\right)^{2k} \\ &= \sum_{k=1}^{+\infty} \frac{1}{2k} \\ &= \frac{1}{2} + \frac{1}{4} + \frac{1}{6} \cdots = +\infty \end{aligned}$$

Vậy $E(Y) = E(Y^+) - E(Y^-)$ không tồn tại.

Định nghĩa 5.5.2. Cho $X_1, X_2, \ldots, X_n$ là các biến ngẫu nhiên liên tục tuyệt đối với hàm mật độ đồng thời f_{X_i} và $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ là một hàm đo được thì:

$$E\left(g\left(X_1, X_2, \ldots, X_n\right)\right) = \int\limits_{\mathbb{R}^n} g\left(x_1, x_2, \ldots, x_n\right) f_{X_1, X_2, \ldots, X_n}\left(x_1, x_2, \ldots, x_n\right) dx$$

Chú ý: điều này đúng nếu

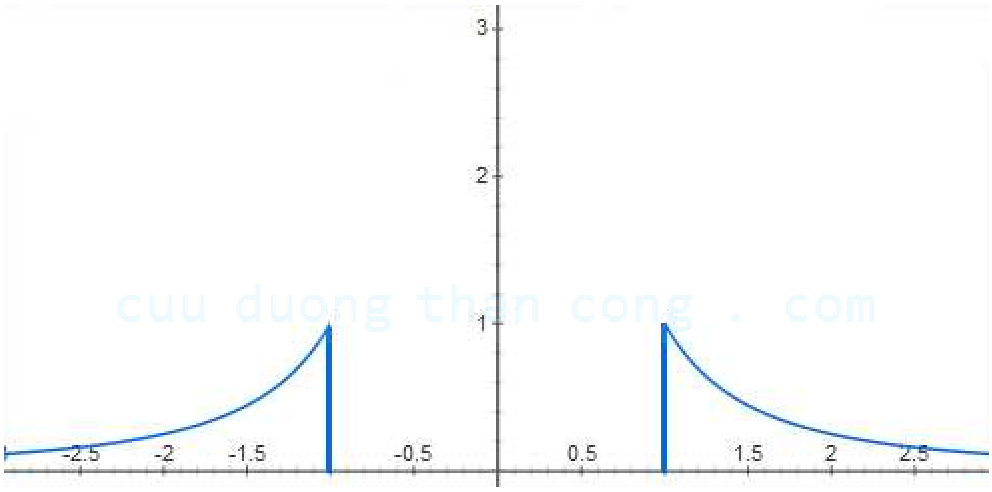
$$\Leftrightarrow \left[\begin{array}{l} g \geq 0 \\ \text{tích phân bên phải hội tụ tuyệt đối} \end{array} \right.$$

Chú ý: kỳ vọng có thể không tồn tại, tồn tại hữu hạn hoặc tồn tại và bằng $\pm\infty$

Ví dụ: Cho biến ngẫu nhiên X có hàm mật độ xác suất:

$$f_X(x) = \begin{cases} \frac{1}{2x^2} & \text{nếu } |x| \geq 1 \\ 0 & \text{nếu } |x| < 1 \end{cases}$$

Tính $E(X)$



Giải

$$\begin{aligned} E\left(X^+\right) &= \int_{-\infty}^{+\infty} x \frac{1}{2x^2} dx = \int_1^{+\infty} \frac{1}{2x} dx \\ &= \lim_{t \to +\infty} \int_1^t \frac{1}{2x} dx = \lim_{t \to +\infty} \frac{1}{2} \ln x|_{x=1}^{x=t} \\ &= \frac{1}{2} \lim_{t \to +\infty} \ln t = +\infty \\ E\left(X^-\right) &= \int_{-\infty}^{+\infty} -x \frac{1}{2x^2} dx = -\int_{-\infty}^{-1} \frac{1}{2x} dx \\ &= \int_{-1}^{-\infty} \frac{1}{2x} dx = \lim_{t \to -\infty} \int_{-1}^{t} \frac{1}{2x} dx \\ &= \lim_{t \to -\infty} \frac{1}{2} \ln x|_{x=1}^{x=t} = -\infty \end{aligned}$$

Vậy không tồn tại $E(X)$.

5.6. Moment

Định nghĩa 5.6.1. Moment cấp k của biến ngẫu nhiên X được định nghĩa là $\alpha_k = E(X_k)$ (k không nhất thiết nguyên) nếu kỳ vọng bên phải tồn tại.

$$\alpha_k = E(X_k) = \begin{cases} \sum x_i^k f_X(X_i) \\ \int x^k f_X(X)\, dx \end{cases}$$

α_1 thường được kí hiệu là m

Định nghĩa 5.6.2. Moment trung tâm cấp $k\,(k > 0)$ của biến ngẫu nhiên X được định nghĩa là:

$$\beta_k = E(X - m)^k \quad \text{nếu } m \text{tồn tại và hữu hạn}$$

nếu kỳ vọng phía bên phải tồn tại. β_2 thường được kí hiệu là σ^2.

Ví dụ: Cho X và Y là hai biến ngẫu nhiên độc lập, cùng phân phối và X có hàm mật độ xác suất:

$$f_X(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Đặt $Z = \max\{X, Y\}$. Tính $\alpha_1, \alpha, \beta_2$

Giải

$$\begin{aligned}
\alpha_1 = m = E(Z) \\
&= E(g(X,Y)) \\
&= \int_0^{+\infty}\int_0^{+\infty} \max\{x,y\}\, f_{X,Y}(x,y)\, dxdy \\
&= \int_0^{+\infty}\int_0^{+\infty} \max\{x,y\}\, e^{-(x+y)} dxdy
\end{aligned}$$

5.7. Tính chất của kỳ vọng.

1. Nếu $X_1, \ldots, X_n$ là các biến ngẫu nhiên độc lập thì $E(X_1, ..., X_n) = E(X_1) ... E(X_n)$.
2. Cho X là biến ngẫu nhiên, $a \in \mathbb{R}$, khi đó $E(aX) = aE(X)$.
3. Nếu X_1, X_2 được định nghĩa trên cùng không gian xác suất $(\Omega, \mathcal{F}, \mathbb{P})$ và $X_1 \leq X_2$, $E(X_1), E(X_2)$ tồn tại thì $E(X_1) \leq E(X_2)$
4. Nếu $X \geq 0$ và $E(X) = 0$ thì $P(X = 0) = 1$
5. Nếu $Var(X) = 0$ thì $P(X) = E(X) = P(X = m) = 1$
6. Nếu $X_1, \ldots, X_n$ là các biến ngẫu nhiên thì $E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n)$

7. Nếu biến ngẫu nhiên X có $E(X) < \infty, Var(X) = \sigma^2$ tồn tại thì $Var(aX+b) = a^2\sigma^2, \forall a, b \in \mathbb{R}$

8. Nếu $X_1, \ldots, X_n$ là các biến ngẫu nhiên độc lập thì $Var(X_1 + \cdots + X_n) = Var(X_1) + \cdots + Var(X_n)$
   Hệ quả:Nếu $X_1, \ldots, X_n$ là các biến ngẫu nhiên độc lập có kỳ vọng hữu hạn và phương sai $\sigma_1^2, \ldots, \sigma_n^2$ thì $Var(a_1X_1 + \cdots + a_nX_n + b) = a_1^2\sigma_1^2 + \cdots + a_n^2\sigma_n^2$

9. Các moment trung tâm $\beta_1, \ldots, \beta_n$ có thể được biểu diễn thông qua các moment $\alpha_1, \ldots, \alpha_n$ với điều kiện $\alpha_1, \ldots, \alpha_{n-1}$ hữu hạn, α_n tồn tại.

$$\begin{aligned} E(X-m)^n &= E\left(\sum_{k=0}^{n} C_n^k (-1)^{n-k} X^k m^{n-k}\right) \\ &= \sum_{k=0}^{n} C_n^k (-1)^{n-k} E\left(X^k\right) m^{n-k} \\ &= \sum_{k=0}^{n} C_n^k((-1)\, m^{n-k}\alpha_k \end{aligned}$$

10. Nếu $E\left(X^k\right) < \infty$ thì $\forall j < k, E\left(X^j\right)$ tồn tại và hữu hạn

5.8. Sự tương quan.

Định nghĩa 5.8.1. Nếu X_1, X_2 là các biến ngẫu nhiên, ta định nghĩa moment đồng thời của X_1, X_2 như sau:

$$\alpha_{jk} = E\left(X_1^j X_2^k\right) \quad j, k > 0$$

và moment trung tâm đồng thời

$$\beta_{jk} = E\left[\left(X_1 - E(X_1)^j\right)(X_2 - E(X_2))^k\right]$$

$$\begin{aligned} \beta_{11} &= E\left[(X_1 - E(X_1))(X_2 - \mathbb{E}(X_2))\right] \\ &= E\left[X_1X_2 - X_1E(X_2) - X_2E(X_1) + E(X_1)E(X_2)\right] \\ &= E(X_1X_2) - E(X_1)E(X_2) \end{aligned}$$

được gọi là hiệp phương sai của hai biến ngẫu nhiên X_1, X_2 và kí hiệu là $\mathbb{C}ov(X_1, X_2)$

Mệnh đề:

1. $\mathbb{C}ov(X, Y) = \mathbb{C}ov(Y, X)$

2. $\mathbb{C}ov(X, X) = \mathbb{V}ar(X)$

3. $\mathbb{C}ov(aX, Y) = a\mathbb{C}ov(X, Y)$

4. $\mathbb{C}ov\left(\sum_{i=1}^{m} X_i, \sum_{j=1}^{n} Y_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} \mathbb{C}ov(X_i, Y_j)$

Định lý 5.8.2. Nếu $E(X_1, X_2)$ tồn tại, $E(X_1)E(X_2) < \infty$ và X_1, X_2 độc lập thì $Cov(X_1, X_2) = 0$
Lưu ý: chiều ngược lại không đúng (chỉ đúng với phân phối chuẩn)

Định lý 5.8.3. Giả sử $E\left(X_1^2\right)E\left(X_2^2\right) < \infty$ thì $E\left(X_1, X_2\right)$ tồn tại và $\left(E\left|X_1X_2\right|\right)^2 < E\left(X_1^2\right)E\left(X_2^2\right)$

Định nghĩa 5.8.4. Giả sử $E\left(X_1^2\right) < \infty, \mathbb{E}(X_2) < \infty$ và $Var\left(X_1\right) = \sigma_1^2 > 0, Var\left(X_2\right) = \sigma_2^2 > 0$ thì đại lượng

$$f\left(X_1, X_2\right) = \frac{\mathbb{C}ov\left(X_1, X_2\right)}{\sigma_1\sigma_2}$$

được gọi là sự tương quan của hai biến ngẫu nhiên X_1, X_2

Định lý 5.8.5.

$$-1 \leq f\left(X_1, X_2\right) \leq 1$$

5.9. Xác suất có điều kiện.

Định nghĩa 5.9.1. Cho X, Y là hai biến ngẫu nhiên, nếu X, Y là các biến ngẫu nhiên rời rạc, hàm xác suất có điều kiện của Y khi biết X nhận giá trị x, ký hiệu $f_{Y|X}\left(y|x\right)$ được định nghĩa là:

$$f_{Y|X}\left(y|x\right) = \frac{f_{X,Y}\left(x,y\right)}{f_X\left(x\right)}, \text{ nếu } f_X\left(x\right) > 0$$

Định nghĩa 5.9.2. Cho X, Y là hai biến ngẫu nhiên liên tục với hàm mật độ đồng thời $f_{X,Y}\left(x,y\right)$, hàm mật độ xác suất có điều kiện của Y khi biết X nhận giá trị x, ký hiệu $f_{Y|X}\left(y|x\right)$ được định nghĩa là:

$$f_{Y|X}\left(y|x\right) = \frac{f_{X,Y}\left(x,y\right)}{f_X\left(x\right)}, \text{ nếu } f_X\left(x\right) > 0$$

Chú ý: Từ định nghĩa trên ta suy ra:

– Nếu X, Y là biến ngẫu nhiên rời rạc thì:

$$\begin{aligned} P\left(Y \in B|X = x\right) &= \sum_{y \in B} P\left(Y = y|X = x\right) \\ &= \sum_{y \in B} f_{Y|X}\left(y|x\right) \end{aligned}$$

– Nếu X, Y là biến ngẫu nhiên liên tục thì:

$$P\left(Y \in B|X = x\right) = \int\limits_B f_{Y|X}\left(y|x\right)dy$$

Trong một số trường hợp ta không chỉ rõ $f_{Y|X}$ nhưng cho biết phân phối có điều kiện của Y khi biết $X = x$ và ký hiệu $Y|X = x \sim B\left(n,n\right)/U\left(0,x\right)$

Ví dụ: Cho X, Y là hai biến ngẫu nhiên có hàm mật độ đồng thời:

$$f_{X,Y}(x,y) = \begin{cases} x+y & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{nơi khác} \end{cases}$$

Giải

$$P\left(X < \frac{1}{4} \middle| Y = \frac{1}{3}\right) = \int\limits_{\{x<\frac{1}{4}\}} f_{X|Y}(x,y)dy$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx = \int_0^1 (x+y)dx = yx + \left.\frac{x^2}{2}\right|_0^1 = y + \frac{1}{2}$$

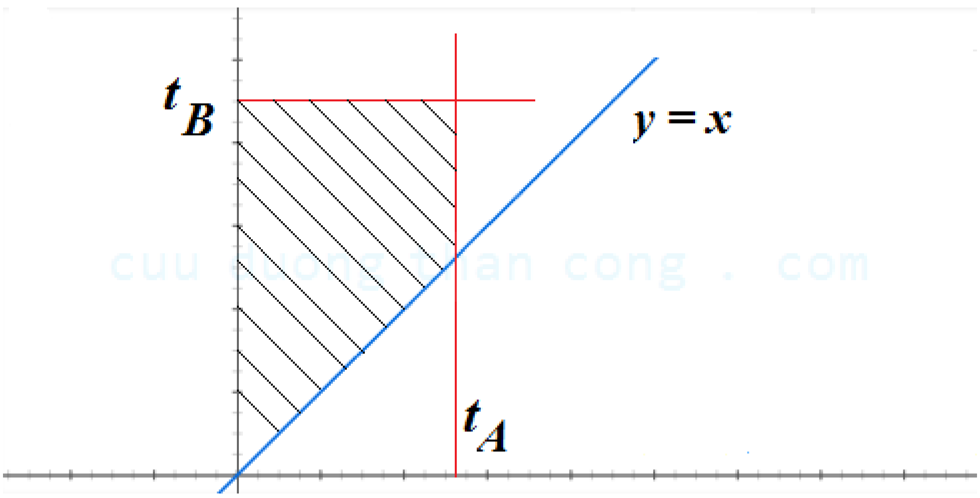
$$\Rightarrow f_{X|Y}\left(x \middle| \frac{1}{3}\right) = \frac{x+y}{\frac{1}{3}+\frac{1}{2}} = \frac{6(x+y)}{5} = \frac{6(x+\frac{1}{3})}{5}$$

$$\Rightarrow P\left(X < \frac{1}{4} \middle| Y = \frac{1}{3}\right) = \int_0^{\frac{1}{4}} \frac{6}{5}\left(x + \frac{1}{3}\right)dx = \frac{6}{5}\left(\frac{x^2}{2} + \frac{1}{3}x\Big|_0^{\frac{1}{4}}\right)$$

$$= \frac{6}{5}\left(\frac{1}{32} + \frac{1}{12}\right) = \frac{11}{80}$$

Ví dụ: Một người tới trạm xe buýt vào thời điểm $t = 0$, có hai tuyến xe buýt A và B đi qua trạm này. Thời điểm đến trạm của tuyến xe buýt A là $X \sim \mathcal{U}(0, t_A)$, thời điểm đến trạm của tuyến xe buýt B có phân phối đều $Y \sim \mathcal{U}(0, t_B)$, với $t_A \le t_B$, Biết rằng các xe tới trạm một cách độc lập với nhau. Tính xác suất A tới trước. $P(X \le Y)$

Giải



$$X \sim \mathcal{U}(0, t_A) \Rightarrow f_X(x) = \begin{cases} \frac{1}{t_A} & x \in [0, t_A] \\ 0 & \text{ngược lại} \end{cases}$$

$$X \sim \mathcal{U}(0, t_B) \Rightarrow f_Y(y) = \begin{cases} \frac{1}{t_B} & y \in [0, t_B] \\ 0 & \text{ngược lại} \end{cases}$$

Do X, Y độc lập nên

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \begin{cases} \frac{1}{t_B}\frac{1}{t_A} & x \in [0, t_A], y \in [0, t_B] \\ 0 & \text{ngược lại} \end{cases}$$

$$\Rightarrow P(X \le Y) = P((X,Y) \in B) \quad \left(B = \left\{(x,y) \in \mathbb{R}^2 \mid 0 \le x \le \tfrac{1}{t_A}, 0 \le y \le \tfrac{1}{t_B}\right\}\right.$$

$$\Rightarrow P(X \le Y) = \int\limits_B f_{X,Y}(x,y)\, dA = \int_0^{t_A}\int_0^{t_B} \frac{1}{t_A t_B} dy dx$$

$$= \frac{1}{t_A t_B}\int_0^{t_A} y|_x^{t_B}\, dx = \frac{1}{t_A t_B}\int_0^{t_A}(t_B - x)\, dx$$

$$= \frac{1}{t_A t_B}\left(t_B x - \frac{x^2}{2}\Big|_0^{t_A}\right) = \frac{1}{t_A t_B}\left(t_B t_A - \frac{t_A^2}{2}\right) = 1 - \frac{t_A}{2t_B}$$

Cách 2:

Như vậy

$$P(X \le Y | X = x) = \int_x^{t_B} \frac{1}{t_B} dy = \frac{1}{t_B}\, y|_x^{t_B} = 1 - \frac{x}{t_B}$$

$$P(X \le Y) = \int_0^{t_A} P(X \le Y | X = x) f_X(x)\, dx$$

$$= \int_0^{t_A}\left(1 - \frac{x}{t_B}\right)\cdot\frac{1}{t_A} dx$$

$$= \frac{1}{t_A}\left(x - \frac{x^2}{2t_B}\Big|_0^{t_A}\right)$$

$$= \frac{1}{t_A}\left(t_A - \frac{t_A^2}{2t_B}\right) = 1 - \frac{t_A}{2t_B}$$

Ta có điều sau

$$
\begin{aligned}
P\left(Y\in B|X\in A\right) &= \frac{P\left(Y\in B, X\in A\right)}{P\left(X\in A\right)}\\
&= \frac{P\left(Y\in B, X\in A\right)}{P\left(X\in A\right)}\\
&= \frac{1}{P\left(X\in A\right)}\int_{-\infty}^{+\infty} P\left(Y\in B, X\in A|X=x\right) f_X\left(x\right)dx\\
&= \frac{1}{P\left(X\in A\right)}\int\limits_A P\left(Y\in B, X\right) f_X\left(x\right)dx\\
&= \frac{1}{P\left(X\in A\right)}\int\limits_A \left(\int\limits_B f_{Y|X}(y|x)dy\right) f_X\left(x\right)dx\\
&\frac{1}{P\left(X\in A\right)}\int\limits_A\int\limits_B f_{X,Y}\left(x,y\right)dydx
\end{aligned}
$$

5.10. Kỳ vọng có điều kiện.

Định nghĩa 5.10.1. Cho hai biến ngẫu nhiên X, Y, kỳ vọng có điều kiện của Y khi biết X nhận giá trị x, ký hiệu $E(Y|X=x)$ được định nghĩa là:

$$
E\left(Y|X=x\right) = \begin{cases} \sum y_i f_{Y|X}\left(y_i|x\right) & X, Y \text{ rời rạc}\\ \int_{-\infty}^{\infty} y f_{Y|X}\left(y|x\right)dy & X, Y \text{ liên tục}\end{cases}
$$

Chú ý: Kỳ vọng có điều kiện của Y đối với X ký hiệu $E\left(Y|X\right)$ là một biến ngẫu nhiên có giá trị là $E\left(Y|X=x\right)$ trên tập $\{X=x\}$

$$
E\left(g\left(X,Y\right)|X=x\right) = \begin{cases} \sum g\left(x, y_i\right) f_{Y|X}\left(y_i|x\right) & X, Y \text{ rời rạc}\\ \int_{-\infty}^{\infty} g\left(x,y\right) f_{Y|X}\left(y|x\right)dy & X, Y \text{ liên tục}\end{cases}
$$

Ví dụ: Cho $X \sim \mathcal{U}\left(0,1\right), Y|X \sim \mathcal{U}\left(x,1\right)$. Tính $E\left(Y|X=x\right), E\left(Y|X\right)$

Giải

$$
\begin{aligned}
f_{Y|X}\left(y|x\right) &= \frac{1}{1-x}\left(0<x<y<1\right)\\
E\left(Y|X=x\right) &= \int_x^1 y\frac{1}{1-x}dy = \left.\frac{y^2}{2\left(1-x\right)}\right|_x^1\\
&= \frac{1-x^2}{2\left(1-x\right)} = \frac{1+x}{2}
\end{aligned}
$$

Vậy $E\left(Y|X\right) = \dfrac{1+X}{2}$.

Định lý 5.10.2. Cho X, Y là các biến ngẫu nhiên sao cho $E\left(X\right), E\left(Y\right)$ tồn tại, khi đó:

– $E\left(X\right) = E\left(E\left(X|Y\right)\right)$
  $E\left(Y\right) = E\left(E\left(Y|X\right)\right)$

– $E\left(r\left(X,Y\right)\right) = E\left(E\left(g\left(X,Y\right)|X\right)\right) = E\left(E\left(g\left(X,Y\right)|Y\right)\right)$

Định nghĩa 5.10.3. Nếu $E\left(Y|X\right)$ tồn tại thì phương sai có điều kiện của biến ngẫu nhiên Y đối với X được định nghĩa là:

$$Var\left(Y|X\right) = E\left[\left(Y - E\left(Y|X\right)\right)^2|X\right]$$

Định lý 5.10.4

$$Var\left(X\right) = E\left(Var\left(X|Y\right)\right) + Var\left(E\left(X|Y\right)\right)$$

5.11. Hàm sinh Moment

Định nghĩa 5.11.1. Hàm sinh moment của biến ngẫu nhiên X được định nghĩa như sau:

$$\begin{aligned} M_X : \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto E\left(e^{tx}\right) \end{aligned}$$

$$M_X\left(t\right) = E\left(e^{tx}\right)\begin{cases}\sum_{x_i} e^{tx_i}P\left(X = x_i\right)\\ \int_{-\infty}^{\infty} e^{tx}f\left(x\right)dx\end{cases}$$

Chú ý: Hàm sinh moment của một biến ngẫu nhiên có thể không tồn tại.
Từ khai triển:

$$e^{tX} = 1 + tX + \frac{\left(tX\right)^2}{2} + \dots + \frac{\left(tX\right)^2}{n}$$

$$\begin{aligned}\Rightarrow E\left(e^{tX}\right) &= 1 + tE\left(X\right) + \frac{t^2}{2!}E\left(X^2\right) + \dots + \frac{t^n}{n!}E\left(X^n\right)\\ &= 1 + m_1t + m_2\frac{t^2}{2!} + \dots + m_n\frac{t^n}{n!}\\ &\Rightarrow \begin{cases}M_X^1\left(0\right) = m_1 = E\left(X\right)\\ M_X^n\left(0\right) = m_1 = E\left(X^n\right)\end{cases}\end{aligned}$$

– $X \sim B(1,p)$:

$$M_X\left(t\right) = \left(Ee^{tX}\right) = e^{0t}P\left(X = 0\right) + e^{1t}P\left(X = 1\right) = \left(1 - p\right) + e^tp$$

– $X \sim B(n,p)$:

$$\begin{aligned}M_X\left(t\right) = E\left(e^{tX}\right) &= \sum_{k=0}^{n} e^{tk}P\left(X = k\right) = \sum_{k=0}^{n} e^{tk}C_n^kp^k\left(1 - p\right)^{n-k}\\ &= \sum_{k=0}^{n} C_n^k\left(e^tp\right)^k\left(1 - p\right)^{n-k} = \left(e^tp + 1 - p\right)^n\end{aligned}$$

– $X \sim \mathcal{P}(\lambda)$. Tính $M_X(t)$, từ đó tính $E(X), E\left(X^2\right), E\left(X^3\right)$

$$\begin{aligned}
M_X(t) &= \sum_{k=0}^{+\infty} e^{tk} P(X=k) \\
&= \sum_{k=0}^{+\infty} e^{tk} \frac{\lambda^k e^{-\lambda}}{k!} = \sum_{k=0}^{+\infty} \frac{\left(e^t\lambda\right)^k e^{-\lambda}}{k!} \\
&= e^{-\lambda} \sum_{k=0}^{+\infty} \frac{\left(e^t\lambda\right)^k}{k!} = e^{-\lambda} e^{e^t\lambda} = e^{\lambda\left(e^t-1\right)} \\
M'_X(t) &= \lambda e^t e^{\lambda\left(e^t-1\right)} \Rightarrow M'_X(0) = \lambda \\
M''_X(t) &= \lambda e^t e^{\lambda\left(e^t-1\right)} + \lambda^2 e^{2t} e^{\lambda\left(e^t-1\right)} \Rightarrow M''_X(0) = E\left(X^2\right) = \lambda^2 + \lambda
\end{aligned}$$

Ví dụ: Cho $X \sim \mathcal{U}(a,b)$

1. Tính M_X
2. Từ công thức của M_X tìm $E(X^n), n \in \mathbb{N}$

Giải

$$\begin{aligned}
M_X(t) &= \int_{-\infty}^{+\infty} e^{tx} f(x)\,dx = \int_a^b e^{tx} \frac{1}{b-a} dx \\
&= \frac{1}{b-a} \int_a^b e^{tx} dx = \frac{1}{b-a} \left(\left. \frac{e^{tx}}{t} \right|_a^b \right) = \frac{1}{b-a} \cdot \frac{\left(e^{tb} - e^{ta}\right)}{t} \\
&= \frac{1}{t(b-a)} \left(\sum_{n=0}^{+\infty} \frac{1}{n!} (tb)^n - \sum_{n=0}^{+\infty} \frac{1}{n!} (ta)^n \right) \\
&= \frac{1}{t(b-a)} \left(\sum_{n=0}^{+\infty} \frac{1}{n!} t^n b^n - \sum_{n=0}^{+\infty} \frac{1}{n!} t^n a^n \right) \\
&= \frac{1}{t(b-a)} \left[1 + \frac{tb}{1!} + \frac{t^2b^2}{2!} + \cdots + \frac{t^nb^n}{n!} + \cdots \right. \\
&\qquad \left. - \left(1 + \frac{ta}{1!} + \frac{t^2a^2}{2!} + \cdots + \frac{t^na^n}{n!} + \cdots \right) \right] \\
&= \frac{1}{t(b-a)} \left[\frac{t(b-a)}{1!} + \frac{t^2\left(b^2-a^2\right)}{2!} + \cdots + \frac{t^n\left(b^n-a^n\right)}{n!} + \cdots \right] \\
&= 1 + \frac{t\left(b^2-a^2\right)}{2!(b-a)} + \cdots + \frac{t^{n-1}\left(b^n-a^n\right)}{(n-1)!n(b-a)} + \frac{t^n\left(b^{n+1}-a^{n+1}\right)}{(n)!(n+1)(b-a)} + \cdots
\end{aligned}$$

$$\Rightarrow E(X^n) = \frac{b^{n+1}-a^{n+1}}{(n+1)(b-a)}$$

Kiểm tra lại:

$$E(X) = \frac{a+b}{2}$$

$$E\left(X^2\right) = Var(X) + [E(X)]^2 = \frac{(b-a)^2}{12} + \left(\frac{b+a}{2}\right)^2$$

$$= \frac{(b-a)^2 + 3(b+a)^2}{12} = \frac{b^3-a^3}{3(b-a)}$$

$$M'_X(0) = \lim_{t\to 0} \frac{M_X(t) - M_X(0)}{t} = \lim_{t\to 0} \frac{\frac{1}{b-a}\cdot\frac{1}{t}\cdot\left(e^{tb}-e^{ta}\right) - 1}{t}$$

$$= \lim_{t\to 0} \frac{\left(e^{tb}-e^{ta}\right) - (tb-ta)}{(b-a)\,t^2} = \lim_{t\to 0} \frac{\left(e^{tb}-e^{ta}\right) - t(b-a)}{(b-a)\,t^2}$$

$$= \lim_{t\to 0} \frac{\left(1+tb+\frac{t^2b^2}{2!} + 0\left(t^2\right)\right) - \left(1+ta+\frac{t}{2} - t\right)}{t^2(b-a)}$$

$$= \lim_{t\to 0} \frac{t(b-a) + \frac{t^2}{2!}\left(b^2-a^2\right) - t(b-a)}{t^2(b-a)} = \frac{1}{2}(b+a) = E(X)$$

$$f(x) = 0\left(t^2\right) \Leftrightarrow \lim_{t\to 0} \frac{f(x)}{t^2} = 0$$

Ví dụ: Cho $X \sim \mathcal{N}(0,1)$. Tính $M_X(t), E(X^n), n \in \mathbb{N}$

Giải

$$M_X(t) = E\left(e^{tX}\right) = \int_{-\infty}^{+\infty} e^{tx}\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{\frac{-x^2}{2}+tx}dx$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(x^2-2tx+t^2\right)+\frac{t^2}{2}}dx$$

$$= \frac{1}{\sqrt{2\pi}}e^{\frac{t^2}{2}}\int_{-\infty}^{+\infty} e^{-1(x-t)^2}dx$$

$$= e^{\frac{t^2}{2}}\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-t)^2}dx = e^{\frac{t^2}{2}}$$

$$M_X(t) = e^{\frac{t^2}{2}} = 1 + \frac{t^2}{2} + \frac{\left(\frac{t^2}{2}\right)^2}{2!} + \cdots + \frac{\left(\frac{t^2}{2}\right)^n}{n!} + \cdots$$

$$E(X^n) = 1 + \frac{t^2}{2} + \frac{t^4}{2^2.2!} + \cdots + \frac{t^{2n}}{2^n.n!} + \cdots$$

$$= 1 + \frac{t^2}{2}\cdot 1 + \frac{t^4}{4!}\cdot\frac{4!}{2^2.2!} + \cdots + \frac{t^{2n}}{(2n)!}\cdot\frac{(2n)!}{2^n.n!} + \cdots$$

$$\begin{cases} 0 & n \text{ lẻ} \\ \frac{(2k)!}{2^k k!} & n = 2k \end{cases}$$

Bổ đề 5.11.2.

1. Nếu $Y = aX + b$ thì $M_Y(t) = e^b M_X(at)$

2. Nếu $X_1, \dots, X_n$ là các biến ngẫu nhiên độc lập và $S = X_1 + \cdots + X_n$ thì

$$M_S(t) = M_{X_1}(t) \dots M_{X_n}(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

$$M_Y(t) = E\left(e^{ty}\right) = E\left(e^{(ax+b)t}\right) = e^{bt}E\left(e^{atx}\right) = e^{bt}M_X(at)$$

$$\begin{aligned} M_S(t) &= E\left(e^{t(X_1+\cdots+X_n)}\right) = E\left(e^{tX_1}, \dots, e^{tX_n}\right) \\ &= E\left(e^{tX_1}\right) \dots E\left(e^{tX_n}\right) = M_{X_1}(t) \dots M_{X_n}(t) \end{aligned}$$

Định nghĩa 5.11.3. Hàm sinh moment của vector ngẫu nhiên $X = (X_1, \dots, X_n)$ được định nghĩa là:

$$\begin{aligned} M_X : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ t = (t_1, \dots, t_n) &\longmapsto E\left(e^{t_1X_1+\cdots+t_nX_n}\right) \end{aligned}$$

Định lý 5.11.4. Cho hai biến ngẫu nhiên X, Y với hàm sinh moment tương ứng là M_X, M_Y. Nếu $M_X(t) = M_Y(t)$ với mọi t trong một lân cận mở chứa 0 thì X, Y cùng phân phối.

Ví dụ: Cho $X_i \sim \mathcal{P}(\lambda_i), i = \overline{1,n}$ và X_i là các biến ngẫu nhiên độc lập. Đặt $Y = X_1 + \cdots + X_n$. Hãy tìm phân phối của Y.

Giải

$$\begin{aligned} M_Y(t) &= M_{X_1}(t) \dots M_{X_n}(t) = e^{\lambda_1(e^t-1)} \dots e^{\lambda_n(e^t-1)} \\ &= e^{(\lambda_1+\cdots+\lambda_n)(e^t-1)} = e^{\lambda'(e^t-1)} \quad \forall t \end{aligned}$$

Ví dụ: Cho X, N là hai biến ngẫu nhiên thỏa $X|N = n \sim B(n,p), N \sim \mathcal{P}(\lambda)$. Tìm phân phối của X.

Giải

Ta có:

$$M_X(t) = E\left(e^{tx}\right) = E\left(e^{tx}|N\right) = E\left(e^{tx}|N = n\right)$$

Vì $X|N = n \sim B(n,p)$:

$$\begin{aligned} &\Rightarrow E\left(e^{tx}|N = n\right) = \left(pe^t + 1 - n\right)^n \\ &\quad \Rightarrow E\left(e^{tx}|N\right) = \left(pe^t + q\right)^n = g(N) \end{aligned}$$

Do đó:

$$\Rightarrow E\left[E\left(e^{tx}|N\right)\right] = \sum_{k=0}^{+\infty} \left(pe^t + q\right)^k \cdot \frac{e^{-\lambda}\lambda}{k!} = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{\left[\left(pe^t + q\right)\lambda\right]^k}{k!}$$
$$= e^{-\lambda} e^{\left(pe^t + q - 1\right)} = e^{\lambda\left(pe^t - p\right)} = e^{\lambda p\left(e^t - 1\right)}$$

Phần VI

Một số bất đẳng thức

6.1 Bất đẳng thức Markov:

Định nghĩa 6.1.1. Giả sử X là biến ngẫu nhiên không âm và $E(X)$ hữu hạn, khi đó:

$$P(X \geq t) \leq \frac{E(X)}{t} \quad \forall t \in \mathbb{R}^*$$

Chứng minh

Trường hợp liên tục tuyệt đối:

$$\begin{aligned} E(X) &= \int_0^{+\infty} xf(x)\,dx = \int_0^t xf(x)\,dx + \int_t^{+\infty} xf(x)\,dx \\ &\geq \int_t^{+\infty} xf(x)\,dx \geq t\int_t^{+\infty} f(x)\,dx = tP(X \geq t) \\ &\Rightarrow P(X \geq t) \leq \frac{E(X)}{t} \end{aligned}$$

6.2 Bất đẳng thức Chebyshev.

Định nghĩa 6.2.1. Cho X là biến ngẫu nhiên có kỳ vọng μ và phương sai σ^2, khi đó:

$$P((|X-\mu|) \geq t) \leq \frac{\sigma^2}{t^2}$$

$$\text{hay } P(|X - E(X)| \geq t) \leq \frac{Var(X)}{t^2} \quad \forall t \in \mathbb{R}_+^*$$

Chứng minh

$$\begin{aligned} Var(X) &= \int_{-\infty}^{+\infty} (x - E(X))^2 f(x)dx \\ &= \int\limits_{|x-E(X)| \leq t} (x - E(X))^2 f(x)\,dx + \int\limits_{|x-E(X)|>t} (x - E(X))^2 f(x)dx \\ &\geq \int\limits_{|x-E(X)|>t} (x - E(X))^2 f(x)dx \\ &\geq \int\limits_{|x-E(X)|>t} t^2 f(x)\,dx \\ &= t^2 P(|X - E(X)| \geq t) \end{aligned}$$

6.3. Bất đẳng thức Hoeffding.

Định nghĩa 6.3.1. Cho $Y_1, \ldots, Y_n$ là các biến ngẫu nhiên độc lập thỏa $E(Y_i) = 0, a_i \leq Y_i \leq b_i \, \forall i$ thì:

$$P\left(\sum Y_i \geq i\right) \leq \varepsilon^{t^2} \prod \exp\left\{\frac{b^2 (b_i - a_i)^2}{8}\right\} \quad \forall \varepsilon > 0, \forall t > 0$$

Ví dụ: Cho $X_1, \ldots, X_{100}$ là các biến ngẫu nhiên độc lập với $X_i \sim B(1,p), p = 0.01$. Đặt $\overline{X} = \frac{1}{100}\sum_{i=1}^{100} X_i$. Hãy ước lượng $P(X \geq 2)$ bằng bất đẳng thức Markov, Chebyshev.

Giải

$$X = X_1 + \cdots + X_{100} \Rightarrow X \sim B(100, p) \Rightarrow \begin{cases} E(X) & = 100p \\ Var(X) & = 100p(1-p) \end{cases}$$

Từ đó ta có:

$$\begin{aligned} P(X \geq 2) &\leq \frac{E(X)}{2} = \frac{100p}{2} = \frac{1}{2} \quad \text{(BĐT Markov)} \\ P(X \geq 2) &= P(X - 100p \geq 2 - 100p) \\ &\leq P(|X - 100p| \geq 2 - 100p) \\ &\leq \frac{100p(1-p)}{(2-100p)^2} = 0.99 \quad \text{(BĐT Chebyshev)} \end{aligned}$$

Hệ quả 6.3.2. Cho $X_1, \ldots, X_n$ là các biến ngẫu nhiên độc lập với $X_i \sim B(1,p)$ thì $P\left(\left|\overline{X} - p\right| \geq \varepsilon\right) = 2e^{-2n\varepsilon^2}$

Ví dụ: Cho $X_1, \ldots, X_{100}$ là các biến ngẫu nhiên độc lập với $X_i \sim B(1,0)$. Hãy so sánh $P\left(\left|\overline{X} - 0.2\right| \geq 1\right)$ bằng hai bất đẳng thức Chebyshev và hệ quả của bất đẳng thức Hoeffding.

Giải

$$\begin{aligned} E(X) &= \frac{1}{100} \cdot 100 \cdot 0.2 = 0.2 \\ Var(X) &= \frac{1}{100^2} \cdot 100 \cdot 0.8 \cdot 0.2 = 0.0016 \\ P\left(\left|\overline{X} - 0.2\right| \geq 1\right) &\leq \frac{0.0016}{1^2} = 0.0016 \end{aligned}$$

6.4. Bất đẳng thức Mill.

Định nghĩa 6.4.1. Cho $Z \sim N(0,1)$ và $t > 0$, khi đó:

$$P(|Z| \geq t) \leq \sqrt{\frac{2}{\pi}} \cdot \frac{e^{\frac{-t^2}{2}}}{t}$$

Ví dụ: Cho $X_1, \ldots, X_{100}$ độc lập, cùng phân phối $X_i \sim N(0,1)$. Hãy chặn $P\left(\left|\overline{X}\right| \geq t\right)$ bằng bất đẳng thức Chebyshev, Mill. Trong đó $\overline{X} = \frac{1}{100}\sum_{i=1}^{100} X_i$

Giải

Ta có:

$$E\left(\overline{X}\right) = \frac{1}{100}\sum_{i=1}^{100} E\left(X_i\right) = 0$$

$$\overline{X} = \frac{X_1 + \cdots + X_{100}}{100} \sim N\left(0, \frac{1}{100}\right)$$

$$\sigma^2 = Var\left(\frac{X_1 + \cdots + X_{100}}{100}\right) = \frac{1}{100^2}\sum_{i=1}^{100} Var\left(X_i\right)$$

$$= \frac{1}{100^2} \cdot 100 \cdot 1 = 0.01$$

Như vậy:

– Theo BĐT Mill:

$$P\left(\left|\overline{X}\right| \geq t\right) = P\left(\frac{|X|}{100} \geq \frac{t}{100}\right)$$

$$\leq \sqrt{\frac{2}{\pi}} \cdot \frac{\exp\left\{-\left(\frac{t}{100}\right)^2 \cdot \frac{1}{2}\right\}}{\frac{t}{100}}$$

– Theo BĐT Chebyshev:

$$P\left(|X| \geq t\right) \leq \frac{\sigma^2}{t^2} = \frac{0.01}{t^2}$$

$$\overline{X} \sim N\left(0, \frac{1}{100}\right)$$

$$\Rightarrow P\left(\left|\overline{X}\right| \geq t\right) = P\left(\frac{\left|\overline{X}\right|}{1/10} \geq \frac{t}{1/10}\right)$$

$$= P\left(\frac{\overline{|X|}}{1/10} \geq 10t\right)$$

$$\leq \sqrt{\frac{2}{\pi}} \cdot \frac{\exp\left\{-\frac{(10t)^2}{2}\right\}}{10t}$$

Ví dụ: Cho $X \sim P(\lambda)$. Dùng bất đẳng thức Chebyshev chứng minh $P(X \geq 2\lambda) \leq \dfrac{1}{\lambda}$

Giải

Ta có:

$$E(X) = \lambda, Var(X) = \lambda$$

Vậy:

$$P(X \geq 2\lambda) = P(X - \lambda \geq 2\lambda - \lambda)$$

$$\leq P\left(|X - \lambda| \geq \lambda\right)$$

$$\leq \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

6.5. Bất đẳng thức liên quan tới kỳ vọng.

Định lý 6.5.1. Cho X, Y là hai biến ngẫu nhiên có phương sai hữu hạn, khi đó:

$$E(XY) \leq \sqrt{E(X^2)\,E(Y^2)}$$

Định lý 6.5.2. Cho X là biến ngẫu nhiên với kỳ vọng $E(X)$. Khi đó:

$$\begin{cases} E(g(X)) \geq g(E(X)) & \text{nếu } g \text{ là một hàm lồi} \\ E(g(X)) \leq g(E(X)) & \text{nếu } g \text{ là một hàm lõm} \end{cases}$$

Chú ý:

– g là một hàm lồi trên $[a, b]$

$$\Leftrightarrow \forall x, y \in [a, b], \forall 0 \leq \alpha \leq 1 : g[\alpha x + (1-\alpha) y] \leq \alpha g(x) + (1-\alpha) g(y)$$

– g là một hàm lõm khi và chỉ khi $-g$ là một hàm lồi

Phần VII

Định lý giới hạn trung tâm và luật số lớn.

7.1. Sự hội tụ của các biến ngẫu nhiên.

Trong phần này $X_1, \dots, X_n$ là các biến ngẫu nhiên được định nghĩa trên cùng một không gian xác suất $\mathcal{F}_1, \dots, \mathcal{F}_n$ là hàm phân phối của $X_1, \dots, X_n$, $\mathcal{F}$ là hàm phân phối của X.

Định nghĩa 7.1.1. Hội tụ hầu khắp nơi
Dãy biến ngẫu nhiên $\{X_n\}$ hội tụ hầu khắp nơi (hầu chắc chắn) về biến ngẫu nhiên X nếu:

$$P\left(\omega \in \Omega : X_n\left(\omega\right) \longrightarrow X\left(\omega\right)\right) = 1$$

Ký hiệu:

$$X_n \xrightarrow{\text{a.s.}} X$$

Định nghĩa 7.1.2. Hội tụ theo xác suất
Dãy biến ngẫu nhiên $X_1, \dots, X_n$ hội tụ theo xác suất về biến ngẫu nhiên X nếu:

$$\forall \varepsilon > 0, P\left(\left|X_n - X\right| > \varepsilon\right) \xrightarrow{n \to \infty} 0$$

Ký hiệu:

$$X_n \xrightarrow{p} X$$

Định nghĩa 7.1.3. Hội tụ theo phân phối
Dãy biến ngẫu nhiên $\{X_n\}$ hội tụ theo phân phối về biến ngẫu nhiên X nếu:

$$F_n\left(x\right) \longrightarrow F\left(x\right), \forall x \in C\left(F_X\right), C(F_X) = \left\{x : F \text{ liên tục tại } x\right\}$$

Ký hiệu:

$$X_n \xrightarrow{d} X$$

Chú ý: Trong trường hợp này $\{X_n\}, X$ có thể được định nghĩa trên các không gian xác suất khác nhau.

Định nghĩa 7.1.4. Hội tụ theo r trung bình
Dãy biến ngẫu nhiên $X_1, \dots, X_n$ hội tụ theo r trung bình về biến ngẫu nhiên X nếu:

$$E\left(\left|X_n - X\right|^r\right) \xrightarrow{n \to \infty} 0$$

Ký hiệu:

$$X_n \xrightarrow{r} X$$

Chú ý: Khi $r = 2$, ta còn gọi là hội tụ theo bình phương trung bình.
Ký hiệu:

$$X_n \xrightarrow{\text{q.m.}} X$$

Ví dụ: Cho $\{X_n\}_{(n\geq 1)}$ thỏa:

$$P\left(X_n = \frac{1}{n}\right) = 1 - \frac{1}{n^2}, P(X_n = n) = \frac{1}{n^2}$$

a. Chứng minh $X_n \xrightarrow{p} X$, trong đó $P(X = 0) = 1$
b. Chứng minh $X_n \xrightarrow{\text{q.m.}} X$

Giải

a. Ta có:

$$\begin{aligned}
&\forall \varepsilon > 0, P(|X_n - X| > \varepsilon) \longrightarrow 0\\
&P(|X_n - X| > \varepsilon) = P(\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon)\\
&\qquad\qquad\qquad\qquad = P(\omega \in \Omega : |X_n(\omega)| > \varepsilon) \longrightarrow 0
\end{aligned}$$

Mặt khác $\exists n_0(\varepsilon) > \varepsilon, \forall n > n_0(\varepsilon)$ sao cho $\dfrac{1}{n(\Omega)} < \varepsilon$.
Ta chọn:

$$n_1 : n_1 > \varepsilon, n_2 : \frac{1}{n_2} < \varepsilon$$
$$n_0 = \max\{n_1, n_2\}$$

Khi đó:

$$\forall n > n_0 \Rightarrow \begin{cases} n > \varepsilon \\ \dfrac{1}{n} < \dfrac{1}{n^2} < \varepsilon \end{cases}$$

Vì vậy:

$$\begin{aligned}
P(\omega \in \Omega : |X_n(\omega)| > \varepsilon) &= P(\omega \in \Omega : X_n(\omega) = n)\\
&= \frac{1}{n^2}
\end{aligned}$$

Vậy:

$$\begin{aligned}
&\forall n > n_0, P(|X_n - X| > \varepsilon) = \frac{1}{n^2} \longrightarrow 0\\
&\Rightarrow \forall \varepsilon > 0, P(|X_n - X| > \varepsilon) \longrightarrow 0\\
&\qquad\qquad \Rightarrow X_n \longrightarrow X
\end{aligned}$$

b. Ta cần chứng minh:

$$E\left(|X_n - X|^{2}\right) \longrightarrow 0$$

Mà

$$\begin{aligned}
E\left(|X_n - X|^2\right) &= E\left(|X_0|^2\right) \\
&= \frac{1}{n^2}\left(1 - \frac{1}{n^2}\right) + n^2 \cdot \frac{1}{n} \\
&= 2 + \frac{1}{n}\left(1 - \frac{1}{n^2}\right) \longrightarrow 0
\end{aligned}$$

Ví dụ: Cho $X_n \sim N\left(0, \frac{1}{n}\right)$. Chứng minh $X_n \xrightarrow{d} X$. Trong đó:

$$F_X(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Giải

Ta có:

$$\begin{aligned}
&F_n(x) \longrightarrow F_X(x)\, \forall x \in C(F_X) = \mathbb{R} \backslash \{0\} \\
&F_n(x) = P(X_0 \leq x) = P\left(X_0\sqrt{n} \leq x\sqrt{n}\right) = \Phi\left(x\sqrt{n}\right)
\end{aligned}$$

Nếu $x > 0$ thì:

$$x\sqrt{n} \xrightarrow{n\to\infty} +\infty \Rightarrow F_n(x) = \Phi\left(x\sqrt{n}\right) \longrightarrow 1 = F_X(x)$$

Nếu $x < 0$ thì:

$$x\sqrt{n} \xrightarrow{n\to\infty} -\infty \Rightarrow F_n(x) = \Phi\left(x\sqrt{n}\right) \longrightarrow 0 = F_X(x)$$

Ví dụ: Cho $X_1, \ldots, X_n, \ldots$ là dãy biến ngẫu nhiên độc lập cùng phân phối có hàm mật độ:

$$f(x) = \begin{cases} \alpha x^{-\alpha-1} & x > 1 \\ 0 & \text{ngược lại} \end{cases}$$

Đặt $Y_n = \max\{X_1, \ldots, X_n\}$, Y_n có hội tụ theo phân phối không? Nếu có hãy xác định phân phối giới hạn.

Giải

Ta có

$$\begin{aligned}
F_{Y_n}(x) &= P\left(\max\{X_1, \ldots, X_n\} \leq X\right) = P(X_1 \leq x, \ldots, X_n \leq x) \\
&= \left[P(X_1 \leq x)\right]^n = \left[F_{X_1}(x)\right]^n
\end{aligned}$$

Với:

$$F_{X_1}(x) = \begin{cases} 1 - x^{-\alpha} & x \geq 1 \\ 0 & x < 1 \end{cases}$$

$$\Rightarrow F_{Y_n}(x) = [F_{X_1}(x)]^n = \begin{cases} (1 - x^{-\alpha}) & x \geq 1 \\ 0 & \text{ngược lại} \end{cases}$$

Ta có:

- $\alpha > 0$:

$$\begin{aligned} &\Rightarrow x^{-\alpha} = \frac{1}{x^{\alpha}} < 1, \forall x > 1 \\ &\Rightarrow \left(1 - x^{-\alpha}\right)^n \longrightarrow 0, \forall x > 1 \\ &\Rightarrow F_{Y_n}(x) \longrightarrow 0, \forall x > 1 \end{aligned}$$

- $x = 1 \Rightarrow F_{Y_n}(x) = 0, \forall n$

- $x < 0 \Rightarrow F_{Y_n}(x) = 0, \forall n$

Vậy $F_{Y_n}(x) \longrightarrow 0, \forall x$

Định lý 7.1.5. Nếu dãy biến ngẫu nhiên $X_1, X_2, \ldots$ hội tụ hầu khắp nơi (hội tụ theo xác suất, hội tụ theo $r-$trung bình, hội tụ theo phân phối) thì biến ngẫu nhiên giới hạn là duy nhất (theo nghĩa hầu khắp)

Định lý 7.1.6. Cho $\{X_n\}, \{Y_n\}$ là các dãy biến ngẫu nhiên:

1. $X_n \xrightarrow{\text{a.s.}} X, Y_n \xrightarrow{\text{a.s.}} Y \Rightarrow X_n + Y_n \xrightarrow{\text{a.s.}} X + Y.$
2. $X_n \xrightarrow{p} X, Y_n \xrightarrow{p} Y \Rightarrow X_n + Y_n \xrightarrow{p} X + Y.$
3. $X_n \xrightarrow{r} X, Y_n \xrightarrow{r} Y \Rightarrow X_n + Y_n \xrightarrow{r} X + Y.$
4. $X_n \xrightarrow{d} X, Y_n \xrightarrow{d} c \Rightarrow X_n + Y_n \xrightarrow{d} X + c.$
5. $X_n \xrightarrow{\text{a.s.}} X, Y_n \xrightarrow{\text{a.s.}} Y \Rightarrow X_n \times Y_n \xrightarrow{\text{a.s.}} X \times Y.$
6. $X_n \xrightarrow{p} X, Y_n \xrightarrow{p} Y \Rightarrow X_n \times Y_n \xrightarrow{p} X \times Y.$
7. $X_n \xrightarrow{d} X, Y_n \xrightarrow{d} c \Rightarrow X_n \times Y_n \xrightarrow{d} cX.$
8. Cho g là một hàm liên tục, nếu:

$$\begin{aligned} X_n \xrightarrow{\text{a.s.}} X &\Rightarrow g(X_n) \xrightarrow{\text{a.s.}} g(X) \\ X_n \xrightarrow{p} X &\Rightarrow g(X_n) \xrightarrow{p} g(X) \\ X_n \xrightarrow{d} X &\Rightarrow g(X_n) \xrightarrow{d} g(X) \end{aligned}$$

7.2. Luật số lớn

Trong phần này, $X_1, \ldots, X_n$ là một dãy biến ngẫu nhiên độc lập, cùng phân phối có $E\left(X_i\right) = \mu, Var\left(X_i\right) = \sigma^2$. Ký hiệu: $\overline{X_n} = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i$. Ta có:

$$E\left(\overline{X_n}\right) = \mu, Var\left(\overline{X_n}\right) = \frac{\sigma^2}{n}$$

Định nghĩa 7.2.1. Luật yếu số lớn (The weak law of large number)
Cho $X_1, X_2, \ldots$ là dãy biến ngẫu nhiên độc lập cùng phân phối, khi đó:

$$X_n \quad \xrightarrow{p} \quad \mu$$

Định nghĩa 7.2.2. Luật mạnh số lớn (The strong law of large number)
Cho $X_1, X_2, \ldots$ là dãy biến ngẫu nhiên độc lập cùng phân phối có $E\left(X_i\right) < \infty$ thì:

$$\overline{X_n} \quad \xrightarrow{\text{a.s.}} \quad \mu$$