

MỤC LỤC

Nội dung	Trang
Chương I: Sơ lược về xác suất và biến ngẫu nhiên	2
I: Định nghĩa, công thức tính xác suất	2
II: Biến ngẫu nhiên, quy luật phân phối xác suất	10
Chương II: Dữ liệu thống kê và các đại lượng thống kê mô tả	23
I: Thu thập dữ liệu và lưu trữ dữ liệu	23
II: Các đại lượng thống kê mô tả	27
Chương III: Ước lượng tham số tổng	31
I. Ước lượng điểm	31
II. Khoảng ước lượng điểm	32
Chương IV: Kiểm định giả thiết thống kê và phân tích phương sai	41
I: Kiểm định giả thiết tham số	41
II: Kiểm định giả thiết phi tham số	71
Chương V: Phân tích hồi quy và tương quan	82
I: Hệ số tương quan và phương trình hồi quy	82
II: Phân Kiểm định hệ số tương quan, sự phù hợp của phương trình hồi quy	84
Tài liệu tham khảo	95
Phụ lục	96

CHƯƠNG I

SƠ LƯỢC XÁC SUẤT, BIẾN NGẪU NHIÊN

Mục tiêu học tập: Sau khi học xong bài này, người học có thể:

- * Hiểu khái niệm xác suất
- * Nắm vững các công thức tính xác suất.
- * Giải được các bài toán cơ bản về xác suất

I. ĐỊNH NGHĨA, CÔNG THỨC TÍNH XÁC SUẤT

1. Biến cố ngẫu nhiên và các phép toán trên biến cố ngẫu nhiên

1.1 Đặt vấn đề

Trong thực tế cho thấy có rất nhiều thí nghiệm khi tiến hành nhiều lần trong cùng điều kiện ban đầu nhưng không dẫn đến cùng kết quả. Chẳng hạn khi tung một con xúc xắc xem như thực hiện một thí nghiệm, khi đó ta không thể đoán trước được chắc chắn kết quả xuất hiện là mặt mấy chấm.

Những hiện tượng khi biết trước các điều kiện ban đầu mà ta không thể xác định chắc chắn kết quả xảy ra của nó gọi là hiện tượng ngẫu nhiên hay phép thử ngẫu nhiên.

Ví dụ: lượng mưa trong năm; đầu tư vào một dự án; tham gia một kỳ thi tuyển sinh; kinh doanh một mặt hàng nào đó;... là các hiện tượng ngẫu nhiên.

1.2 Biến cố ngẫu nhiên, Không gian biến cố sơ cấp

a. Biến cố sơ cấp

Khi thực hiện một phép thử ngẫu nhiên, mỗi kết quả có thể xảy ra của nó được gọi là biến cố sơ cấp.

Tập hợp tất cả các biến cố sơ cấp của phép thử gọi là không gian các biến cố sơ cấp. Kí hiệu : Ω

Ví dụ:

Khi gieo một con xúc xắc. Gọi e_i là kết quả xuất hiện mặt i chấm ($i=1;2;3;4;5;6$).

Khi đó: + Phép thử này có 6 biến cố sơ cấp : $e_1; e_2; e_3; e_4; e_5; e_6$.

+ Không gian các biến cố sơ cấp $\Omega = \{e_1; e_2; e_3; e_4; e_5; e_6\}$

Ví dụ:

Khi gieo một hạt giống. Gọi N là kết quả nảy mầm; K là kết quả không nảy mầm

Khi đó: + Phép thử này có 2 biến cố sơ cấp : $N; K$.

+ Không gian các biến cố sơ cấp $\Omega = \{N; K\}$

b. Biến cố ngẫu nhiên(gọi tắt là biến ngẫu nhiên)

Khi thực hiện phép thử ngẫu nhiên, mỗi kết cục có thể xảy ra hoặc không thể xảy ra trong kết quả của phép thử gọi là biến cố ngẫu nhiên. Biến cố ngẫu nhiên thường kí hiệu: A, B, C, D, ...

Ví dụ:

Khi gieo một con xúc xắc. Gọi A là kết cục mặt chẵn xuất hiện; B là kết cục mặt lẻ xuất hiện; C là kết cục mặt chia hết cho 3 xuất hiện; ...

Khi đó: + A, B, C, ... là các biến cố ngẫu nhiên

* Biến cố ngẫu nhiên A là tập hợp gồm một số biến cố sơ cấp. Do đó biến cố ngẫu nhiên A là tập hợp con của Ω .

Ví dụ :

* Chọn các mệnh đề đúng trong các mệnh đề sau

- a) Biến cố ngẫu nhiên là kết cục luôn xảy ra trong phép thử ngẫu nhiên.
- b) Phép thử ngẫu nhiên là biến cố ngẫu nhiên.
- c) Biến cố sơ cấp là biến cố ngẫu nhiên
- d) Biến cố ngẫu nhiên là phép thử ngẫu nhiên.

* Tung đồng thời 3 đồng tiền gồm hai mặt S, N. Xác định các phần tử của Ω . Xác định 3 biến cố ngẫu nhiên mà không phải là biến cố sơ cấp.

c. Biến cố chắc chắn, biến cố không thể.

Biến cố nào mà luôn xảy ra trong phép thử gọi là biến cố chắc chắn(kí hiệu Ω); Biến cố nào mà không thể xảy ra trong phép thử gọi là biến cố không thể(Kí hiệu ϕ)

1.3 Các phép toán trên biến cố

1.3.1. quan hệ giữa các biến cố

* Biến cố A được gọi là kéo theo biến cố B, kí hiệu $A \subseteq B$ nếu A xảy ra thì kéo theo B cũng xảy ra.

* Biến cố A và biến cố B được gọi là bằng nhau, kí hiệu $A = B$ nếu A kéo theo B và B kéo theo A.

Ví dụ:

Tung một con xúc xắc một lần, với $\Omega = \{e_1; e_2; e_3; e_4; e_5; e_6\}$

Gọi A là biến cố mặt chẵn xuất hiện; B là biến cố mặt lẻ xuất hiện; C là biến cố mặt chia hết cho 3 xuất hiện.

* Các kết quả sau kết quả nào đúng :

- | | | | |
|------------------------|-----------------------------|---------------------------------------|----------------------|
| a) $\{e_1\} \subset A$ | b) $\{e_2\} \subset A$ | c) $A = \{e_2; e_4; e_6\}$ | d) $A \subset B$ |
| e) $C \subset A$ | f) $\{e_2; e_5\} \subset B$ | g) $A \subset \{e_1; e_2; e_4; e_6\}$ | h) $A \cap B = \phi$ |

* Xác định các phần tử cho các biến cố $A, B, C, A \cap B, A \cap C, B \cap C, A \cup B, A \cup C, B \cup C$ và mô tả bằng lời các biến cố ngẫu nhiên này

1.3.2 Các phép toán

Cho A và B là hai biến cố ngẫu nhiên của cùng một phép thử.

a. Phép cộng: Tổng của hai biến cố A và B , kí hiệu $A \cup B$ là biến cố xảy ra khi và chỉ khi ít nhất một trong hai biến cố A, B xảy ra.

b. Phép nhân: Tích của hai biến cố A và B , kí hiệu $A \cap B$ là biến cố xảy ra khi và chỉ khi hai biến cố A, B đồng thời xảy ra.

c. Phép trừ: Hiệu của hai biến cố A và B , kí hiệu $A \setminus B$ là biến cố xảy ra khi và chỉ khi biến cố A xảy ra mà biến cố B không xảy ra.

Định nghĩa :

+ Ta gọi $\bar{A} = \Omega \setminus A$ là biến cố đối lập của biến cố A

+ Hai biến cố A, B được gọi là xung khắc nếu $A \cap B = \emptyset$

Chú ý:

Những tính chất của phép cộng, nhân và trừ giống như các tính chất của phép hợp, giao và hiệu của các tập hợp

Yêu cầu SV:

Xét không gian biến cố sơ cấp $\Omega = \{e_1, e_2, e_4, e_6\}$

Gọi A là biến cố xuất hiện mặt chẵn

B là biến cố xuất hiện mặt lẻ

C là biến cố xuất hiện mặt chia hết cho 3

Đáp án nào đúng, đáp án nào sai:

- a) $B = \bar{A}$
- b) A, B xung khắc
- c) $C = A \cap B$
- d) $A \setminus B$ là biến cố xuất hiện mặt chẵn
- e) $A \setminus C$ là biến cố xuất hiện mặt hai chấm hoặc bốn chấm
- f) $A \setminus C$ là biến cố xuất hiện mặt hai chấm
- g) $A \cup C$ là biến cố xuất hiện mặt chẵn hoặc ba chấm
- h) $B = \{e_2\} \cup \{e_3\} \cup \{e_5\}$

2. Hệ đầy đủ các biến cố:

Định nghĩa:

Dãy n biến cố $B_1, B_2, \dots, B_n$ lập thành một hệ đầy đủ các biến cố nếu nó thỏa mãn các điều kiện sau:

$$a) B_1 \cup B_2 \cup \dots \cup B_n = \Omega$$

$$b) B_i \cap B_j = \phi, \forall i \neq j$$

Yêu cầu SV:

Các đáp án sau đâu đúng, đâu sai:

1) Cho $\Omega = \{e_1, e_2, \dots, e_n\}$, khi đó hệ $e_1, e_2, \dots, e_n$ lập thành hệ đầy đủ

2) Gieo đồng thời 2 đồng tiền gồm hai mặt S, N.

Gọi NN là biến cố hai đồng tiền xuất hiện mặt ngửa.

SS là biến cố hai đồng tiền xuất hiện mặt sấp.

SN là biến cố đồng tiền thứ nhất xuất hiện mặt sấp, đồng tiền thứ 2 xuất hiện mặt ngửa.

NS là biến cố đồng tiền thứ nhất xuất hiện mặt ngửa, đồng tiền thứ 2 xuất hiện mặt sấp.

A là biến cố có một đồng tiền xuất hiện mặt sấp.

$$a) \Omega = \{NN; NS; SN; SS\}$$

b) Phép thử này có 4 biến cố sơ cấp

c) Hệ biến cố NN, NS, SN, SS là hệ đầy đủ

$$d) A = \{NS; SN\}$$

e) Hệ biến cố NN, A, SS lập thành hệ đầy đủ.

$$f) A = NS \cup SN$$

3. Các định nghĩa xác suất

3.1 Định nghĩa xác suất cổ điển

Định nghĩa

Với không gian biến cố sơ cấp Ω hữu hạn phần tử, các biến cố sơ cấp đồng khả năng. A là một biến cố trong không gian Ω . Khi đó xác suất (khả năng) biến cố A xảy ra được xác định :

$$P(A) = \frac{n(A)}{n(\Omega)}$$

Trong đó: + $n(A)$ là số biến cố sơ cấp (kết quả) có trong A (hay là số kết quả thuận lợi cho A xảy ra)

+ $n(\Omega)$ là số biến cố sơ cấp (kết quả) của không gian Ω (hay là số kết quả có thể xảy ra).

Ví dụ: Tung một con xúc xắc cân đối và đồng chất.

Gọi e_i là biến cố xuất hiện mặt i chấm ($i=1,2,\dots,6$)

A là biến cố xuất hiện mặt chẵn.

B là biến cố xuất hiện mặt chia hết cho 3

Ta thấy: + Các e_i đồng khả năng vì $P(e_i) = \frac{1}{6} \quad \forall i = 1,2,\dots,6$

+ $A = \{e_2, e_4, e_6\}$: có 3 kết quả (biến cố sơ cấp) thuận lợi cho A xảy ra.

+ $B = \{e_3, e_6\}$: có 2 kết quả (biến cố sơ cấp) thuận lợi cho B xảy ra.

+ $\Omega = \{e_1; e_2; e_3; e_4; e_5; e_6\}$: Có 6 kết quả (biến cố sơ cấp) có thể xảy ra.

$$\text{Do đó: } P(A) = \frac{n(A)}{n(\Omega)} = \frac{3}{6} = 0.5; \quad P(B) = \frac{n(B)}{n(\Omega)} = \frac{2}{6} = 0.333$$

Ví dụ:

1) Một đợt xổ số phát hành 10^6 vé số, trong đó có 1 giải đặc biệt (6 số); 10 giải nhất (5 số), 10 giải nhì (5 số), 20 giải ba (5 số); 70 giải tư (5 số); 100 giải năm (4 số); 300 giải sáu (4 số); 1000 Giải bảy (3 số); 10000 giải tám (2 số); 9 giải phụ đặc biệt và 45 giải khuyến khích. Một người mua ngẫu nhiên một tờ vé số. Tìm xác suất để người đó:

a) Trúng giải đặc biệt; giải nhất; giải tư; giải tám.

b) trúng số.

2) Khi lai hai cây đậu có kiểu gen Aa. Tính xác suất để thế hệ con mang kiểu gen:

a) aa

b) AA

c) Dị hợp tử

d) đồng hợp tử

3) Một hộp gồm 5 bi trắng, 4 bi đỏ. Từ hộp đó lấy ngẫu nhiên cùng ra 2 bi.

a) Không gian biến cố sơ cấp có bao nhiêu phần tử.

b) Gọi B là biến cố lấy được hai bi đỏ. Tìm $P(B)$

c) Gọi C là biến cố lấy được hai bi khác màu. Tìm $P(C)$

d) Gọi D là biến cố lấy được hai bi cùng màu. Tìm $P(D)$

3.2 Định nghĩa xác suất tần suất

Qua định nghĩa ở mục 3.1 ta thấy nó đòi hỏi không gian biến cố sơ cấp Ω hữu hạn phần tử và lại đồng khả năng. Vì vậy để khắc phục nhược điểm đó ta xét định nghĩa sau:

Giả sử một phép thử có thể lặp lại n lần độc lập, trong đó biến cố A xuất hiện m lần trong n lần thực hiện phép thử. Khi đó ta gọi $f = \frac{m}{n}$ là tần suất xuất hiện biến cố A. Người ta kiểm chứng được

khi số lần lặp n càng lớn thì tỉ số $\frac{m}{n}$ tiến về một giá trị cố định p nào đó,

Ví dụ: Nhà toán học Pearson và Buffon đã làm thực nghiệm gieo nhiều lần một đồng tiền cân đối và đồng chất. kết quả được ghi lại như sau:

Người làm thí nghiệm	Số lần gieo	Số lần xuất hiện mặt ngửa	$f = \frac{m}{n}$
Buffon	4040	2048	0.508
Pearson(lần 1)	12000	6019	0.5016
Pearson(lần 2)	24000	12012	0.5005

Với bảng thực nghiệm trên cho thấy xác suất để mặt ngửa xuất hiện là $p = 0.5$

Định nghĩa

Khi số lần lặp n của phép thử càng lớn, tần suất $\frac{m}{n}$ của biến cố A tiến về một số cố định p , ta nói biến cố A ổn định ngẫu nhiên và p chính là xác suất của biến cố A .

Và như vậy khi n đủ lớn ta có thể xấp xỉ $p \approx \frac{m}{n}$, nghĩa là: $P(A) \approx \frac{m}{n}$

Ví dụ: Để biết xác suất bắn trúng mục tiêu của một xạ thủ là bao nhiêu, người ta tiến hành cho xạ thủ đó bắn n viên đủ lớn (mỗi lần bắn xem như thực hiện một phép thử), sau đó ghi nhận số viên đạn trúng mục tiêu (giả sử m viên trúng mục tiêu).

Khi đó: $f = \frac{m}{n}$ được xem là xác suất trúng mục tiêu của xạ thủ đó

4. Các công thức tính xác suất

4.1 Công thức cộng

Cho n biến cố ngẫu nhiên $A_1, A_2, \dots, A_n$ trên cùng không gian biến cố sơ cấp Ω

Khi đó:

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k) - \sum_{1 \leq k < j \leq n} P(A_k \cap A_j) + \sum_{1 \leq k < j < l \leq n} P(A_k \cap A_j \cap A_l) + \dots + (-1)^{n-1} P(A_1 \cap A_2 \cap \dots \cap A_n)$$

* Nếu các biến cố $A_1, A_2, \dots, A_n$ đôi một xung khắc thì $P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k)$

* Với hai biến cố A, B :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$P(A \cup B) = P(A) + P(B), \text{ (Với } A, B \text{ xung khắc)}$$

* Với ba biến cố A, B, C :

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$
$$P(A \cup B \cup C) = P(A) + P(B) + P(C), \text{ (Với } A, B, C \text{ đôi một xung khắc)}$$

Ví dụ:

1) Từ một hộp gồm 3 bi trắng, 5 bi đỏ lấy ngẫu nhiên cùng lúc ra 3 bi.

Gọi A là biến cố lấy được 2 đỏ, 1 trắng

B là biến cố lấy được 2 trắng, 1 đỏ

Tìm $P(A)$, $P(B)$, $P(A \cup B)$

2) Có 3 bức thư khác nhau và 3 phong bì có ghi địa chỉ sẵn, cho ngẫu nhiên 3 bức thư vào 3 phong bì đó. Tìm xác suất trong 3 bức thư đó có ít nhất một bức thư gửi đúng địa chỉ

4.2 Xác suất có điều kiện, công thức nhân

a. Xác suất điều kiện

Ví dụ: Từ bộ bài Lútukho(52 lá), rút ngẫu nhiên ra 1 lá.

Gọi A là biến cố rút được lá hai

B là biến cố rút được lá đỏ

Tìm: a. $P(A)$, $P(B)$, $P(A \cap B)$

b. $P(A|B)$: Xác suất lá rút được lá hai, biết lá rút được là lá đỏ

Giải

$$a) P(A) = \frac{4}{52} = \frac{1}{13}, \quad P(B) = \frac{26}{52} = \frac{1}{2}, \quad P(A \cap B) = \frac{2}{52} = \frac{1}{26}$$

$$b) P(A|B) = \frac{n(A \cap B)}{n(B)} = \frac{2}{26} = \frac{1}{13}$$

* Ta gọi $P(A|B)$ là xác suất của biến cố A với điều kiện biến cố B đã xảy ra và nó được tính bởi công thức

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

* Hai biến cố A và B gọi là độc lập nếu $P(A|B) = P(A)$; $P(B|A) = P(B)$

b. Công thức nhân

* Từ công thức xác suất điều kiện ta có:

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

* Nếu A, B độc lập thì $P(A \cap B) = P(A)P(B)$

* Nếu $A_1, A_2, \dots, A_n$ là các biến cố cùng không gian Ω thì:

$$P\left(\bigcap_{k=1}^n A_k\right) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \dots P(A_n|A_1 \cap \dots \cap A_{n-1})$$

* Nếu $A_1, A_2, \dots, A_n$ là các biến cố độc lập thì: $P\left(\bigcap_{k=1}^n A_k\right) = \prod_{k=1}^n P(A_k)$

Chú ý: Nếu không có gì nhầm lẫn thì ta có thể sử dụng kí hiệu $A+B$ thay cho $A \cup B$; $A.B$ thay cho $A \cap B$

4.3 Công thức xác suất đầy đủ và công thức Bayes

Trong không gian Ω cho hệ đầy đủ các biến cố $A_1, A_2, \dots, A_n$, A là một biến cố bất kỳ của Ω , Khi đó ta có:

$$a) P(A) = P(A_1)P(A|A_1) + P(A_2)P(A|A_2) + \dots + P(A_n)P(A|A_n),$$

(Công thức xác suất đầy đủ)

b) Nếu $P(A) > 0$ thì $P(A_k|A) = \frac{P(A_k)P(A|A_k)}{P(A)}$, $k=1,2,\dots,n$, (Công thức Bayes)

Chứng minh

a) Ta có:

$$A = A \cap \Omega = A \cap \bigcup_{k=1}^n A_k, \text{ Vì } A_1, A_2, \dots, A_n \text{ là hệ đầy đủ}$$

$$A = \bigcup_{k=1}^n (A \cap A_k) \Rightarrow P(A) = \sum_{k=1}^n P(A \cap A_k), \text{ Vì } A_1, A_2, \dots, A_n \text{ Xung khắc đôi một}$$

$$P(A) = \sum_{k=1}^n P(A_k)P(A|A_k).$$

b) Ta có:

$$P(A_k|A) = \frac{P(A \cap A_k)}{P(A)} = \frac{P(A_k)P(A|A_k)}{P(A)}$$

Yêu cầu SV

1) Từ một hộp gồm 10 bi trắng, 5 bi đỏ, lấy lần lượt không hoàn lại ra 2 bi.

a) Tính xác suất 2 bi lấy ra cùng màu đỏ

b) Tính xác suất 2 bi lấy ra khác màu nhau

2) Có hai lô sản phẩm, lô 1 có 100 sản phẩm trong đó có 10 phế phẩm; lô 2 có 90 sản phẩm trong đó có 5 phế phẩm.

a) Lấy ngẫu nhiên mỗi lô 1 sản phẩm. Tìm xác suất trong 2 sản phẩm lấy ra có 1 phế phẩm

b) Chọn ngẫu nhiên 1 lô, rồi từ lô đó lấy ngẫu nhiên ra 2 sản phẩm. Tìm xác suất trong 2 sản phẩm lấy ra có 1 phế phẩm

4.4 Công thức xác suất nhị thức

Cho n phép thử độc lập (kết quả xảy ra hay không xảy ra của phép thử này không ảnh hưởng đến kết quả xảy ra hay không xảy ra của phép thử khác), mỗi phép thử ta chỉ quan tâm đến hai biến cố A và $\bar{A}$ và $P(A) = p$ (không đổi với mỗi phép thử)

Xác suất để biến cố A xuất hiện k lần trong n lần thực hiện phép thử được xác định:

$$P_n(k) = C_n^k p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n$$

Chứng minh

Gọi B là biến cố trong n lần thực hiện phép thử có k lần biến cố A xảy ra

$$\Rightarrow B = A \dots A \bar{A} \dots \bar{A} + A \dots A \bar{A} \bar{A} \bar{A} \dots \bar{A} + \dots, \text{ (có } C_n^k \text{ hạng tử)}$$

$$\Rightarrow P(B) = P(A \dots A \bar{A} \dots \bar{A}) + P(A \dots A \bar{A} \bar{A} \bar{A} \dots \bar{A}) + \dots, \text{ (có } C_n^k \text{ số hạng)}$$

$$\Rightarrow P(B) = [P(A)]^k [P(\bar{A})]^{n-k} + [P(A)]^k [P(\bar{A})]^{n-k} + \dots, \text{ (có } C_n^k \text{ số hạng)}$$

$$\Rightarrow P(B) = C_n^k p^k (1-p)^{n-k}$$

Yêu cầu SV

Tung 20 lần một con xúc xắc cân đối và đồng chất. Tìm xác suất

- Có 5 lần xuất hiện mặt 3 chấm
- Có 8 lần xuất hiện mặt chẵn
- Có ít nhất 2 lần xuất hiện mặt chẵn

II. BIẾN NGẪU NHIÊN

1. Khái niệm biến ngẫu nhiên và hàm phân phối

1. 1. Khái niệm biến ngẫu nhiên:

Ví dụ : Tung 3 lần một đồng tiền cân đối và đồng chất Khi đó ta có $\Omega = \{ NNN, NNS, NSN, SNN, NSS, SSN, SSS \}$

Trong đó: N là biến cố xuất hiện mặt ngửa trong mỗi lần tung

S là biến cố xuất hiện mặt sấp trong mỗi lần tung

Trên không gian Ω ta xác định một hàm X lấy giá trị trên R như sau:

$$X: \Omega \rightarrow R$$

$$\omega \rightarrow X(\omega) : \text{số lần xuất hiện mặt ngửa}$$

$$\text{Ta thấy : } X(SSS) = 0$$

$$X(SSN) = X(SNS) = X(NSS) = 1$$

$$X(SNN) = X(NSN) = X(NNS) = 2$$

$$X(NNN) = 3$$

Như vậy tập giá trị của $X(\omega) : \{ 0, 1, 2, 3 \}$

Trong ví dụ trên X được gọi là biến ngẫu nhiên và ta cũng thấy rằng: $\forall x \in R$ luôn tồn tại biến cố $A = \{ \omega : X(\omega) < x \}$

Chẳng hạn:

$$+ x \leq 0 \Rightarrow A = \emptyset$$

$$+ 0 < x \leq 1 \Rightarrow A = \{ SSS \}$$

$$+ 1 < x \leq 2 \Rightarrow A = \{ SSS, SNS, NSS, SSN \}$$

$$+ 2 < x \leq 3 \Rightarrow A = \{ SSS, SNS, NSS, SSN, SNN, NSN, NNS \}$$

$$+ x > 3 \Rightarrow A = \Omega$$

Dựa vào đặc điểm trên, ta có định nghĩa biến ngẫu nhiên như sau:

Định nghĩa

Biến ngẫu nhiên X là một hàm xác định trên không gian biến cố sơ cấp Ω và nhận giá trị trong R sao cho $\forall x \in R$ tồn tại biến ngẫu nhiên $A = \{ \omega : X(\omega) < x \}$

+ Biến ngẫu nhiên thường kí hiệu: $X, Y, Z, \dots$

+ Giá trị của biến ngẫu nhiên kí hiệu: $x, y, z, \dots$

+ Nếu không có gì nhầm lẫn thì $X(\omega) = x$, đôi khi ta viết $X = x$

Ta có thể hiểu biến ngẫu nhiên là đại lượng nhận giá trị trong tập số thực R , phụ thuộc vào kết quả của phép thử.

Ví dụ: Ta có $X(SSS) = 0$, ta có thể viết: $X = 0$, còn $A = \{ \omega : X(\omega) < x \} \{ \omega : X(\omega) < x \}$ ta viết $A = (X < x)$

Định nghĩa

a) Biến ngẫu nhiên X được gọi là biến ngẫu nhiên rời rạc nếu tập giá trị của X hữu hạn hoặc vô hạn đếm được

b) Biến ngẫu nhiên X được gọi là liên tục nếu tập giá trị của X là khoảng (a, b) , a có thể là $-\infty$, b có thể là $+\infty$

Yêu cầu SV:

Hãy xác định các biến ngẫu nhiên cho các ví dụ sau; tìm miền giá trị của nó và tính xác suất ứng với từng giá trị của nó.

a) Bắn không hạn chế vào mục tiêu, bắn cho tới khi nào có viên đạn trúng mục tiêu thì dừng lại

b) Từ một hộp có 7 bi đỏ, 3 bi xanh và 10 bi vàng lấy lần lượt có hoàn lại 4 viên bi

1.2. Hàm phân phối của biến ngẫu nhiên:

Định nghĩa

Cho X là biến ngẫu nhiên, khi đó luôn tồn tại $P(\{ \omega : X(\omega) < x \}) \forall x$ và ta gọi

$F(x) = P(X < x)$: là hàm phân phối xác suất của biến ngẫu nhiên X

Ví dụ: Bắn 3 viên đạn độc lập vào mục tiêu Gọi X là số viên đạn trúng đích Xác suất bắn trúng mỗi viên là 0,6

+ X là biến ngẫu nhiên, tập giá trị: $\{0, 1, 2, 3\}$

+ Không gian biến cố sơ cấp $\Omega = \{ \overline{AAA}, \overline{AA\bar{A}}, \overline{A\bar{A}\bar{A}}, \overline{A\bar{A}A}, \overline{AA\bar{A}}, \overline{A\bar{A}A}, \overline{AA\bar{A}}, \overline{AA\bar{A}} \}$

Trong đó A là biến cố bắn trúng đích

Ta có:

$$+ P(X = 0) = 0,4^3$$

$$+ P(X = 1) = 3.0,4^3.0,6$$

$$+ P(X = 2) = 3.0,4.0,6^2$$

$$+ P(X = 4) = 0,6^3$$

Ta có hàm phân phối:

$$+ F(x) = P(X \leq x) = \begin{cases} P(\phi), & x \leq 0 \\ P(X = 0), & 0 < x \leq 1 \\ P(X = 0) + P(X = 1), & 1 < x \leq 2 \\ P(X = 0) + P(X = 1) + P(X = 2), & 2 < x \leq 3 \\ 1, & x > 3 \end{cases}$$

$$= \begin{cases} 0, & x \leq 0 \\ 0,4^3, & 0 < x \leq 1 \\ 0,4^3 + 3.0,4^3.0,6, & 1 < x \leq 2 \\ 0,4^3 + 3.0,4^3.0,6 + 3.0,4.0,6^2, & 2 < x \leq 3 \\ 1, & x > 3 \end{cases}$$

2. Các tính chất hàm phân phối:

i) Hàm phân phối là hàm đơn điệu tăng

ii) Hàm phân phối $F(x)$ liên tục trái, nghĩa là $\lim_{x \rightarrow a} F(x) = F(a)$

iii) $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow +\infty} F(x) = 1$

*** Yêu cầu:**

1) Giả sử X có hàm phân phối

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 < x \leq 1 \\ 1, & x > 1 \end{cases}$$

a) Vẽ đồ thị hàm $F(x)$

b) Tính $P(-1 \leq x < \frac{1}{2})$ và $P(0 < x \leq 1)$

2) Giả sử X có hàm phân phối:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-ax}, & x > 0 \end{cases}$$

a) Tìm a và vẽ đồ thị hàm F(x)

b) Tính P(-1 ≤ x < 1)

3) Phân phối rời rạc và phân phối liên tục:

3.1. Phân phối rời rạc:

3.1.1. Bảng phân phối xác suất

Cho X là biến ngẫu nhiên rời rạc nhận các giá trị: $x_1, x_2, \dots, x_n, \dots$ với xác suất tương ứng như sau:

X	x_1	x_2	...	x_n	...
P(X = x_i)	P_1	P_2	...	P_n	...

Trong đó: $P_1 + P_2 + \dots + P_n + \dots = 1$

+ Bảng trên được gọi là bảng phân phối xác suất của X

+ Nếu $x_1 < x_2 < \dots < x_n < \dots$ thì hàm phân phối của X có dạng:

$$F(x) = \begin{cases} 0 & \text{nếu } x \leq x_1 \\ P_1 & \text{nếu } x_1 < x \leq x_2 \\ P_1 + p_2 & \text{nếu } x_2 < x \leq x_3 \\ \vdots & \\ P_1 + p_2 + \dots + p_k & \text{nếu } x_k < x \leq x_{k+1} \end{cases}$$

Yêu cầu:

Một gia đình có ba người con, giả sử xác suất sinh con trai và sinh con gái là như nhau. Gọi X là số con trai của gia đình đó. Tìm phân phối xác suất (bảng phân phối xác suất) và hàm phân phối xác suất của X

3.1.2. Hàm mật độ xác suất của X

Cho X là biến ngẫu nhiên rời rạc nhận các giá trị: $x_1, x_2, \dots, x_n, \dots$, hàm số được định nghĩa: $f(x) = P(X=x), x = x_1, x_2, \dots, x_n, \dots$ được gọi là hàm mật độ xác suất của X

Chú ý: Bảng phân phối xác suất của X còn gọi là hàm mật độ xác suất của X dưới dạng bảng.

Yêu cầu

1) Bắn 5 viên đạn độc lập với nhau vào một mục tiêu (trong điều kiện như nhau), xác suất bắn trúng mục tiêu của mỗi lần bắn là như nhau và bằng 0,2. Gọi X là số viên đạn bắn trúng mục tiêu.

a) Tìm phân phối xác suất của X, cho biết X thuộc dạng phân phối nào?

b) muốn mục tiêu bị phá hủy phải có ít nhất 3 viên đạn trúng mục tiêu. Tìm xác suất để mục tiêu bị phá hủy

3.2. Phân phối liên tục

3.2.1 Hàm mật độ xác suất

Cho X là biến ngẫu nhiên liên tục có hàm phân phối xác suất $F(x)$. Hàm số $f(x)$ được gọi là hàm mật độ xác suất của X nếu thỏa:

$$F(x) = \int_{-\infty}^x f(t)dt, \quad x \in R$$

+ Tại những điểm x làm cho $f(x)$ liên tục thì $F'(x) = f(x)$

+ Hàm mật độ xác suất của X tồn tại là duy nhất

3.2.2 Tính chất của hàm mật độ xác suất

+ $f(x) \geq 0, \forall x$

$$+ \int_{-\infty}^{+\infty} f(x)dx = 1$$

$$+ P(\alpha \leq X < \beta) = \int_{\alpha}^{\beta} f(x)dx$$

+ $P(X = \alpha) = 0, P(\alpha \leq X < \beta) = P(\alpha \leq X \leq \beta) = P(\alpha < X < \beta)$

Yêu cầu

1) Cho biến ngẫu nhiên X có hàm mật độ

$$f(x) = \begin{cases} me^{-\lambda x}, & x > 0, \lambda > 0 \\ 0, & x \leq 0 \end{cases}$$

a) Tìm tham số m

b) Tìm hàm phân phối xác suất của X và tính $P(0 < X < 1)$

4. Các đặc trưng của biến ngẫu nhiên

4.1 Kỳ vọng(trung bình) :

Kỳ vọng của X đặc trưng cho giá trị trung tâm của biến ngẫu nhiên X

$$E(X) = \mu = \begin{cases} \sum_{x_i} x_i P(X = x_i) & , X \text{ rời rạc} \\ \int_{-\infty}^{+\infty} xf(x)dx & , X \text{ liên tục} \end{cases}$$

Tính chất

- + $E(C) = C$, (C hằng số)
- + $E(CX) = CE(X)$
- + Nếu X, Y có kỳ vọng thì $E(X+Y) = E(X)+E(Y)$
- + Nếu X, Y độc lập và có kỳ vọng thì $E(XY) = E(X)E(Y)$

4.2 Phương sai: Đặc trưng cho mức độ phân tán của các giá trị của X so với kỳ vọng

$$Var(X) = \sigma^2 = E(X - \mu)^2 = E(X^2) - \mu^2$$

$$\text{Trong đó: } E(X^2) = \begin{cases} \sum_{x_i} x_i^2 P(X = x_i) & , X \text{ rời rạc} \\ \int_{-\infty}^{+\infty} x^2 f(x)dx & , X \text{ liên tục} \end{cases}$$

Tính chất

- + $Var(C) = 0$, (C hằng số)
- + $Var(CX) = C^2E(X)$
- + Nếu X, Y độc lập và có phương sai thì $Var(X+Y) = Var(X)+Var(Y)$

4.3 Mod

Mod là giá trị của X(kí hiệu x_{mod}) mà tại đó hàm mật độ đạt giá trị lớn nhất.

- + Trường hợp biến ngẫu nhiên rời rạc thì $P(X=x_{\text{mod}})$ là lớn nhất
- + x_{mod} có thể có duy nhất một giá trị cũng có thể có nhiều hơn một giá trị.

Yêu cầu

1) Cho biến ngẫu nhiên X có hàm mật độ xác suất:

$$f(x) = \begin{cases} e^{-x} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$$

Tìm trung bình μ , phương sai σ^2 và x_{Mod}

2) Trong hộp gồm 7 sản phẩm tốt và 3 sản phẩm hỏng, lấy ngẫu nhiên ra 3 sản phẩm. Gọi X là số sản phẩm hỏng có trong 3 sản phẩm lấy ra. Tìm trung bình μ , phương sai và x_{Mod}

5. Các quy luật phân phối quan trọng

5.1 Phân phối nhị thức

Biến ngẫu nhiên X được gọi là có phân phối nhị thức nếu hàm mật độ xác suất của nó có dạng:

$$f(x) = P(X = x) = C_n^x p^x (1-p)^{n-x}, x = 0, 1, \dots, n$$

Kí hiệu: $X \sim B(n, p)$, n và p gọi là hai tham số của phân phối nhị thức

+ Nếu $X \sim B(n, p)$ thì $E(X) = np$; $Var(X) = np(1-p)$.

* Đặc biệt: Nếu n = 1 thì phân phối B(1, p) gọi là phân phối Bernouli

5.2 Phân phối poisson

Biến ngẫu nhiên X được gọi là có phân phối Poisson nếu hàm mật độ xác suất của nó có dạng:

$$f(x) = P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}, x = 0, 1, 2, \dots, \lambda > 0$$

Kí hiệu: $X \sim P(\lambda)$, λ gọi là tham số của phân phối Poisson

+ Nếu $X \sim P(\lambda)$ thì $E(X) = \lambda$; $Var(X) = \lambda$.

5.3 Mối liên hệ giữa phân phối nhị thức và phân phối Poisson

Định lý: Cho X có phân phối nhị thức B(n, p).

$$\text{Nếu } np \xrightarrow{n \rightarrow \infty} \lambda, p \xrightarrow{n \rightarrow \infty} 0 \text{ thì } P(X = x) = C_n^x p^x (1-p)^{n-x} \xrightarrow{n \rightarrow \infty} \frac{\lambda^x}{x!} e^{-\lambda}$$

Chứng minh

Do $np \xrightarrow{n \rightarrow \infty} \lambda \Rightarrow \lambda \approx np$, (khi n đủ lớn)

$$\Rightarrow p = \frac{\lambda}{n}$$

Ta có:

$$P(X = x) = C_n^x p^x (1-p)^{n-x} = \frac{n(n-1)\dots(n-x+1)}{x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

$$\lim_{n \rightarrow \infty} P(X = x) = \frac{\lambda^x}{x!} \lim_{n \rightarrow \infty} \frac{n(n-1)\dots(n-x+1)}{n^x} \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{\lambda^x}{x!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = \frac{\lambda^x}{x!} e^{-\lambda}.$$

* Như vậy khi n khá lớn, p khá nhỏ thì ta xấp xỉ

$$P(X = x) = C_n^x p^x (1-p)^{n-x} \approx \frac{\lambda^x}{x!} e^{-\lambda}, \text{ với } \lambda \approx np$$

Ví dụ:

1) Bắn 5 viên đạn độc lập với nhau vào một mục tiêu (trong điều kiện như nhau), xác suất bắn trúng mục tiêu của mỗi lần bắn là như nhau và bằng 0,2. Gọi X là số viên đạn bắn trúng mục tiêu.

a) Tìm phân phối xác suất của X, cho biết X thuộc dạng phân phối nào?

b) muốn mục tiêu bị phá hủy phải có ít nhất 3 viên đạn trúng mục tiêu. Tìm xác suất để mục tiêu bị phá hủy

2) Một lô bóng đèn điện tử gồm 10000 bóng, xác suất để mỗi bóng hỏng là 0,001. gọi X là số bóng đèn hỏng của lô hàng

a) Xác định dạng phân phối xác suất của X

b) Tìm xác suất trong lô có đúng 3 bóng hỏng; ít nhất 4 bóng hỏng

5.4 Phân phối chuẩn

Biến ngẫu nhiên X được gọi là có phân phối chuẩn nếu hàm mật độ xác suất của nó có dạng:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in R$$

Kí hiệu: $X \sim N(\mu, \sigma^2)$, với μ, σ^2 gọi là hai tham số của phân phối chuẩn

+ Nếu $X \sim N(\mu, \sigma^2)$ thì $E(X) = \mu$; $Var(X) = \sigma^2$.

* Đặc biệt: nếu $\mu = 0, \sigma^2 = 1$ thì phân phối $N(0;1)$ gọi là phân phối chuẩn tắc

+ Hàm mật độ của phân phối $N(0;1)$:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in R$$

+ Nếu $X \sim N(0;1)$ thì

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \Phi(\beta) - \Phi(\alpha)$$

(Trong đó: $\varphi(t) = \int_0^t \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$ là hàm Laplace)

Hoặc
$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \Phi(\beta) - \Phi(\alpha)$$

(Trong đó: $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ là hàm PPXS của X)

Định lý chuẩn hóa

Nếu $X \sim N(\mu, \sigma^2)$ thì $Z = \frac{X - \mu}{\sigma} \sim N(0;1)$

Định lý Moivre-Laplace

Nếu $X \sim B(n, p)$, n khá lớn, p không quá gần 0 và 1 thì $X \sim N(np; np(1-p))$

Yêu cầu: Điều tra ngẫu nhiên 10000 trẻ em, giả sử xác suất sinh con trai và con gái như nhau và bằng 1/2. Gọi X là số trẻ em trai. Tính xác suất để X nằm trong khoảng 4000 đến 6000; dưới 5000; trên 6000.

5.5 Phân phối Gamma và khi bình phương

5.5.1 Hàm Gamma

Hàm gamma được xác định: $\Gamma(t) = \int_0^{+\infty} x^{t-1} e^{-x} dx$, với $t > 0$

+ Áp dụng phương pháp tích phân từng phần ta có: $\Gamma(t+1) = t \cdot \Gamma(t)$, với $t > 0$

5.5.2 Phân phối Gamma

Biến ngẫu nhiên X được gọi là có phân phối Gamma nếu hàm mật độ xác suất của nó có dạng:

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}, \text{ với } x > 0$$

Kí hiệu: $X \sim G(\alpha, \beta)$, với α, β gọi là hai tham số của phân phối Gamma

+ Nếu $X \sim G(\alpha, \beta)$ thì $M(t) = \frac{1}{(1 - \beta t)^\alpha}$; $E(X) = \alpha\beta$; $Var(X) = \alpha\beta^2$.

5.6 Phân phối khi bình phương

Phân phối khi bình phương là phân phối Gamma $G(\alpha, \beta)$, với $\alpha = \frac{r}{2}$, $\beta = 2$, trong đó r

$= 1, 2, 3, \dots$

$$f(x) = \frac{1}{\Gamma(\frac{r}{2})2^{\frac{r}{2}}} x^{\frac{r}{2}-1} e^{-\frac{x}{2}}, \text{ với } x > 0$$

Kí hiệu: $X \sim \chi^2(r)$, với r gọi là tham số (bậc tự do) của phân phối khi bình phương

+ Nếu $X \sim \chi^2(r)$ thì $M(t) = \frac{1}{(1-2t)^{\frac{r}{2}}}$; $E(X) = r$; $Var(X) = 2r$.

Định lý:

Nếu $X \sim \chi^2(r)$, $Y \sim \chi^2(s)$ và X, Y là hai biến ngẫu nhiên độc lập thì $T = X+Y$ là biến ngẫu nhiên có phân phối $\chi^2(r+s)$

Định lý:

Nếu $X_1, X_2, \dots, X_n$ là n biến ngẫu nhiên độc lập có phân phối $N(0;1)$ thì

$$T = X_1^2 + X_2^2 + \dots + X_n^2 \text{ là biến ngẫu nhiên có phân phối } \chi^2(n)$$

5.7 Phân phối Student

Cho hai biến ngẫu nhiên độc lập $X \sim N(0;1)$, $Y \sim \chi^2(n)$, khi đó phân phối của biến ngẫu nhiên $T = \frac{X}{\sqrt{\frac{Y}{n}}}$ được gọi là phân phối Student

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi \cdot n} \Gamma(\frac{n}{2})} \cdot \frac{1}{(1 + \frac{x^2}{n})^{\frac{n+1}{2}}}$$

Kí hiệu: $X \sim T(n)$, với n gọi là tham số (bậc tự do) của phân phối Student

5.8 Phân phối Fisher

Cho hai biến ngẫu nhiên độc lập $X \sim \chi^2(m)$, $Y \sim \chi^2(n)$, khi đó phân phối của biến ngẫu nhiên $F = \frac{\frac{X}{m}}{\frac{Y}{n}}$ được gọi là phân phối Fisher

$$f(x) = \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} \cdot \left(\frac{n}{m}\right)^{\frac{n}{2}} \cdot \frac{x^{\frac{n}{2}-1}}{(1 + \frac{n}{m}x)^{\frac{m+n}{2}}}, \quad x > 0$$

Kí hiệu: $X \sim F(m,n)$, với m, n gọi là hai tham số (bậc tự do) của phân phối Fisher

Bài Tập củng cố chương I

1) Một lớp có 50 sinh viên (trong đó có 30 nam và 20 nữ). Chọn ngẫu nhiên một nhóm gồm 4 sinh viên. Tính các xác suất:

- a) Có 2 nam trong số 4 sinh viên được chọn
- b) Có ít nhất một sinh viên nam trong số 4 sinh viên được chọn
- c) Có ít nhất 2 sinh viên nam trong số 4 sinh viên được chọn
- d) Không có sinh viên nam trong số 4 sinh viên được chọn

2) Một phân xưởng có 60 công nhân, trong đó có 40 nữ và 20 nam. Tỉ lệ công nhân nữ tốt nghiệp phổ thông trung học là 15%. Còn tỉ lệ này đối với nam là 20%. Gặp ngẫu nhiên một công nhân của phân xưởng. Tính xác suất để gặp người công nhân tốt nghiệp phổ thông trung học.

3) Một hộp đựng 3 bi đỏ và 7 bi xanh. Lấy ngẫu nhiên từ hộp ra một bi. Nếu bi lấy ra màu đỏ thì bỏ vào hộp một bi xanh. Nếu bi lấy ra màu xanh thì bỏ vào hộp một bi màu đỏ. Sau đó từ hộp ta lấy tiếp ra một bi

- a) Tìm xác suất để bi lấy ra lần sau là bi đỏ
- b) Nếu hai bi lấy ra (lấy lần thứ nhất và lần thứ hai) cùng màu. Tìm xác suất để hai bi này cùng màu xanh

4) Một người có 5 chìa khóa nhưng chỉ có 2 chìa khóa mở được cửa. Người đó thử từng chìa (thử xong nếu không mở được khóa để riêng chìa khóa đó ra). Tính xác suất để lần thứ hai người đó mở được khóa

5) Có 3 xạ thủ cùng bắn vào một bia. Mỗi xạ thủ bắn 1 viên. Xác suất bắn trúng bia của xạ thủ thứ nhất, thứ hai, thứ ba tương ứng là: 0,6; 0,7; 0,8. Tính xác suất để hai viên trúng bia

6) Hộp thứ nhất có 8 chai thuốc (trong đó có 3 chai kém phẩm chất). Hộp thứ hai có 5 chai thuốc (trong đó có 2 chai kém phẩm chất). Lấy ngẫu nhiên từ mỗi hộp ra một chai. Tìm xác suất lấy được hai chai thuốc tốt

7) Hộp thứ nhất có 7 sản phẩm loại 1 và 3 sản phẩm loại 2; hộp thứ hai có 5 sản phẩm loại 1 và 3 sản phẩm loại 2. Lấy ngẫu nhiên 1 sản phẩm ở hộp thứ nhất bỏ vào hộp thứ hai rồi sau đó từ hộp thứ hai lấy ngẫu nhiên ra một sản phẩm.

- a) Tìm xác suất sản phẩm lấy ra là loại 1
 - b) Biết sản phẩm lấy ra là loại 2. Tìm xác suất sản phẩm đó được bỏ từ hộp 1 sang
- 8)** Trong một hộp có 12 bóng đèn, trong đó có 3 bóng hỏng. lấy ngẫu nhiên có thứ tự không hoàn lại 3 bóng để dùng. Tìm xác suất để cả 3 bóng đều không hỏng