

MỞ ĐẦU

Trong các giáo trình toán, vật lý ở nhà trường phổ thông người ta thường chỉ xét bài toán trong đó kết quả phép toán được xác định một cách duy nhất. Chẳng hạn, nếu ta thả một hòn đá thì nó sẽ rơi với một gia tốc không đổi. Vị trí của hòn đá ở mọi thời điểm đều có thể tính được. Tuy nhiên còn có nhiều bài toán mà kết quả các kết quả thực hiện trong chúng không được xác định một cách duy nhất, nhưng lại có ý nghĩa lớn lao về mặt khoa học cũng như việc áp dụng trong kỹ thuật, Kinh tế, Y học,... Chẳng hạn, nếu ta gieo đồng tiền thì không thể nói trước rằng khi đồng tiền rơi xuống mặt đất, mặt sấp hay mặt ngửa của đồng tiền sẽ lên trên, ở đây kết quả của phép thử thực hiện không được xác định một cách duy nhất. Hình như trong các bài toán như vậy ta không nên nói trước một điều gì xác định, tuy nhiên ngay với thực tiễn của những trò chơi thông thường cũng chứng tỏ một điều ngược lại là, với một số khá lớn lần gieo đồng tiền thì ta thấy gần một nửa số lần rơi mặt sấp và một nửa số lần rơi mặt ngửa, đây là một quy luật xác định. Trong lý thuyết xác suất người ta nghiên cứu các quy luật dạng đó. Chính việc thiết lập các bài toán cũng được thay đổi căn bản. Chúng ta quan tâm không phải là kết quả của một phép thử xác định mà là cái nhận được sau nhiều lần lặp lại phép thử đó. Nói một cách khác, trong lý thuyết xác suất ta nghiên cứu tính quy luật của các biến cố ngẫu nhiên hàng loạt.

Lý thuyết xác suất xuất hiện và phát triển trong quá trình giải quyết một loạt các bài toán riêng lẻ mang tính trò chơi và ứng dụng. Các kiến thức đầu tiên chúng ta biết được có quan hệ với việc giải các bài toán về trò chơi xuất hiện từ thế kỷ XVI — XVII (D. Cardano, Huyghens, B. Pascal, P. Ferma,...). Sau đó các bài toán ứng dụng bắt đầu xuất hiện và phát triển (đáng kể đầu tiên là các bài toán về đề phòng tai nạn và thiên tai). Dần dần được tách ra một lĩnh vực các bài toán với hình thái riêng biệt cũng như phương pháp giải chúng, hình thành các định nghĩa đầu tiên và các định lý. Định lý đầu tiên thiết lập mối quan hệ giữa lý thuyết và thực hành và là phần đầu nhóm các định lý có tên “các Định lý giới hạn” của lý thuyết xác suất do Bernoulli (1654 - 1705) chứng minh cuối thế kỷ 17. Sau đó sự phát triển của lý thuyết xác suất được tiếp tục trong các công trình của A. Moivre (1667 - 1754), P. Laplace (1749 - 1827), K. Gauss (1777 - 1855), Poisson (1781 - 1840), và đặc biệt trong các công trình của nhà toán học Nga P.L. Chebushhev (1821 - 1894), và các học trò của ông ta A.A. Markov (1856 — 1922), A. M. Liapunov (1857 - 1918). Trong thế kỷ XX sự phát triển lớn nhất của lý thuyết xác suất và việc trình bày nó một cách hoàn thiện như một khoa học toán học đã được giới thiệu trong các công trình của các nhà toán học Xô viết.

Hơn 300 năm phát triển, đến nay nội dung và phương pháp của xác suất thống kê rất phong phú, được áp dụng rộng rãi trong nhiều lĩnh vực. Vì vậy, việc học tập, nghiên cứu môn xác suất thống kê đã trở thành nhu cầu không thể thiếu đối với sinh viên của nhiều ngành của các trường Đại học cũng như của các cán bộ nghiên cứu của hầu hết các ngành khoa học kỹ thuật.

Để nâng cao chất lượng đào tạo, đáp ứng với nhu cầu của sự phát triển xã hội và tạo điều kiện thuận lợi cho sinh viên học tập nghiên cứu môn học này, chúng tôi biên soạn cuốn sách **Xác suất & thống kê**. Qua cuốn sách nhỏ này, chúng tôi mong muốn và hy vọng các bạn sinh viên sẽ đạt kết quả cao trong học tập cũng như áp dụng được các phương pháp của xác suất thống kê trong công việc của mình sau này.

Đối với các bác sỹ, các dược sỹ, các nhà nhà kinh tế, các nhà doanh nghiệp và các chuyên gia nghiệp vụ quản lý, biết thu thập, xử lý các thông tin nghề nghiệp là yêu cầu

không thể thiếu được. Toán học nói chung, lý thuyết xác suất thống kê nói riêng, là công cụ nghiên cứu rất hữu hiệu. Đối với sinh viên các ngành Y khoa, sinh học, kinh tế, kỹ thuật, mục đích cuối cùng của học toán là sử dụng được công cụ này trong công việc của mình. Do đó cuốn sách được viết theo quan điểm thực hành, chú trọng việc vận dụng các phương pháp của xác suất thống kê trong thực tế mà không đi sâu vào việc chứng minh cơ sở lý thuyết toán học một cách chặt chẽ.

Với tinh thần ứng dụng, tốc độ, dễ hiểu và dễ áp dụng vào thực tiễn, cuốn sách chia làm hai phần: phần 1 “**Sơ lược về lý thuyết xác suất**” chỉ trình bày trong hai chương.

Chương 1. CÁC KHÁI NIỆM CƠ BẢN VỀ XÁC SUẤT

Chương 2. LƯỢNG NGẪU NHIÊN HÀM PHÂN PHỐI

Cuối mỗi chương chúng tôi cũng đưa ra một số bài tập nhằm cho sinh viên vận dụng lý thuyết đã học một cách thành thạo, và thấy được phần nào ứng dụng cụ thể của nó vào thực tiễn.

Phần 2 “**Thống kê toán học**” trình bày trong 4 chương

Chương 3. MẪU VÀ CÁCH BIỂU DIỄN MẪU

Chương 4. LÝ THUYẾT ƯỚC LƯỢNG

Chương 5. KIỂM ĐỊNH GIẢ THUYẾT THỐNG KÊ

Chương 6. TƯƠNG QUAN VÀ HỒI QUI

Đặc biệt cuối chương 1 phần 2 và cuối sách chúng tôi hướng dẫn cách sử dụng máy tính bỏ túi Casio fx 500MS trong việc tính toán một vài tham số trong xác suất thống kê phục vụ cho việc thi cử và nghiên cứu sau này khi chưa có đủ điều kiện.

Vì khả năng có hạn, nên cuốn sách khó tránh khỏi những sai sót, mong các bạn đọc và đồng nghiệp đóng góp để chúng tôi hoàn thiện hơn nữa.

Tác giả

Phần I SƠ LƯỢC LÝ THUYẾT XÁC SUẤT

Lý thuyết xác suất là một bộ môn Toán học nghiên cứu những quy luật ngẫu nhiên và những hiện tượng số lớn. Nó xác lập những quy luật tất nhiên ẩn dấu sau những hiện tượng mang tính ngẫu nhiên. Khi nghiên cứu một số lớn hiện tượng tương tự, việc nắm bắt những quy luật này sẽ cho phép dự báo các hiện tượng ngẫu nhiên đó sẽ xảy ra như thế nào. Các phương pháp của lý thuyết xác suất được ứng dụng rộng rãi trong việc giải quyết những bài toán thuộc các lĩnh vực khác nhau của khoa học Tự nhiên Kinh tế và Xã hội.

Chương 1 CÁC KHÁI NIỆM CƠ BẢN CỦA LÝ THUYẾT XÁC SUẤT

§1 PHÉP THỬ - SỰ KIỆN - XÁC SUẤT CỦA SỰ KIỆN

1.1. *Khái niệm về phép thử, sự kiện (biến cố) liên kết với phép thử*

Khi nghiên cứu một hiện tượng nào đó, người ta cần phải chuẩn bị một số điều kiện để tiến hành thí nghiệm, khi đó ta nói rằng người ta đã chuẩn bị một phép thử ngẫu nhiên. Vậy phép thử ngẫu nhiên là sự thực hiện một loạt điều kiện xác định nào đó với mục đích xác định trước, Người ta thường ký hiệu phép thử ngẫu nhiên là phép thử (G). Phép thử ngẫu nhiên (G) có thể là một thí nghiệm lặp lại trong các điều kiện bên ngoài giống hệt nhau. Chẳng hạn tung một đồng xu rơi xuống mặt bàn (phép thử (G)) quan sát xem mặt sấp lên trên hay mặt ngửa lên trên...

Một phép thử (G) sau khi thực hiện xong nó có nhiều kết cục có thể xảy ra, mỗi kết cục đó được gọi là một sự kiện sơ cấp của phép thử. Tập hợp các sự kiện sơ cấp của phép thử (G) gọi là không gian các sự kiện sơ cấp của (G), và kí hiệu là Ω , mỗi sự kiện sơ cấp của phép thử (G) xem như 1 điểm của không gian. Một tập hợp con của Ω được gọi là một sự kiện (hay biến cố) của phép thử (G). Ta gọi một sự kiện liên kết phép thử (G) là một sự kiện có thể xảy ra hoặc có thể không xảy ra tùy thuộc vào kết quả khi (G) thực hiện. Sự kiện như vậy gọi là sự kiện ngẫu nhiên.

Khi phép thử (G) thực hiện, một sự kiện nào đó được gọi là xảy ra khi và chỉ khi chỉ cần ít nhất một biến cố sơ cấp chứa trong sự kiện đó xảy ra là đủ.

Ví dụ. Phép thử (G) gieo một con xúc xắc xuống mặt bàn. Gọi e_i là kết quả chỉ mặt có i chấm lên trên ($i=1,2,3,4,5,6$) thì không gian các biến cố sơ cấp của phép thử là:

$$\Omega = \{e_1, e_2, e_3, e_4, e_5, e_6\}$$

Tập $A = \{e_3, e_6\}$ là biến cố chỉ xuất hiện mặt có chấm là bội của 3 lên trên sau khi gieo con xúc xắc. Biến cố A được gọi là xảy ra khi phép thử (G) tiến hành, nếu (G) thực hiện mặt 3 chấm lên trên hay mặt 6 chấm lên trên. Các sự kiện của phép thử thông thường chia làm 3 loại chính:

+ Sự kiện bất khả, kí hiệu là $\emptyset$ hoặc V là sự kiện mà khi phép thử thực hiện nhất thiết nó không xảy ra.

+ Sự kiện chắc chắn, kí hiệu là Ω là sự kiện mà khi phép thử thực hiện nhất thiết nó phải xảy ra.

+ Sự kiện ngẫu nhiên, kí hiệu bằng các chữ in hoa như A, B... là các sự kiện mà khi phép thử thực hiện nó có thể xảy ra cũng có thể không xảy ra.

1.2. *Quan hệ, các phép toán trên các sự kiện*

Người ta định nghĩa quan hệ giữa các sự kiện và các phép toán trên chúng cũng giống như các phép toán trên tập hợp, vì vậy mà sử dụng các phép toán như trong lý thuyết tập hợp.

1.2.1. Sự kiện kéo theo

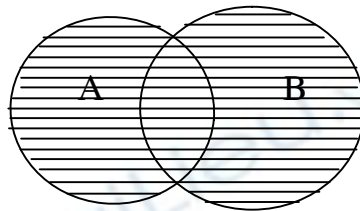
Sự kiện A gọi là kéo theo sự kiện B nếu A xảy ra thì B cũng xảy ra. Kí hiệu là $A \subset B$.

1.2.2. Sự kiện tương đương

Hai sự kiện A và B gọi là tương đương khi và chỉ khi $A \subset B$ và $B \subset A$.

1.2.3. Tổng các sự kiện

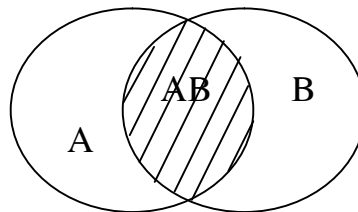
Sự kiện C được gọi là tổng các sự kiện A và B, ký hiệu $A+B = C$, hoặc $A \cup B = C$ khi và chỉ khi C xảy ra thì ít nhất 1 trong 2 sự kiện A hoặc B xảy ra.



Tổng quát. Cho n sự kiện $A_1, A_2, \dots, A_n$. Tổng của n sự kiện A_i là sự kiện C, kí hiệu $C = \bigcup_{i=1}^n A_i \Leftrightarrow C \text{ xảy ra thì } \exists A_i \text{ xảy ra } (i = 1; 2; \dots; n)$.

1.2.4. Tích các sự kiện

Tích của hai sự kiện A và B là sự kiện kí hiệu là AB hoặc $A \cap B$ thoả mãn: AB xảy ra $\Leftrightarrow$ cả A và B đồng thời xảy ra.

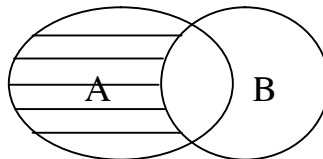


Tổng quát. Tích của n sự kiện $A_1, A_2, \dots, A_n$ là sự kiện kí hiệu $\bigcap_{i=1}^n A_i$ thoả mãn:

$\bigcap_{i=1}^n A_i$ xảy ra $\Leftrightarrow$ tất cả A_i đều xảy ra ($i = 1; 2; \dots; n$).

1.2.5. Hiệu của hai sự kiện

Sự kiện E được gọi là hiệu của hai sự kiện A và B, kí hiệu $E = A \setminus B$ nếu E xảy ra khi A xảy ra mà B không xảy ra.



1.2.6. Quan hệ giữa các sự kiện

i) Hai sự kiện A và B được gọi là xung khắc, nếu A xuất hiện thì B không xuất hiện và ngược lại. Nếu A, B là hai sự kiện xung khắc, ta kí hiệu $A \cap B = \emptyset$.

ii) Hai sự kiện A và B gọi là đối lập nếu $\begin{cases} A \cap B = \emptyset \\ A \cup B = \Omega \end{cases}$, khi đó sự kiện đối lập của A ký hiệu là $\bar{A}$.

iii) Hệ n sự kiện $A_1, A_2, \dots, A_n$ gọi là hệ sự kiện đầy đủ nếu:

$$\begin{cases} A_i \cap A_j = \emptyset, \forall i \neq j \\ \bigcup_{i=1}^n A_i = \Omega \end{cases}$$

Ví dụ 1. Phép thử (G) gieo một con xúc xắc, gọi e_i ($i = 1, 2, \dots, 6$) là sự kiện chỉ xuất hiện mặt i chấm lên trên sau khi gieo. A là sự kiện chỉ mặt có số chấm chẵn lên trên, B là sự kiện chỉ mặt có số chấm là bội của 3 lên trên, thì:

$$\Omega = \{e_1, e_2, \dots, e_6\} \text{ và } e_6 = A \cap B; \quad A = e_2 \cup e_4 \cup e_6.$$

Ví dụ 2. Hai xạ thủ cùng bắn mỗi người bắn một viên vào bia. Gọi $A_i :=$ “Người thứ i bắn trúng bia” ($i=1, 2$). Hãy viết các biến cố sau qua A_1, A_2 .

- Chỉ có xạ thủ thứ nhất bắn trúng bia: $A_1 \bar{A}_2$.
- Có đúng một xạ thủ bắn trúng bia: $A_1 \bar{A}_2 \cup \bar{A}_1 A_2$.
- Có ít nhất một xạ thủ bắn trúng bia: $A_1 \cup A_2$.
- Cả hai xạ thủ đều bắn trúng bia: $A_1 A_2$.
- Không có xạ thủ nào bắn trúng bia: $\bar{A}_1 \bar{A}_2$.
- Có không quá một xạ thủ bắn trúng bia: $\bar{A}_1 \bar{A}_2 \cup A_1 \bar{A}_2 \cup \bar{A}_1 A_2$.
- Chỉ ra một vài nhóm biến cố đầy đủ: $\{A_1, \bar{A}_1\}$ hoặc $\{A_2, \bar{A}_2\}$ hoặc

$$\{A_1 A_2, A_1 \bar{A}_2, \bar{A}_1 A_2, \bar{A}_1 \bar{A}_2\}.$$

§2 CÁC ĐỊNH NGHĨA CỦA XÁC SUẤT

Chúng ta thấy rằng, khi có phép thử ngẫu nhiên (G) được thực hiện thì các biến cố ngẫu nhiên $A, B, C, \dots$ liên kết với (G) có thể xảy ra hoặc không xảy ra. Do đó vấn đề đặt ra là: *Làm sao đo được mức độ xảy ra của một biến cố ngẫu nhiên nào đó?* Để giải quyết vấn đề này người ta tìm cách gán cho mỗi biến cố A liên kết với (G) một số ký hiệu $P(A)$ thỏa mãn 3 tính chất sau:

- $P(\Omega) = 1; \quad P(\emptyset) = 0.$
- $P(A) \in [0, 1].$
- Nếu A, B là hai biến cố xung khắc thì $P(A \cup B) = P(A) + P(B).$

Thì số $P(A)$ đó gọi là xác suất của biến cố A . Ba tính chất trên gọi là ba tính chất của xác suất.

Vậy xác suất của một biến cố là một số thực thuộc đoạn $[0, 1]$, chỉ mức độ xảy ra khách quan của biến cố (sự kiện) đó khi phép thử được tiến hành. Để đạt được mục đích đó chúng ta đưa ra định nghĩa xác suất trong một số trường hợp hay gặp sau đây

2.1. Định nghĩa cổ điển của xác suất

Xét phép thử (G) có số kết quả có thể xảy ra là n và các kết quả là đồng khả năng, trong n kết quả đó có m kết quả thuận lợi cho sự kiện A xảy ra thì xác suất của sự kiện A là số thực kí hiệu $P(A)$ và định nghĩa là $P(A) = \frac{m}{n}$.

Ví dụ 3. Gieo một con xúc xắc (Phép thử (G)) thì $\Omega = \{e_1, e_2, \dots, e_6\}$ do con xúc xắc cân đối và đồng chất nên các kết quả e_i ($i = 1, 2, 3, 4, 5, 6$) đồng khả năng xảy ra nên số khả năng của (G) là $n = 6$.

Gọi A là biến cố chỉ xuất hiện mặt có chấm là bội của 3 thì số khả năng thuận lợi cho A xảy ra là 2, vì nếu mặt 3 chấm xuất hiện hoặc mặt 6 chấm xuất hiện thì A xuất hiện vậy $m = 2$. Theo định nghĩa cổ điển của xác suất thì xác suất của biến cố A là:

$$P(A) = \frac{m}{n} = \frac{2}{6} = \frac{1}{3}$$

Ví dụ 4. Một thùng kín trong đó có 3 bi trắng và 4 bi đen, các bi làm đồng chất, cùng độ lớn độ nhẵn (gọi là đồng khả năng). Lấy ngẫu nhiên 3 bi cùng một lúc. Tìm xác suất để lấy được 2 bi đen và 1 bi trắng.

Phép thử (G) là lấy ngẫu nhiên một lúc 3 bi, do các bi đồng khả năng được lấy nên số cách lấy là $C_7^3 = \frac{7!}{3!(7-3)!} = \frac{7.6.5.4!}{3!4!} = 35$

$\Rightarrow$ Số đồng khả năng là $n = 35$.

Gọi A là biến cố lấy được 2 bi đen và một bi trắng, nên số cách lấy bi đen là C_4^2 , số cách lấy bi trắng là C_3^1 . Theo luật tích, số cách lấy cùng một lúc 3 bi được hai bi đen và 1 bi trắng là: $C_4^2 \times C_3^1 = 6 \times 3 = 18 \Rightarrow$ số khả năng thuận lợi cho A là $m = 18$.

Vậy theo định nghĩa cổ điển của xác suất ta có: $P(A) = \frac{18}{35}$.

2.2. Định nghĩa xác suất theo quan điểm thống kê

Xét một phép thử (G) liên kết với sự kiện A , lặp lại phép thử (G) n lần độc lập, Chúng ta thấy có k lần xuất hiện sự kiện A . Khi đó tỉ số được gọi là tần suất xuất hiện sự kiện A trong n lần lặp lại phép thử (G). Chúng ta nhận thấy rằng tần suất $f_n(A)$ có các tính chất sau:

- 1) $f_n(\Omega) = 1, f_n(V) = 0$
- 2) $0 \leq f_n(A) \leq 1$
- 3) A, B xung khắc thì $f_n(A \cup B) = f_n(A) + f_n(B)$

(Tự kiểm tra tại sao ?)

Và $f_n(A)$ thay đổi nếu n thay đổi hoặc thực hiện phép thử trong n lần khác. Tuy nhiên bằng thực nghiệm người ta chứng minh được rằng với n khá lớn thì $f_n(A)$ ổn định quanh một giá trị p nào đó, giá trị p đó theo quan điểm thống kê gọi là xác suất của sự kiện A .

Định nghĩa. Xác suất của sự kiện A là trị số ổn định của tần suất $f_n(A)$ khi số lượng phép thử tăng lên vô hạn.

Chẳng hạn hai nhà thống kê Buffon và Pearson đã thí nghiệm gieo đồng tiền nhiều lần, kết quả ở bảng sau:

Người gieo	Số lần gieo	Số lần sấp	Tần xuất
Buffon	4040	2048	0,5080
Pearson	12000	6019	0,5016
Pearson	24000	12012	0,5005

Qua kết quả trên cho chúng ta thấy tần suất xuất hiện mặt sấp (S) ổn định xung quanh giá trị $p = 0,5$ khi số lượng phép thử n tăng lên, nên ta nói rằng xác suất xuất hiện mặt sấp khi gieo đồng tiền là $P(S) = 0,5$.

2.3. Định nghĩa xác suất theo quan điểm hình học

Xét phép thử (G) lấy ngẫu nhiên một điểm trên đoạn $[0, 1]$, thì không gian các sự kiện sơ cấp của phép thử Ω ở đây là vô hạn kết quả không đếm được. Trong trường hợp này ta không thể xây dựng xác suất của sự kiện A trên cơ sở xác suất của các sự kiện sơ cấp P_i (vì các điểm trên đoạn thẳng coi như đồng khả năng và các $p_i = 0$).

Nhưng ta thấy rằng nếu sự kiện A là một đoạn thẳng nào đó nằm trong đoạn $[0, 1]$ thì A càng lớn xác suất để một điểm rơi vào trong A càng lớn, vì thế ta xem xác suất của một điểm rơi vào miền A là $P(A) = \frac{\text{Sé dui Säan } A}{\text{Sé dui Säan } [0,1]}$.

Dễ thấy rằng $P(A)$ có các tính chất của xác suất. Mở rộng kết quả trên cho trường hợp một điểm rơi vào miền phẳng hay khối không gian ta có định nghĩa như sau:

Định nghĩa xác suất theo quan điểm hình học

Giả sử Ω là tập hợp các điểm nào đó (đoạn thẳng, miền phẳng, mảnh mặt cong hay khối không gian), và A là tập con của Ω , khi đó xác suất để một điểm rơi vào miền A là: $P(A) = \frac{\text{Sé So cña } A}{\text{Sé So cña } \Omega}$, độ đo ở đây là độ dài, diện tích hay thể tích...

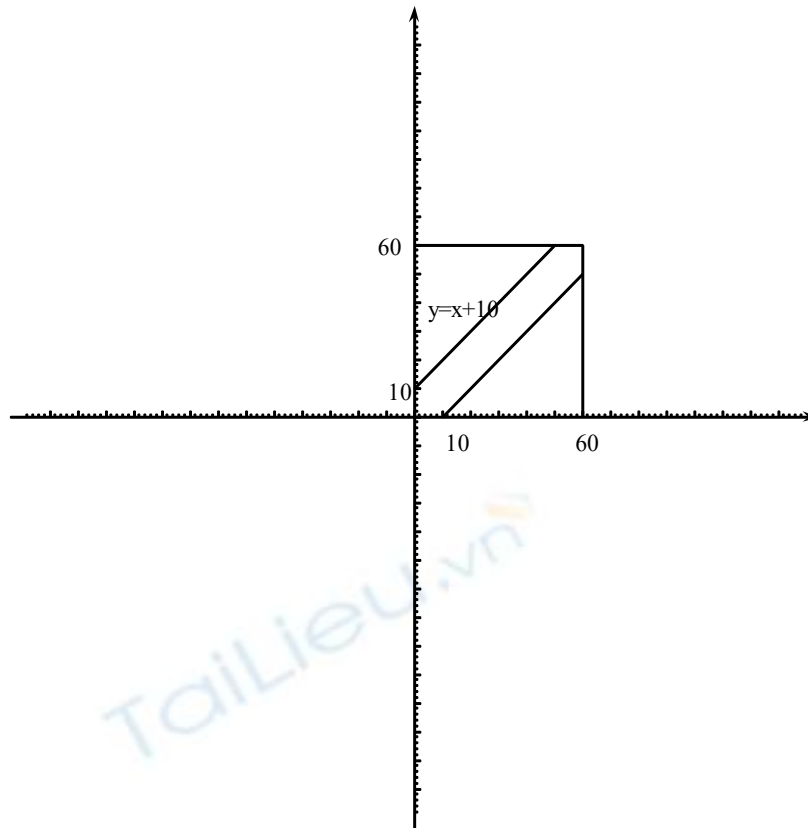
Ví dụ 5. Hai người hẹn gặp nhau tại một địa điểm đã định trong khoảng thời gian từ 19 đến 20 giờ. Hai người đến chỗ hẹn độc lập nhau và quy ước rằng khi đến chỗ hẹn sẽ đợi nhau 10 phút, nếu người kia không đến thì sẽ bỏ đi. Tính xác suất để họ gặp nhau.

Giải Ta biểu diễn thời điểm đến chỗ hẹn của người thứ nhất là một điểm trên trục hoành, người thứ hai trên trục tung. Như vậy thời điểm đến của cả hai người được biểu diễn bằng một điểm có tọa độ là cặp (x, y) nằm trong hình vuông $0 \leq x \leq 60; 0 \leq y \leq 60$, đơn vị tính là phút.

Để hai người gặp nhau các thời điểm đến x và y của mỗi người phải thỏa mãn bất đẳng thức $|x - y| \leq 10$, Hay $x - 10 \leq y \leq x + 10$.

Các điểm thỏa mãn bất đẳng thức trên được biểu diễn bởi các điểm nằm giữa hai đường thẳng $y = x - 10$ và $y = x + 10$ (Hình vẽ). Vậy theo định nghĩa xác suất Hình học ta có

$$P(A) = \frac{\text{Sé So cña } A}{\text{Sé So cña } \Omega} = \frac{60 \times 60 - 50 \times 50}{60 \times 60} = \frac{11}{36}$$



2.4. Sơ lược một số khái niệm của giải tích kết hợp

2.4.1. Chỉnh hợp

Cho một tập hợp X có n phần tử khác nhau. Một cách chọn ra k phần tử khác nhau có thứ tự từ n phần tử của tập hợp X gọi là một chỉnh hợp chập k của n phần tử ($k \leq n$). Số chỉnh hợp chập k của n phần tử kí hiệu và tính theo công thức:

$$A_n^k = n(n-1)(n-2)\dots(n-k+1)$$

Ví dụ 6. Cho $X = \{1, 2, 3, 4, 5\}$ gồm 5 chữ số 1; 2; 3; 4; 5. Hỏi có thể tạo nên bao nhiêu số gồm 3 chữ số đôi một khác nhau từ năm chữ số trên.

Một số như đề ra là một bộ có thứ tự gồm 3 chữ số đôi một khác nhau lấy từ 5 chữ số đã cho. Do đó số có thể tạo thành là: $A_5^3 = 5 \times 4 \times 3 = 60$ số

2.4.2. Hoán vị

Một hoán vị của n phần tử của tập hợp X gồm n phần tử khác nhau là một chỉnh hợp chập n của n . Kí hiệu số hoán vị của n phần tử là: $P_n = A_n^n = n!$

2.4.3. Tổ hợp

Một tổ hợp chập k của n phần tử của tập hợp X gồm n phần tử khác nhau là một cách chọn ra k phần tử khác nhau của X không phân biệt thứ tự.

Số tổ hợp chập k của n phần tử kí hiệu và tính theo công thức sau:

$$C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!}$$

Người ta chứng minh được rằng: $C_{n-1}^{k-1} + C_{n-1}^k = C_n^k$

Qui ước $C_n^0 = 1$ ta có công thức khai triển nhị thức như sau:

$$(a+b)^n = \sum_{k=0}^n C_n^k a^{n-k} b^k$$

Ví dụ 7. Một nhóm học viên có 5 người, trong đó có 3 nam và 2 nữ. Muốn chọn 3 học viên đi lao động trong đó có 2 nam và 1 nữ. Hỏi có bao nhiêu cách chọn.

Số cách chọn 2 nam trong 3 nam là: $C_3^2 = 3$

Số cách chọn 1 nữ trong 2 nữ là: $C_2^1 = 2$

$\Rightarrow$ Số cách chọn 3 người có 2 nam và 1 nữ là: $C_3^2 \times C_2^1 = 6$

2.4.4. Luật tích

Giả sử để thực hiện việc A ta phải thực hiện liên tiếp k bước:

Bước thứ 1: có m_1 cách thực hiện.

Bước thứ 2: có m_2 cách thực hiện.

.....

Bước thứ k : có m_k cách thực hiện.

Khi đó số cách thực hiện việc A là $m = m_1 \times m_2 \times \dots \times m_k$.

§3 CÁC ĐỊNH LÝ CƠ BẢN CỦA XÁC SUẤT

3.1. Định lý cộng xác suất

Nếu A, B là 2 biến cố xung khắc thì $p(A \cup B) = p(A) + p(B)$.

Nếu A, B là 2 biến cố bất kì thì $p(A \cup B) = p(A) + p(B) - p(AB)$

Tổng quát: Cho n biến cố $A_1, A_2, \dots, A_n$

Nếu n biến cố $A_1, A_2, \dots, A_n$ này xung khắc từng đôi thì ta có:

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

Nếu n biến cố $A_1, A_2, \dots, A_n$ bất kì thì ta có công thức:

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_i A_j) + \sum_{i < j < k} P(A_i A_j A_k) - \dots + (-1)^{n-1} P\left(\bigcap_{i=1}^n A_i\right)$$

Chúng minh công thức trên bằng phương pháp qui nạp.

3.2. Xác suất có điều kiện, định lý nhân xác suất

3.2.1. Định nghĩa

Nếu B là một biến cố có xác suất $P(B) \neq 0$ thì xác suất có điều kiện của biến cố A với điều kiện của biến cố B đã xảy ra được định nghĩa là:

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

3.2.2. Định lý nhân xác suất

Nếu $P(A) > 0$, $P(B) > 0$ với A, B là 2 biến cố bất kì thì:

$$P(AB) = P(B) \times P\left(\frac{A}{B}\right) = P(A) \times P\left(\frac{B}{A}\right)$$

Ví dụ 8. Trong kho có 96% sản phẩm đúng qui cách. Trong số sản phẩm đúng qui cách có 70% sản phẩm loại I. Lấy ngẫu nhiên một sản phẩm, tính xác suất để lấy được sản phẩm loại I.

Gọi A là biến cố lấy ngẫu nhiên một sản phẩm thì được sản phẩm loại I, B là biến cố chỉ lấy ngẫu nhiên một sản phẩm thì được sản phẩm đúng qui cách, thì ta có $P(B) = 96\%$.

$$\text{Cần tính } P(AB) = P(B) \times P\left(\frac{A}{B}\right) = 0,96 \times 0,70 = 0,672$$

3.2.3. Sự độc lập các biến cố

Nếu việc xảy ra hay không xảy ra của biến cố B không ảnh hưởng gì đến việc xảy ra biến cố A , lẽ dĩ nhiên A, B là 2 biến cố độc lập và ta viết:

$$p\left(\frac{A}{B}\right) = p\left(\frac{A}{\bar{B}}\right) = p(A)$$

Từ đó suy ra: A, B là hai biến cố độc lập $\Leftrightarrow p(AB) = p(A) \times p(B)$

Nhận xét. Nếu A, B là 2 biến cố độc lập thì A và $\bar{B}$; $\bar{A}$ và B ; $\bar{A}$ và $\bar{B}$ cũng độc lập.

Tổng quát: Cho n biến cố $A_1, A_2, \dots, A_n$ gọi là độc lập nếu mỗi biến cố trong chúng độc lập với tích một số các biến cố còn lại. Từ đó tổng quát định lý nhân cho n biến cố

$$A_1, A_2, \dots, A_n \text{ là: } P\left(\bigcap_{i=1}^n A_i\right) = P(A_1) \cdot P\left(\frac{A_2}{A_1}\right) \cdot P\left(\frac{A_3}{A_1 A_2}\right) \dots P\left(\frac{A_n}{\bigcap_{i=1}^{n-1} A_i}\right)$$

Chúng minh công thức này bằng qui nạp (độc giả tự chứng minh xem như bài tập).

3.3. Công thức xác suất toàn phần, công thức Bayes

Giả sử $A_1, A_2, \dots, A_n$ là một hệ sự kiện đầy đủ, B là một sự kiện bất kì thì:

$$B = B \cap \left(\bigcup_{i=1}^n A_i\right) = BA_1 + BA_2 + \dots + BA_n$$

Do các $A_i, (i=1, 2, \dots, n)$ xung khắc từng đôi nên các BA_i cũng xung khắc từng đôi. Áp dụng công thức cộng xác suất ta có:

$$P(B) = P(BA_1) + P(BA_2) + \dots + P(BA_n).$$

Áp dụng công thức nhân ta có:

$$P(B) = P(A_1)P\left(\frac{B}{A_1}\right) + P(A_2)P\left(\frac{B}{A_2}\right) + \dots + P(A_n)P\left(\frac{B}{A_n}\right)$$

$$\text{hay : } P(B) = \sum_{i=1}^n P(A_i)P\left(\frac{B}{A_i}\right) \quad (1)$$

Công thức (1) gọi là công thức xác suất toàn phần.

Bây giờ nếu phép thử đã thực hiện, biết sự kiện B đã xảy ra, tìm xác suất xuất hiện sự kiện $A_i, (i=1, 2, \dots, n)$, tức là cần tìm các xác suất $P\left(\frac{A_i}{B}\right), (i=1, 2, \dots, n)$?

Theo Định lý nhân xác suất thì:

$$P(A_i B) = P(A_i) \cdot P\left(\frac{B}{A_i}\right) = P(B) \cdot P\left(\frac{A_i}{B}\right) \Rightarrow P\left(\frac{A_i}{B}\right) = \frac{P(A_i) \cdot P\left(\frac{B}{A_i}\right)}{P(B)} \Rightarrow$$

$$\Rightarrow P\left(\frac{A_i}{B}\right) = \frac{P(A_i) \cdot P\left(\frac{B}{A_i}\right)}{\sum_{i=1}^n P(A_i) \cdot P\left(\frac{B}{A_i}\right)} \quad (2) \quad i=1,2,\dots,n$$

(2) gọi là công thức Bayes.

Ý nghĩa của công thức (2) là: Sau khi thí nghiệm biết sự kiện B đã xảy ra, ta tính xác suất $P\left(\frac{A_i}{B}\right)$ là xác suất B xảy ra do yếu tố A_i tác động là bao nhiêu? Vì thế $P\left(\frac{A_i}{B}\right)$ gọi là xác suất *hậu nghiệm* để phân biệt với các xác suất *tiền nghiệm* $P(A_i)$.

Ví dụ 9. Một trạm cấp cứu bỏng có 80% bệnh nhân bỏng do nóng và 20% bệnh nhân bỏng do hoá chất. Loại bỏng do nóng có 30% bị biến chứng, loại bỏng do hoá chất có 50% bị biến chứng.

a) Tìm xác suất để khi mở tập hồ sơ ra lấy ngẫu nhiên 1 bệnh án thì gặp bệnh án của bệnh nhân bị biến chứng?

b) Tìm xác suất để khi mở tập hồ sơ ra lấy ngẫu nhiên 1 bệnh án thì gặp bệnh án của bệnh nhân bị biến chứng do nóng gây ra?

Gọi A_1 là sự kiện lấy ngẫu nhiên một bệnh án thì gặp bệnh án của bệnh nhân bị bỏng do nóng. A_2 là biến cố lấy ngẫu nhiên một bệnh án thì gặp bệnh án của bệnh nhân bị bỏng do hoá chất, B là biến cố lấy ngẫu nhiên một bệnh án thì gặp bệnh án của bệnh nhân bị biến chứng. Khi đó $\{A_1, A_2\}$ lập thành hệ sự kiện đầy đủ và $B = BA_1 + BA_2$.

$$a) P(B) = P(A_1)P\left(\frac{B}{A_1}\right) + P(A_2)P\left(\frac{B}{A_2}\right) = \frac{80}{100} \times \frac{30}{100} + \frac{20}{100} \times \frac{50}{100} = 0,34$$

b) Theo công thức (2) cần tính:

$$P\left(\frac{A_1}{B}\right) = \frac{P(A_1) \cdot P\left(\frac{B}{A_1}\right)}{P(B)} = \frac{\frac{80}{100} \times \frac{30}{100}}{\frac{34}{100}} = 0,706.$$

3.4. Dãy phép thử độc lập, công thức Bernoulli

Xét một phép thử (G) , A là một biến cố liên kết với (G) $P(A) = p$ và $P(\bar{A}) = 1 - p = q$, một phép thử như vậy được gọi là phép thử Bernoulli.

Lặp lại (G) n lần độc lập. Tìm xác suất biến cố B chỉ sự kiện A xuất hiện đúng k ($k = 0, 1, 2, \dots, n$) lần, kí hiệu xác suất này là $P_n(k)$. Lặp lại (G) n lần như trên gọi là thực hiện dãy phép thử độc lập Bernoulli.

Gọi B là biến cố trong n lần lặp lại (G) , sự kiện A xuất hiện k lần. Ta thấy rằng B có nhiều cách thực hiện:

Chẳng hạn: $B = \underbrace{AA \dots A}_k + \underbrace{AA \dots A}_{n-k} + \underbrace{AA \dots A}_{k-1} + \underbrace{AA \dots A}_{n-k-2} + \dots$

Mỗi biến cố trong tổng ứng với một cách chọn k chữ số A trong n chữ $A, \bar{A}$ nên số số hạng của tổng là C_n^k , vì thế có C_n^k cách thực hiện B mà mỗi cách có

$$P\left(\underbrace{AA \dots A}_k + \underbrace{AA \dots A}_{n-k} + \dots\right) = p^k q^{n-k}.$$

$$\Rightarrow P_k(n) = P(B) = C_n^k p^k q^{n-k}. \quad (3)$$

Công thức (3) gọi là công thức Bernoulli thứ nhất.

Để tính xác suất trong n phép thử Bernoulli độc lập, biến cố A xuất hiện từ k_1 đến k_2 lần ($0 \leq k_1 < k_2 \leq n$) ta dùng Định lý cộng xác suất, và ký hiệu xác suất này là $P_n(k_1, k_2)$, ta có công thức:

$$P_n(k_1, k_2) = \sum_{k=k_1}^{k_2} P_n(k) = \sum_{k=k_1}^{k_2} C_n^k p^k q^{n-k} \quad (4)$$

Công thức (4) gọi là công thức Bernoulli thứ 2.

Chú ý. Do $(p+q)^n = \sum_{k=0}^n C_n^k p^k q^{n-k}$ (công thức khai triển nhị thức) $\Rightarrow P_n(k)$ chính là hệ số p^k của khai triển nhị thức trên, vì vậy công thức (3) còn được gọi là công thức xác suất nhị thức.

3.5. Số có khả năng nhất

Nếu trong dãy phép thử độc lập Bernoulli tồn tại số k_0 sao cho $P_n(k_0) \geq P_n(k)$, $\forall k=1, 2, \dots, n$, thì số k_0 đó được gọi là số có khả năng nhất của phép thử. Để tìm số k_0 ta tính: $(n+1)p$.

Nếu $(n+1)p \in N$ thì có 2 giá trị k_0 là $k_0 = (n+1)p$ và $k_0 = (n+1)p - 1$.

Nếu $(n+1)p \notin N$ thì k_0 là số nguyên lớn nhất không vượt quá $(n+1)p$.

Ví dụ 10. Khi lại chuột trắng và chuột xám ở dòng thuần thì F_1 mọi con chuột đều xám

(xám là tính trội) ở thế hệ F_2 có $\frac{3}{4}$ số chuột xám và $\frac{1}{4}$ là trắng. Giả sử F_2 cho ta được 5 con chuột. Tính xác suất sao cho:

- Có 3 con xám và hai con trắng.
- Có ít nhất một con trắng.

Gọi A là biến cố xuất hiện một con chuột ở F_2 là xám $\Rightarrow P(A) = \frac{3}{4}, P(\bar{A}) = \frac{1}{4}$ ta

cần tính các xác suất sau:

$$a) P_5(3) = C_5^3 \left(\frac{3}{4}\right)^3 \left(\frac{1}{4}\right)^2 = \frac{135}{512} = 0,264$$

$$b) 1 - p_5(0) = 1 - \frac{243}{1024} = 0,763$$

3.6. Xác suất trong chẩn đoán

Công việc hàng ngày của BS là chẩn đoán. Chẩn đoán tức là xác suất. Ta theo dõi tiến trình chẩn đoán xem xác suất biểu hiện ở những vị trí nào.

Một người đến khám bệnh vì triệu chứng X, với triệu chứng này BS nghi ngờ và "đoán" bị bệnh B. Câu hỏi xuất hiện là khả năng bị bệnh B là bao nhiêu? Khả năng này gọi là chỉ số nghi ngờ.

Qua phỏng vấn người bệnh hoặc người nhà người bệnh, hoặc làm xét nghiệm (nói chung làm thủ tục chẩn đoán T), kết quả T có thể là dương tính T^+ hoặc âm tính T^- liên quan đến bệnh B.

Câu hỏi: Khi có kết quả T chỉ số nghi ngờ có thay đổi không? Khi nào thì cho làm xét nghiệm T? Liệu kết quả T trả về có đáng tin cậy không? Nội dung phần này sẽ trả lời những câu hỏi đó.

3.6.1. Độ chính xác của một xét nghiệm

Có hai tham số đánh giá một xét nghiệm đó là độ nhạy (Sensitivity) và độ chuyên (Speccificity)

3.6.1.1. Độ nhạy (Sensitivity) là khả năng xét nghiệm T báo dương tính (T^+) đối với người bệnh B, ký hiệu: $P(T^+/B^+)$ và gọi là dương thật (true positive).

Âm giả là tỷ lệ xét nghiệm T báo âm tính (T^-) đối với người bị bệnh B, ký hiệu: $P(T^-/B^+)$. Như vậy ta có $P(T^+/B^+) + P(T^-/B^+) = 1$

3.6.1.2. Độ chuyên (Speccificity) là tỷ lệ xét nghiệm T báo âm tính (T^-) trên người không bị bệnh B và ký hiệu: $P(T^-/B^-)$, còn gọi là âm thật.

Dương giả là tỷ lệ xét nghiệm T báo dương tính (T^+) trên người không bị bệnh B và ký hiệu: $P(T^+/B^-)$, tương tự ta có $P(T^-/B^-) + P(T^+/B^-) = 1$.

3.6.1.3. Xác định độ nhạy và độ chuyên. Chọn hai nhóm: nhóm bị bệnh B ký hiệu B^+ , nhóm không bị bệnh B ký hiệu B^- . Cả hai nhóm cho làm xét nghiệm T. Tùy thuộc vào kết quả xét nghiệm T trên hai nhóm mà ta có độ nhạy và độ chuyên của xét nghiệm T như sau:

		Bệnh B		
		B^+	B^-	
Xét nghiệm T	T^+	a	b	a+b
	T^-	c	d	c+d
		a+c	b+d	

Độ nhạy: $P(T^+/B^+) = \frac{a}{a+c}$; Độ chuyên: $P(T^-/B^-) = \frac{d}{b+d}$

dương giả: $P(T^+/B^-) = \frac{b}{b+d}$; âm giả: $P(T^-/B^+) = \frac{c}{a+c}$

Nhận xét: Độ nhạy và độ chuyên của xét nghiệm không phụ thuộc vào tỷ lệ bệnh đang lưu hành.

Ví dụ. Có hai xét nghiệm T_1 , T_2 : T_1 có độ nhạy 93% và độ chuyên 95%. T_2 dương giả 7%; âm giả 5%. T_1 dùng sàng lọc người có nguy cơ bị bệnh B, T_2 dùng chẩn đoán bệnh này trên những người mà T_1 cho kết quả dương tính.

Một người từ dân số có tỷ lệ bệnh B là 0,001, cho người này làm xét nghiệm T_1 kết quả dương tính, cho làm xét nghiệm T_2 cũng dương tính. Tính khả năng người này bị bệnh B.

Giải: Từ gt có: $P(T_1^+/B^+) = 0,93$; $P(T_1^-/B^-) = 0,95$

$$P(T_2^+/B^-) = 0,07; \quad P(T_2^+/B^+) = 0,05.$$

$$P(T_1^+) = P(B^+)P(T_1^+/B^+) + P(B^-)P(T_1^+/B^-) = 0,001 \times 0,93 + (1-0,001) \times (1-0,95) = 0,05088.$$

$$P(B^+/T_1^+) = \frac{P(B^+)P(T_1^+/B^+)}{P(T_1^+)} = \frac{0,001 \times 0,93}{0,05088} = 0,0183.$$

Cho làm xét nghiệm T_2 , kết quả T_2 dương tính, vậy ta có

$$P(T_2^+) = P(B^+)P(T_2^+/B^+) + P(B^-)P(T_2^+/B^-) =$$

$$0,0183 \times (1-0,05) + 0,07 \times (1-0,183) = 0,0861.$$

$$P(B^+/T_2^+) = \frac{P(B^+)P(T_2^+/B^+)}{P(T_2^+)} = \frac{(1-0,05) \times 0,018}{0,0861} = 0,2017.$$

Kết luận người này mắc bệnh B là 20,17%.

3.6.2. Giá trị tiên đoán (predictive value)

+ $P(T^-/B^-)$ là khả năng bị bệnh B nếu xét nghiệm T dương tính. Ký hiệu PV^+ .

+ $P(T^+/B^-)$ là khả năng bị bệnh B nếu xét nghiệm T âm tính. Ký hiệu PV^- .

+ $LR^- = \frac{P(T^+/B^+)}{1 - P(T^+/B^-)}$ càng lớn hơn 1 thì xét nghiệm càng nhạy.

+ $LR^+ = \frac{1 - P(T^+/B^+)}{1 - P(T^+/B^-)}$ càng bé hơn 1 thì xét nghiệm càng chuyên. Đây là hai thông số

của một xét nghiệm.

+ Tỷ cơ hội của một biến cố A là số ký hiệu: $LR(A)$ là tỷ số giữa xác suất xảy ra biến cố A đối với người bị bệnh B so với xác suất xảy ra biến cố A đối với người không bị bệnh B. Do đó khi một bệnh nhân có biến cố A xảy ra mà tỷ cơ hội $LR(A)$ càng lớn hơn 1 thì sự nghi ngờ người đó bị bệnh B càng cao.

BÀI TẬP CHƯƠNG 1

1. Một lô hàng có 100 sản phẩm, trong đó có 15 phế phẩm. Lấy ngẫu nhiên ra 20 sản phẩm. Tìm xác suất để cho trong 20 sản phẩm lấy ra có :
 - a. 5 phế phẩm.
 - b. Bị cả 15 phế phẩm.
 - c. Có đúng 10 chính phẩm.
2. Một em bé có 8 bìa hình vuông với các chữ N, N, N, A, A, H, H, O. Tìm xác suất để em bé trong khi sắp ngẫu nhiên theo hàng ngang, em thu được từ NHANHNAO.
3. Một em bé có một hộp chứa 2 bi trắng và 4 bi đỏ. Em rút hủi hoạ từng viên bi một cho đến viên cuối cùng. Tìm xác suất để viên bi cuối cùng là đỏ. Nếu chọn ngẫu nhiên một bi, tìm xác suất nhận được bi đỏ.
4. Một người có một hộp kín, trong đó có chứa 3 viên bi không rõ màu sắc, nhưng chúng đồng khả năng được chọn. Người đó bỏ thêm vào hộp 1 bi đỏ đồng khả năng được lấy với các bi trong hộp. Xáo đều rồi lấy ngẫu nhiên từ hộp ra một bi. Tìm xác suất để lấy được bi đỏ.
5. Hai xạ thủ bắn vào một bia một cách độc lập nhau, mỗi người bắn bốn lần mỗi lần một phát. Xác suất bắn trúng đích của người thứ nhất và người thứ hai trong mỗi lần bắn lần lượt là 0,6 và 0,9. Mỗi người sẽ được coi là bắn đạt yêu cầu nếu bắn trúng từ ba phát trở lên.
 - a) Hãy tính xác suất bắn đạt yêu cầu của từng người.
 - b) Kết quả bắn kiểm tra cho biết trong hai xạ thủ trên chỉ có một xạ thủ bắn đạt yêu cầu. Tìm xác suất để người bắn không đạt yêu cầu là người thứ nhất.
6. Biết rằng tỷ lệ người mắc bệnh nào đó ở một địa phương nào đó là 3%. Người ta sử dụng một phản ứng mà nếu người bị bệnh thì phản ứng luôn luôn dương tính, nếu không bị bệnh thì phản ứng dương tính với xác suất 0,20.
 - a. Tìm xác suất phản ứng dương tính.
 - b. Tìm xác suất bị bệnh, không bị bệnh trong nhóm người có phản ứng dương tính.
 - c. Qua phương pháp thử này có thể ước lượng tỷ lệ mắc bệnh là bao nhiêu.
7. Một bác sỹ có tiếng về chữa một bệnh nào đó. Xác suất chữa khỏi bệnh là 0,80. Có người nói rằng cứ 10 người đến chữa thì chắc chắn có 8 người khỏi. Điều khẳng định đó có đúng không?

Tìm xác suất sao cho bác sỹ đó chữa 10 người thì có 8 người khỏi.
8. Trong một vùng có tỷ lệ nữ là 55%. Trong đợt dịch bệnh, khả năng mắc bệnh của nam là 8%, của nữ là 3%. Hỏi tỷ lệ mắc bệnh chung của cả vùng ?
9. Có một bệnh nhân mà bác sỹ chẩn đoán là mắc bệnh A với xác suất 0,5; mắc bệnh B với xác suất 0,35 và mắc bệnh C với xác suất là 0,15. Để có thêm thông tin chẩn đoán bác sỹ đã cho xét nghiệm sinh hoá. Sau 3 lần thử thấy có 2 lần dương tính. Hãy cho biết nên chẩn đoán bệnh nhân mắc bệnh nào? Biết rằng khả năng dương tính của mỗi lần xét nghiệm với bệnh A, B, C tương ứng là 0,1; 0,2; 0,6.
10. Một bình chứa 12 bi, trong đó có 4 bi trắng. Một bình khác chứa 20 bi, trong đó có 14 bi trắng. Ta làm thí nghiệm như sau:
 - Bước 1: lấy ngẫu nhiên từ mỗi bình ra một bi;
 - Bước 2: sau đó lấy ngẫu nhiên một bi trong hai bi vừa lấy được.

Tính xác suất để trong hai bi lấy ra được ở bước 1 có đúng một bi trắng, biết rằng bi lấy ra được ở bước 2 là bi trắng.

11. Cho n cái hộp, mỗi hộp chứa m bi trắng và k bi đỏ. Lấy hũ họa 1 bi từ hộp 1 bỏ vào hộp 2 sau đó lấy hũ họa 1 bi từ hộp 2 bỏ vào hộp 3, cứ tiếp tục lấy hũ họa 1 bi từ hộp 3 bỏ vào hộp 4,Tìm xác suất để viên bi cuối cùng lấy ra từ hộp n là trắng.

12. Một học sinh viết xong n bức thư rồi bỏ vào n bì thư, dán lại và trên mỗi bì thư ghi một địa chỉ khác nhau cần gửi, rồi gửi đi. Tìm xác suất sao cho có ít nhất một lá thư đến đúng địa chỉ. Gọi xác suất đó là P_n . Tìm $\lim_{n \rightarrow \infty} P_n$.

13. Gieo hũ họa một điểm lên một đoạn thẳng có độ dài 30cm. Tìm xác suất để điểm đó rơi vào một đoạn con có độ dài 10 cm hoàn toàn nằm trong đoạn đã cho.

14. Cho đoạn thẳng với độ dài a . Chọn ngẫu nhiên trên đoạn này hai điểm, khi đó ta có đoạn thẳng được chia làm ba đoạn nhỏ. Tìm xác suất để ba đoạn thu được lập thành các cạnh của một tam giác.

15. Cho hình vuông với các đỉnh $A(0;0)$, $B(0; 1)$, $C(1; 0)$, $D(1; 1)$. Gieo ngẫu nhiên một điểm $M(X, Y)$ trong hình vuông đó.

a. Chứng minh rằng: $P \{X < x, Y < y\} = P \{X < x\}P \{Y < y\} = xy$.

b. Tìm $P \{|X - Y| < z\}$ với $1 \leq z \leq 1$.

c. Tìm $P \{XY < z\}$ với $1 \leq z \leq 1$.

d. Tìm $P \{\max(X, Y) < z\}$ với $1 \leq z \leq 1$.

16. Trong cơ quan nọ có 3 chiếc ô tô. Khả năng có sự cố của mỗi ô tô tương ứng là 0,15; 0,20 và 0,10.

a. Tìm xác suất cả ba ô tô cùng bị hỏng.

b. Tìm xác suất có ít nhất một cái hoạt động được.

c. Tìm khả năng cả ba ô tô cùng hoạt động được.

d. Tìm khả năng có không quá 2 ô tô bị hỏng.

17. Một người có ba chỗ ưa thích như nhau để câu cá. Xác suất câu được cá ở những chỗ đó tương ứng là 0,6; 0,7 và 0,8. Biết rằng ở mỗi một chỗ người đó đã thả câu ba lần và chỉ câu được một con cá. Tìm xác suất để cá câu được ở chỗ thứ nhất.

18. Tín hiệu thông tin được phát ba lần với xác suất thu được mỗi lần là 0,4.

a. Tìm xác suất để nguồn thu nhận được thông tin đó.

b. Nếu muốn xác suất thu được thông tin lên 0,9 thì phải phát bao nhiêu lần.

Chương 2 BIẾN NGẪU NHIÊN VÀ HÀM PHÂN PHỐI

§1 BIẾN NGẪU NHIÊN

1.1. Khái niệm biến ngẫu nhiên

Biến ngẫu nhiên là đại lượng nhận giá trị thực với xác suất xác định. Người ta kí hiệu các biến ngẫu nhiên bằng các chữ in hoa như: X, Y, Z hoặc các chữ Hy Lạp như: $\xi, \eta, \dots$ và các giá trị nhận được của nó bằng các chữ thường như: $x_1, x_2, \dots, y_1, y_2, \dots$

Biến ngẫu nhiên có hai loại: Biến ngẫu nhiên rời rạc và Biến ngẫu nhiên liên tục.

Biến ngẫu nhiên được gọi là rời rạc nếu các giá trị có thể nhận được của nó hữu hạn hoặc vô hạn đếm được. Biến ngẫu nhiên liên tục là các biến ngẫu nhiên mà các giá trị nhận được của chúng là vô hạn không đếm được, có thể lấp đầy một khoảng (a, b) của một trục số thực.

Ví dụ 1. Phép thử (G) bắn liên tiếp 4 phát súng vào 1 mục tiêu. Gọi X là đại lượng chỉ số phát đạn trúng mục tiêu thì X là biến ngẫu nhiên rời rạc, các giá trị nhận được của chúng là: 0, 1, 2, 3, 4.

Ví dụ 2. Đường dây điện nối 2 tổng đài A, B cách nhau 50 mét bỗng nhiên bị đứt. Gọi X là đại lượng chỉ khoảng cách từ điểm đứt đến A (đơn vị đo là mét), thì X là biến ngẫu nhiên liên tục (tính chính xác đến 0,001 mét).

1.2. Dãy phân phối xác suất của biến ngẫu nhiên rời rạc

Cho biến ngẫu nhiên rời rạc X có thể nhận n giá trị $x_1, x_2, \dots, x_n$ với xác suất $P(X = x_i) = p_i$ ($i = 1, 2, \dots, n$). Dãy phân phối xác suất của X (hay qui luật phân phối của X) là:

X	x_1	x_2		x_n
P	p_1	p_2		p_n

Trong đó $\sum_{i=1}^n p_i = 1$

Ví dụ 3. X là biến ngẫu nhiên chỉ số chấm của các mặt trên con xúc xắc, khi gieo 1 con xúc xắc thì dãy phân phối của X là:

X	1	2	3	4	5	6
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

1.3. Hàm phân phối

Định nghĩa. Hàm phân phối của biến ngẫu nhiên X là hàm $F(x)$ được xác định theo công thức công thức: $F(x) = P(X < x)$.

$\Rightarrow$ Nếu X là biến ngẫu nhiên rời rạc có dãy phân phối xác suất:

X	x_1	x_2		x_n
P	p_1	p_2		p_n

Thì: $F(x) = P(X < x) = \sum_{x_i < x} P(X = x_i)$

Ví dụ 4. Lập hàm phân phối của biến ngẫu nhiên rời rạc X có dãy phân phối xác suất:

X	-1	1	2	3
p	0,1	0,3	0,4	0,2

Ta có hàm phân phối của X là:

$$F(x) = P(X < x) = \sum_{x_i < x} p_i$$

Cụ thể: $x \leq -1$ thì $F(x) = P(X < x) = 0$

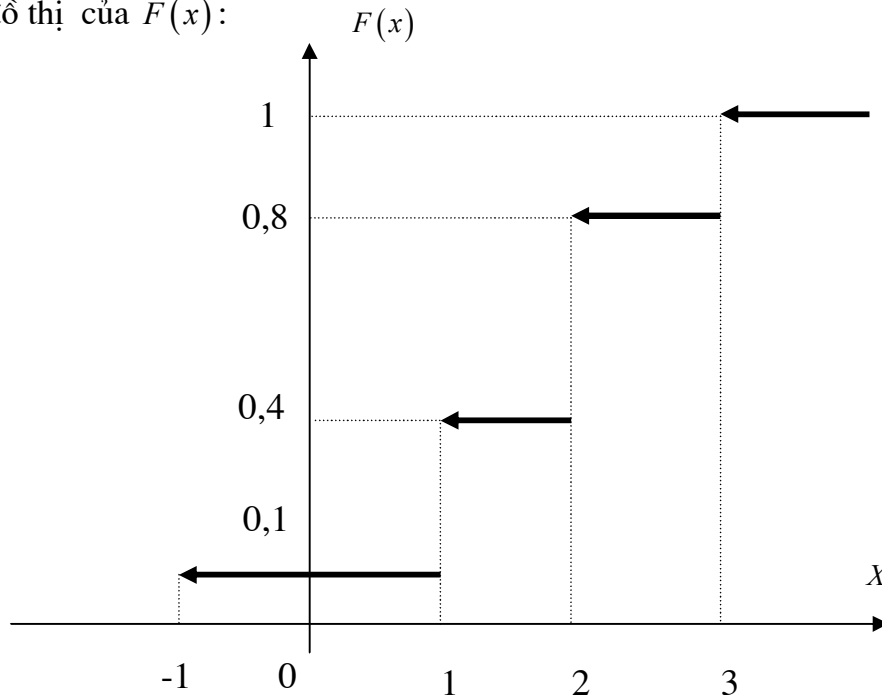
Nếu $-1 < x \leq 1$ thì $F(x) = P(X < x) = P(X = -1) = 0,1$. Tương tự

$1 < x \leq 2$ thì $F(x) = 0,1 + 0,3 = 0,4$.

$2 < x \leq 3$ thì $F(x) = 0,1 + 0,3 + 0,4 = 0,8$; $x > 3$ thì $F(x) = 1$.

Vậy ta có $F(x) = \begin{cases} 0, & \text{Nếu } x \leq -1 \\ 0,1 & \text{Nếu } -1 < x \leq 1 \\ 0,4 & \text{Nếu } 1 < x \leq 2 \\ 0,8 & \text{Nếu } 2 < x \leq 3 \\ 1 & \text{Nếu } x > 3 \end{cases}$

Nên đồ thị của $F(x)$:



Đồ thị $F(x)$ có dạng hình bậc thang.

Các tính chất của hàm phân phối $F(x)$

i) $0 \leq F(x) \leq 1, \forall x$

$$\text{ii)} \quad P(\alpha \leq X < \beta) = F(\beta) - F(\alpha)$$

$$\text{iii)} \quad \alpha < \beta \Rightarrow F(\alpha) \leq F(\beta), \text{ tức là hàm phân phối không giảm.}$$

$$\text{iv)} \quad F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0, \quad F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1.$$

Chứng minh. Các tính chất i), iv) suy trực tiếp từ định nghĩa. Chúng ta chứng minh các tính chất ii), iii).

Gọi A là biến cố chỉ $\{X < \alpha\}$, B là biến cố chỉ $\{\alpha \leq X < \beta\}$, C là biến cố chỉ $\{X < \beta\} \Rightarrow C = A + B$ và $AB = \emptyset$, áp dụng công thức cộng xác suất có:

$$P(C) = P(A) + P(B).$$

$$F(\beta) = F(\alpha) + P(\alpha \leq X < \beta) \Rightarrow P(\alpha \leq X < \beta) = F(\beta) - F(\alpha) \Rightarrow \text{ii}).$$

Từ $F(\beta) = F(\alpha) + P(\alpha \leq X < \beta)$, $(\beta > \alpha) \Rightarrow F(\beta) > F(\alpha)$ vì $P(\alpha \leq X < \beta) \geq 0 \Rightarrow \text{iii}).$

Từ tính chất iii) $\Rightarrow$ Nếu $F(x)$ liên tục thì:

$$\lim_{\beta \rightarrow \alpha} P(\alpha \leq X < \beta) = P(X = \alpha) = \lim_{\beta \rightarrow \alpha} [F(\beta) - F(\alpha)] = 0$$

Tức là nếu $F(x)$ liên tục thì $P(X = \alpha) = 0 \quad \forall \alpha$.

1.4. Hàm mật độ

1.4.1. Định nghĩa. Nếu lượng ngẫu nhiên X liên tục có hàm phân phối $F(x)$ liên tục đạo hàm của hàm phân phối $F(x)$ là $F'(x) = f(x)$ gọi là hàm mật độ của X .

1.4.2. Tính chất

$$\text{i)} \quad f(x) \geq 0 \quad \forall x$$

$$\text{ii)} \quad F(x) = \int_{-\infty}^x f(t) dt$$

$$\text{iii)} \quad P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x) dx$$

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

Ý nghĩa Tính chất iv) nói lên diện tích của hình phẳng giới hạn bởi trục Ox và đường cong hàm mật độ: $y = f(x)$ bằng 1. Tính chất ii) nói lên xác suất $P(\alpha < X < \beta)$ là diện tích hình thang cong giới hạn bởi các đường thẳng: $x = \alpha$, $x = \beta$ trục Ox và đường cong hàm mật độ $y = f(x)$.

Chú ý. Một hàm $y = f(x)$ thoả mãn 4 tính chất trên là hàm mật độ của một biến ngẫu nhiên X nào đó.

§2 CÁC ĐẶC TRƯNG CỦA BIẾN NGẪU NHIÊN

2.1. Số kỳ vọng

2.1.1. Định nghĩa

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

X	x_1	x_2		x_n
P	p_1	p_2		p_n

Số kỳ vọng của X là số kí hiệu $M(X)$ hoặc $E(X)$ và xác định như sau:

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

Nếu X là biến ngẫu nhiên liên tục có hàm mật độ $f(x)$ thì kỳ vọng của X là:

$$M(X) = \int_{-\infty}^{+\infty} xf(x)dx$$

Ví dụ 5. Cho X là biến ngẫu nhiên có dãy phân phối xác suất là:

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

Thì $M(X) = (1 \times 0,2) + (2 \times 0,5) + (3 \times 0,2) + (4 \times 0,1) = 2$.

Ví dụ 6. X là biến ngẫu nhiên phân phối đều trên $[a, b]$ thì hàm mật độ của X là:

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

$$\Rightarrow M(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$$

Ý nghĩa. Số kỳ vọng là số trung bình theo xác suất của biến ngẫu nhiên. Nếu xem $X_1, X_2, \dots, X_n$ là hệ chất điểm tại đó có đặt các khối lượng $p_1, p_2, \dots, p_n$ thì kỳ vọng chính là trọng tâm của hệ chất điểm.

2.1.2. Tính chất

- $M(C) = C$ (C là biến ngẫu nhiên hằng số)
- $M(CX) = CM(X)$ (C là hằng số)
- $M(X+Y) = M(X) + M(Y)$ (X, Y là hai biến ngẫu nhiên)
- X, Y là hai biến ngẫu nhiên độc lập nếu luật phân phối của X không phụ thuộc vào luật phân phối của Y và ngược lại thì $M(X.Y) = M(X).M(Y)$

2.2. Mốt và trung vị

2.2.1. Mốt của biến ngẫu nhiên ký hiệu $\text{Mod}(X)$ là trị số nhận được của biến ngẫu nhiên có xác suất cực đại (đối với biến ngẫu nhiên rời rạc) hay trị số có mật độ xác suất cực đại (đối với biến ngẫu nhiên liên tục).

2.2.2. Trung vị (median) của biến ngẫu nhiên X là số ký hiệu M_e sao cho:

$$P(X < M_e) = P(X > M_e)$$